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*

April 21, 2021

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.

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

Contents

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 [15] and the introduction books [14, 37]. Composition of flows formulas are also very useful in particular settings like Hamiltonian systems [17] or in the presence of a small perturbation of a reference flow [54].
- **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. [9, 11, 22, 25].
- **Differential equations on Lie groups.** Sometimes, the state itself of the differential equation belongs to a Lie group, as in [41]. 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 [44].

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

$$\dot{x}(t) = X(t)x(t), \tag{1.1}$$

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 Duhamel or Chen-Fliess expansion

One of the most straightforward approaches to solving (1.1) consists in what can be seen as an iterated application of Duhamel's principle. For small times, starting from the initial approximation $x(t) \approx x(0)$, one then enhances the approximation by plugging the approximation 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 [26, 33]. 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 2.16 and Remark 8.7 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 (1.1) has a deep link with the famous Campbell [23], Baker [8], Hausdorff [40], Dynkin [31] 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 [19], "chacun considère que les démonstrations de ses prédécesseurs ne sont pas convaincantes". We therefore encourage the reader to dive into the fascinating retrospectives [1] and [18] 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:

$$e^{X_2} e^{X_1} = \exp\left(X_1 + X_2 + \frac{1}{2}[X_2, X_1] + \dots \right). \tag{1.2}$$

When more than two exponentials are multiplied, say e^{X_1} through e^{X_n}, one can of course iterate the formula (1.2) 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 (1.1).

Magnus performed a breakthrough by deriving in [53] the first formal representation of the solution to (1.1) as the exponential of a series, of which the first terms are

$$x(t) = \exp\left(\int_0^t X(\tau_1)\, \mathrm{d}\tau_1 + \frac{1}{2} \int_0^t \int_0^{\tau_1} [X(\tau_1), X(\tau_2)]\, \mathrm{d}\tau_2\, \mathrm{d}\tau_1 + \cdots \right) x(0). \tag{1.3}$$

This formula can be seen as the continuous counterpart of the CBHD formula and highlights important structural properties of the solutions to (1.1) (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. [69]), often used for numerical splitting methods which attempts to give a meaning to the equality

$$e^{X_1 + X_2} = \lim_{n \to +\infty} \left(e^{\frac{X_1}{n}} e^{\frac{X_2}{n}} \right)^n, \tag{1.4}$$

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 [53], 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

$$e^{X_1 + X_2} = e^{X_1} e^{X_2} \exp\left(-\frac{1}{2}[X_1, X_2] \right) \exp\left(\frac{1}{3}[X_2, [X_1, X_2]] + \frac{1}{6}[X_1, [X_1, X_2]] \right) \cdots \tag{1.5}$$

In the context of differential equations such as (1.1), a nice formula is Sussmann's infinite product expansion, introduced in [68]. 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

$$x(t) = e^{\xi_1 X_1} e^{\xi_2 X_2} e^{\xi_{12} [X_1, X_2]} e^{\xi_{112} [X_1, [X_1, X_2]]} e^{\xi_{212} [X_2, [X_1, X_2]]} \cdots x(0), \tag{1.6}$$

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 [67, 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 [2, 3, 35] in which they derive an exponential representation of flows, very similar to Magnus' expansion, using the *chronological calculus* framework) and by Strichartz (see [65] 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 true objects (say matrices, operators or vector fields), 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 paper

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

- We give in Section 2 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 iterated Duhamel or 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 Section 2.4), 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 paper.
- 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 and for Sussmann's infinite product expansion. 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 nilpotent assumption. In this case, convergence requires analyticity, contrary to the proofs we give for the other expansions under a full nilpotent 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 Magnus expansions, which disprove the convergence of these expansions even for analytic vector fields and in very weak senses. We state an open problem concerning the convergence of Sussmann's infinite product for analytic vector fields.
- 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.
- 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**. 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. 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) and introduce a new mixed expansion which isolates the role of a perturbation in the dynamics.

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 [42, 60] for thorough introductions to Lie algebras and free Lie algebras.

Definition 2.1 (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 2.2 (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. 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 2.3 (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 [20, Chapter 3, Section 2.7, Definition 2]). $\mathcal{A}(X)$ can be seen as a graded algebra:*

$$\mathcal{A}(X) = \bigoplus_{n\in\mathbb{N}} \mathcal{A}_n(X), \tag{2.1}$$

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 2.4 (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 consider $\mathcal{L}(X)$, the* free Lie algebra *generated by X over the field $\mathbb{K}$, which is defined as the Lie subaglebra 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 [60, Theorem 0.4]). $\mathcal{L}(X)$ is a graded Lie algebra:*

$$\mathcal{L}(X) = \bigoplus_{n\in\mathbb{N}} \mathcal{L}_n(X), \qquad [\mathcal{L}_m(X), \mathcal{L}_n(X)] \subset \mathcal{L}_{m+n}(X) \tag{2.2}$$

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

Definition 2.5 (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 2.6 (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 2.5. Then the canonical surjection $\sigma_m : \mathcal{L}(X) \to \mathcal{N}_m(X)$ is a Lie algebra homomorphism.*

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

Lemma 2.7. *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 Formal brackets and evaluation

Definition 2.8 (Formal brackets)**.** *For X as above, we consider* $\mathrm{Br}(X)$ *the* set of formal 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 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 couple $(b_1, b_2) \in \mathrm{Br}(X)^2$ 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 $\mathrm{ad}_{b_1}(b_2)$ which allows iterated left 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 $n(b) = n_1(b) + \cdots + n_q(b) = |b| - n_0(b)$.

Definition 2.9 (Subspaces of brackets)**.** *When $I = [\![0, q]\!]$ and $\ell \in \mathbb{N}$, S_ℓ denotes the subset of* $\mathrm{Br}(X)$ *defined by*

$$S_\ell := \{b \in \mathrm{Br}(X);\ \ n(b) \leq \ell\}. \tag{2.3}$$

With this convention, S_ℓ is a subset of $\mathrm{Br}(X)$*, which is different from a convention commonly used in control theory where one refers to the vector space spanned by our S_ℓ in $\mathcal{L}(X)$.*

Remark 2.10. *There is a natural "evaluation" mapping* eval *from* $\mathrm{Br}(X)$ *to $\mathcal{L}(X)$ defined by induction by* $\mathrm{eval}(X_i) := X_i$ *for $X_i \in X$ and* $\mathrm{eval}([b_1, b_2]) := [\mathrm{eval}(b_1), \mathrm{eval}(b_2)]$*. Through this mapping,* $\mathrm{Br}(X)$ *spans $\mathcal{L}(X)$ over $\mathbb{K}$, i.e. $\mathcal{L}(X) = \mathrm{span}_{\mathbb{K}} \mathrm{eval}(\mathrm{Br}(X))$. 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 $\mathcal{L}(X)$. Nevertheless, we will sometimes implicitly evaluate the formal brackets of* $\mathrm{Br}(X)$ *in $\mathcal{L}(X)$, omitting the mapping* eval*.*

More precisely, the eval *map extends to a surjective algebra homomorphism from the nonassociative free algebra over X (which is the free vector space over* $\mathrm{Br}(X)$*, whose elements are formal (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* eval *is precisely the ideal generated by the relations that define anticommutativity and the Jacobi identity in $\mathcal{L}(X)$.*

2.1.3 Formal series, exponential and logarithms

Definition 2.11 (Formal series)**.** *We consider the (unital associative) algebra $\widehat{\mathcal{A}}(X)$ of formal 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 $S \in \widehat{\mathcal{A}}(X)$ and $\sigma \in I^*$, $\langle S, X_\sigma\rangle$ denotes the coefficient of X_σ in S: $S = \sum_{\sigma\in I^*} \langle S, X_\sigma\rangle X_\sigma$.*

Remark 2.12. *The definition of $\widehat{\mathcal{A}}(X)$ can be made more rigorous by considering $val : \mathcal{A}(X) \to \mathbb{N} \cup \{\infty\}$ by $val(a) = \inf\{n \in \mathbb{N} \mid a \in \bigoplus_{k\geq n} \mathcal{A}_k(X)\}$. Then $(a, b) \mapsto e^{-val(b-a)}$ 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. A formal series $\sum_{n\in\mathbb{N}} a_n$ with $a_n \in \mathcal{A}_n(X)$ thus converges in the metric space $\widehat{\mathcal{A}}(X)$, which justifies the notations of Definition 2.11. To avoid confusion, 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

$$\exp(a) := \sum_{m\geq 0} \frac{a^m}{m!}, \tag{2.4}$$

$$\log(1 + a) := \sum_{m\geq 1} \frac{(-1)^{m-1}}{m} a^m. \tag{2.5}$$

Since a has zero constant term, one checks that the right-hand sides of (2.4) and (2.5) indeed define formal series of $\widehat{\mathcal{A}}(X)$. In particular, $\log(\exp(a)) = a$ and $\exp(\log(1 + a)) = 1 + a$.

Lemma 2.13. *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 (2.4), one has:

$$(\exp(a))_r = \sum_{k=1}^{r} \sum_{r_1+\ldots r_k=r} \frac{a_{r_1} \ldots a_{r_k}}{k!} = a_r + \Theta_r\left(a_1, \ldots a_{r-1}\right), \tag{2.6}$$

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

$$a(t) := \sum_{i \in I} a_i(t) X_i. \tag{2.7}$$

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 $x^\star$,

$$\begin{cases} \dot{x}(t) = x(t)a(t), \\ x(0) = x^\star, \end{cases} \tag{2.8}$$

whose solutions are precisely defined in the following way.

Definition 2.14 (Solution to a formal ODE)**.** *Let $a_i \in L^1(\mathbb{R}_+; \mathbb{K})$ for $i \in I$ and define a by* (2.7)*. Let $x^\star \in \widehat{\mathcal{A}}(X)$ with homogeneous components $x_n^\star \in \mathcal{A}_n(X)$. The solution to the formal ODE* (2.8) *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 $n \in \mathbb{N}^*$,*

$$x_n(t) = x_n^\star + \int_0^t x_{n-1}(\tau) a(\tau) \, \mathrm{d}\tau. \tag{2.9}$$

Iterating this integral formula yields the following series expansion, which is the most direct way to compute the solution to (2.8) and was popularized by the works [26, 33].

Lemma 2.15 (Iterated Duhamel or Chen-Fliess series expansion)**.** *In the context of Definition 2.14, the solution to* (2.8) *with initial data $x^\star = 1$ can be expanded as*

$$x(t) = \sum_{\sigma \in I^*} \left(\int_0^t a_\sigma \right) X_\sigma, \tag{2.10}$$

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

$$\int_0^t a_\sigma := \int_{0 < \tau_1 < \cdots < \tau_n < t} a_{\sigma_1}(\tau_1) \cdots a_{\sigma_n}(\tau_n) \, \mathrm{d}\tau. \tag{2.11}$$

Proof. Expansion (2.10) is a direct consequence of the iterated application of (2.9) and of the definition of X_σ in Definition 2.2 and can be proved by induction on the length of σ. □

Remark 2.16. *Despite its simplicity, the Chen-Fliess series expansion 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* (2.8) *and every $t \geq 0$, one has the identity*

$$\langle x(t), X_1 X_2 \rangle + \langle x(t), X_2 X_1 \rangle - \langle x(t), X_1 \rangle \langle x(t), X_2 \rangle = 0 \tag{2.12}$$

- *In the context of nonlinear ordinary differential equations, the representation* (2.10) *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* (2.10) *will not be invariant by diffeomorphism (see Remark 8.7), 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.

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 (2.11) yields the elegant formula $x(t) = x^\star \exp(ta)$, with the notation of (2.4). Of course, it is no longer valid for a time-varying dynamic (because the indeterminates do not commute *a priori*), but one can wish to find an object of which the flow is the exponential, the so-called "*logarithm of the flow*". In this section, we recall and prove Theorem 2.26, which states that the logarithm of flows of formal linear differential equations is given by explicit Lie brackets. Our proof follows the method proposed in [65, Section 3] and relies on well-known algebraic results, which we recall, for the sake of giving a self-contained presentation. Another related approach, relying on Ree's theorem and shuffle relations satisfied by the Chen-Fliess series coefficients is developed in [47, 48, 59, 60].

2.3.1 Notations for indexes

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

Definition 2.17. *For $m, r \in \mathbb{N}^*$, we define the set of ordered positive partitions of size m of r,*

$$\mathbb{N}_r^m := \{\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_m) \in (\mathbb{N}^*)^m;\ \ \mathbf{r}_1 + \cdots + \mathbf{r}_m = r\}\,, \tag{2.13}$$

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

$$\mathcal{T}_{\mathbf{r}}(t) := \Big\{\tau = (\tau_1, \dots, \tau_r) \in (0,t)^r;\ \ \forall j \in [\![1, m]\!],\ \ 0 < \tau_{R_j} < \cdots < \tau_{R_{j-1}+1} < t\Big\}, \tag{2.14}$$

where, for $j \in [\![1, m]\!]$,

$$R_j := \sum_{i=1}^{j} \mathbf{r}_i. \tag{2.15}$$

Example 2.18. *The sets $\mathcal{T}_{\mathbf{r}}(t)$ will be used as integration domains, and can be pictured as products of pyramidal domains. As examples, we compute the integration domains for $r \leq 3$. One has*

$$\mathcal{T}_{(1)}(t) = \{\tau = (\tau_1) \in (0,t)^1\}, \tag{2.16}$$
$$\mathcal{T}_{(2)}(t) = \{\tau = (\tau_1, \tau_2) \in (0,t)^2;\ \ 0 < \tau_2 < \tau_1 < t\}, \tag{2.17}$$
$$\mathcal{T}_{(1,1)}(t) = \{\tau = (\tau_1, \tau_2) \in (0,t)^2\}, \tag{2.18}$$
$$\mathcal{T}_{(3)}(t) = \{\tau = (\tau_1, \tau_2, \tau_3) \in (0,t)^3;\ \ 0 < \tau_3 < \tau_2 < \tau_1 < t\}, \tag{2.19}$$
$$\mathcal{T}_{(2,1)}(t) = \{\tau = (\tau_1, \tau_2, \tau_3) \in (0,t)^3;\ \ 0 < \tau_2 < \tau_1 < t\}, \tag{2.20}$$
$$\mathcal{T}_{(1,2)}(t) = \{\tau = (\tau_1, \tau_2, \tau_3) \in (0,t)^3;\ \ 0 < \tau_3 < \tau_2 < t\}, \tag{2.21}$$
$$\mathcal{T}_{(1,1,1)}(t) = \{\tau = (\tau_1, \tau_2, \tau_3) \in (0,t)^3\}. \tag{2.22}$$

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

$$\mathcal{T}_{(2,2)}(t) = \{\tau = (\tau_1, \tau_2, \tau_3, \tau_4) \in (0,t)^4;\ \ 0 < \tau_2 < \tau_1 < t \text{ and } 0 < \tau_4 < \tau_3 < t\}. \tag{2.23}$$

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 2.19 (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*

$$\mathrm{Log}_M\{F\}(t) := \sum_{r=1}^{M} \frac{1}{r} \sum_{m=1}^{r} \frac{(-1)^{m-1}}{m} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \int_{\tau \in \mathcal{T}_{\mathbf{r}}(t)} [\cdots [F(\tau_r), F(\tau_{r-1})], \dots F(\tau_1)]\, \mathrm{d}\tau. \tag{2.24}$$

Remark 2.20. *In such an abstract setting, the right-hand side of* (2.24) *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* (2.24) *as an abstract formula or notation. We will check, each time we use it, that we can give a meaning to the integrals.*

2.3.2 Preliminary algebraic results

Define a 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 $1 \leq i \leq q$, and, for $1 \leq i_1, \ldots, i_k \leq q$ with $k \in \mathbb{N}^*$,

$$\beta(X_{i_1} X_{i_2} \cdots X_{i_k}) := [\cdots [X_{i_1}, X_{i_2}], \ldots, X_{i_k}]. \tag{2.25}$$

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 [30], Specht [63] and Wever [72] states that, if a polynomial is a Lie element, then it is equal to its left normed bracketing.

Lemma 2.21 (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 [60, Theorem 1.4]. □

Example 2.22. *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, $\beta([X_1, X_2]) = (X_1X_2 - X_2X_1) - (X_2X_1 - X_1X_2) = 2[X_1, X_2]$.*

Let $\mathcal{A}(X) \otimes \mathcal{A}(X)$ be the tensor product of 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 [20, Chapter 3, Section 4.1, Definition 1] for a precise construction). Define an homomorphism Δ from $\mathcal{A}(X)$ to $\mathcal{A}(X) \otimes \mathcal{A}(X)$ 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 [60, Proposition 1.2] for more detail). The homomorphism Δ can then be used to characterize Lie elements, as in the following result, which was proposed by Friedrichs in [34], then proved by multiple authors in the same period [27, 32, 52, 53].

Lemma 2.23 (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 [60, Theorem 1.4]. □

Example 2.24. *The element X_1X_2 does not belong to $\mathcal{L}$. And indeed,*

$$\begin{aligned} \Delta(X_1X_2) = \Delta(X_1)\Delta(X_2) &= (X_1 \otimes 1 + 1 \otimes X_1)(X_2 \otimes 1 + 1 \otimes X_2) \\ &= X_1X_2 \otimes 1 + X_1 \otimes X_2 + X_2 \otimes X_1 + 1 \otimes X_1X_2 \\ &\neq X_1X_2 \otimes 1 + 1 \otimes X_1X_2. \end{aligned} \tag{2.26}$$

On the contrary, the element $[X_1, X_2] = X_1X_2 - X_2X_1$ belongs to $\mathcal{L}$. And indeed,

$$\begin{aligned} \Delta([X_1, X_2]) &= \Delta(X_1X_2) - \Delta(X_2X_1) \\ &= (X_1X_2 \otimes 1 + X_1 \otimes X_2 + X_2 \otimes X_1 + 1 \otimes X_1X_2) \\ &\quad - (X_2X_1 \otimes 1 + X_2 \otimes X_1 + X_1 \otimes X_2 + 1 \otimes X_2X_1) \\ &= [X_1, X_2] \otimes 1 + 1 \otimes [X_1, X_2]. \end{aligned} \tag{2.27}$$

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 series over $\mathcal{A}(X) \otimes \mathcal{A}(X)$. For such series with zero constant term, one can define, as in (2.4), an exponential, say $\exp_{\otimes}$, which also verifies a uniqueness property such as Lemma 2.13. One can then derive a criterion to determine whether the logarithm of a formal series is a Lie element.

Corollary 2.25. *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 [60, Theorem 3.2]. By linearity and degree preservation, Lemma 2.23 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,

$$\begin{aligned} \log a \in \widehat{\mathcal{A}}(X) &\iff \Delta(\log(a)) = \log(a) \otimes 1 + 1 \otimes \log(a) \\ &\iff \exp_\otimes(\Delta(\log(a))) = \exp_\otimes(\log(a) \otimes 1 + 1 \otimes \log(a)) \\ &\iff \Delta(\exp(\log(a))) = \exp_\otimes(\log(a) \otimes 1) \exp_\otimes(1 \otimes \log(a)) \\ &\iff \Delta(a) = ((\exp \log a) \otimes 1)(1 \otimes (\exp \log a)) = a \otimes a, \end{aligned} \tag{2.28}$$

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. □

2.3.3 Formal linear differential equations

Theorem 2.26. *For $t \in \mathbb{R}_+$ and $x^\star \in \widehat{\mathcal{A}}(X)$, the solution x to (2.8) satisfies*

$$x(t) = x^\star \exp\left(\mathrm{Log}_\infty\{a\}(t)\right), \tag{2.29}$$

with the notation of Definition 2.19.

Proof. First, by linearity, it suffices to prove (2.29) for $x^\star = 1$. For $t \in \mathbb{R}_+$, to show that $\log(x(t))$ is a Lie series, thanks to Corollary 2.25, it suffices to check that $\Delta(x(t)) = x(t) \otimes x(t)$. We proceed using the same trick as in [65]. At the initial time $\Delta(x(0)) = \Delta(1) = 1 \otimes 1 = x(0) \otimes x(0)$. Then, on the one hand

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta(x) = \Delta(\dot{x}) = \Delta(xa) = \Delta(x)\Delta(a) = \Delta(x)(a \otimes 1 + 1 \otimes a). \tag{2.30}$$

On the other hand,

$$\frac{\mathrm{d}}{\mathrm{d}t}(x \otimes x) = \dot{x} \otimes x + x \otimes \dot{x} = (xa) \otimes x + x \otimes (xa) = (x \otimes x)(a \otimes 1 + 1 \otimes a). \tag{2.31}$$

Hence, both quantities satisfy the same formal differential equation with the same initial condition, so they are equal for every $t \in \mathbb{R}_+$ and $\log(x(t)) \in \widehat{\mathcal{L}}(X)$.

Repeated integration of (2.9) yields, for every $t \in \mathbb{R}_+$,

$$x(t) = 1 + \sum_{r \geq 1} \int_{0 < \tau_r < \cdots < \tau_1 < t} a(\tau_r) \cdots a(\tau_1) \, \mathrm{d}\tau. \tag{2.32}$$

Hence, recalling the definitions (2.13) of $\mathbb{N}_r^m$ and (2.14) of $\mathcal{T}_{\mathbf{r}}(t)$, one has

$$\log(x(t)) = \sum_{r=1}^{+\infty} \sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\mathcal{T}_{\mathbf{r}}(t)} a(\tau_r) a(\tau_{r-1}) \cdots a(\tau_1) \, \mathrm{d}\tau. \tag{2.33}$$

Since $\log(x(t)) \in \widehat{\mathcal{L}}(X)$, applying Lemma 2.21 to each of its homogeneous components in $\mathcal{A}_r$ proves

$$\log(x(t)) = \sum_{r=1}^{+\infty} \frac{1}{r} \sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\mathcal{T}_{\mathbf{r}}(t)} [\cdots [a(\tau_r), a(\tau_{r-1})], \ldots a(\tau_1)] \, \mathrm{d}\tau. \tag{2.34}$$

Recalling the notation (2.19) and taking the exponential concludes the proof of (2.29). □

2.3.4 Coordinates of the first kind

Although the expansion (2.34) 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 2.27 (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 *when the elements of $\mathcal{B}$ are the evaluation of formal brackets in $\mathrm{Br}(X)$, that will be identified. Then, for every $n \in \mathbb{N}^*$, we use the following notations $\mathcal{B}_n = \{b \in \mathcal{B}; |b| = n\}$ and $\mathcal{B}_{[\![1,n]\!]} = \{b \in \mathcal{B}; |b| \leq n\}$.*

Proposition 2.28. *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 (2.8) satisfies*

$$x(t) = x^\star \exp\left(\sum_{b\in\mathcal{B}} \zeta_b(t,a)b\right). \tag{2.35}$$

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 (2.34) 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 (2.35) is understood in the sense of a well-defined formal 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 2.29 (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 $\mathcal{L}(X)$.*

Remark 2.30. *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*

$$\mathrm{Br}_\lambda(X) := \{b \in \mathrm{Br}(X);\ \ \forall i \in I, n_i(b) = \lambda_i\}. \tag{2.36}$$

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 2.31. *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.*

2.3.5 Campbell Baker Hausdorff Dynkin formula

Corollary 2.32. *Let X be a finite set, $n \in \mathbb{N}^*$ and $y_1, \ldots, y_n \in \widehat{\mathcal{L}}(X)$ without constant term. There exists a unique $w \in \widehat{\mathcal{L}}(X)$ such that*

$$e^{y_1} \cdots e^{y_n} = e^w. \tag{2.37}$$

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

$$\mathrm{CBHD}_\infty(y_1, \ldots, y_n) = \sum_{b\in\mathcal{B}} \alpha_b y_b \tag{2.38}$$

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

Proof. We prove that (2.37) holds with

$$w := \mathrm{Log}_\infty \left\{ \sum_{j=1}^{n} y_j 1_{[j-1,j]} \right\}(n) \tag{2.39}$$

in the sense of Definition 2.19.

Step 1: Proof when $X = \{X_1, \dots, X_n\}$ and $y_j = X_j$ for $j \in [\![1, n]\!]$. The solution to (2.8) 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 2.26, w solves (2.37). By injectivity of the exponential (see Lemma 2.13), it is the unique solution. By Proposition 2.28, the equality (2.38) 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 algebra homomorphism $\widehat{\mathcal{A}}(Y) \to \widehat{\mathcal{A}}(X)$, which is also a Lie algebra homomorphism $\widehat{\mathcal{L}}(Y) \to \widehat{\mathcal{L}}(X)$, that we still denote Λ. Indeed Lemma 2.7 ensures the extension as an algebra homomorphism $\mathcal{A}(Y) \to \widehat{\mathcal{A}}(X)$ (resp. a Lie algebra homomorphism $\mathcal{L}(Y) \to \widehat{\mathcal{L}}(X)$). 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 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 algebra homomorphism Λ to the relation $e^{Y_1} \cdots e^{Y_n} = e^W$ we get (2.37). By applying the Lie algebra homomorphism Λ to the relation $W = \sum_{b \in \mathcal{B}} \alpha_b b$ we get (2.38). □

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 2.33. *There exists a family of maps $F_{q,h} : \widehat{\mathcal{L}}(X)^q \to \widehat{\mathcal{L}}(X)$ for $q \in \mathbb{N}^*$ and $h = (h_1, \dots, h_q) \in (\mathbb{N}^*)^q$ such that*

- *$F_{q,h}(y_1, \dots, y_q)$ is a linear combination of brackets of $y_1, \dots, y_q$ involving y_j exactly h_j times, for $j \in [\![1, q]\!]$,*
- *for every $n \geq 2$, $y_1, \dots, y_n \in \widehat{\mathcal{L}}(X)$,*

$$\mathrm{CBHD}_\infty(y_1, \dots, y_n) = \sum_{\substack{q \in [\![1,n]\!] \\ j_1 < \cdots < j_q \in [\![1,n]\!] \\ h = (h_1, \dots, h_q) \in (\mathbb{N}^*)^q}} F_{q,h}(y_{j_1}, \dots, y_{j_q}). \tag{2.40}$$

For $q = 1$, $F_{1,(1)}(y_1) = y_1$ and $F_{1,(h_1)}(y_1) = 0$ for $h_1 \geq 2$. For $q = 2$ and $h_1 + h_2 \leq 4$,

$$\begin{aligned}
F_{2,(1,1)}(y_1, y_2) &= \frac{1}{2}[y_1, y_2] & F_{2,(2,2)}(y_1, y_2) &= -\frac{1}{24}[y_2, [y_1, [y_1, y_2]]] \\
F_{2,(2,1)}(y_1, y_2) &= \frac{1}{12}[y_1, [y_1, y_2]] & F_{2,(3,1)}(y_1, y_2) &= 0 \\
F_{2,(1,2)}(y_1, y_2) &= \frac{1}{12}[y_2, [y_2, y_1]] & F_{2,(1,3)}(y_1, y_2) &= 0.
\end{aligned} \tag{2.41}$$

For higher order terms, there is no elegant general closed form.

Proof. Using the same Lie algebra homomorphism arguments as in the proof of Corollary 2.32, 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 (2.38) where the terms are grouped by their homogeneity with respect to y_1 and y_2. This defines the maps $F_{1,(1)}(y_1) = y_1$ 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 (2.41).

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

$$F_{n,h}(y_1, \dots, y_n) := \sum_{m | h_1, \dots, h_{n-1}} F_{2,(m,h_n)}\big(F_{n-1,(\frac{h_1}{m}, \dots, \frac{h_{n-1}}{m})}(y_1, \dots, y_{n-1}), y_n\big). \tag{2.42}$$

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

$$\begin{aligned}
&\mathrm{CBHD}_\infty(y_1, \dots, y_n) \\
&= \mathrm{CBHD}_\infty(\mathrm{CBHD}_\infty(y_1, \dots, y_{n-1}), y_n) \\
&= \mathrm{CBHD}_\infty(y_1, \dots, y_{n-1}) + y_n + \sum_{g \in (\mathbb{N}^*)^2} F_{2,g}(\mathrm{CBHD}_\infty(y_1, \dots, y_{n-1}), y_n) \\
&= y_n + \sum_{\substack{q \in [\![1, n-1]\!] \\ j_1 < \dots < j_q \in [\![1, n-1]\!] \\ h' \in (\mathbb{N}^*)^q}} \left(F_{q,h'}(y_{j_1}, \dots, y_{j_q}) + \sum_{g \in (\mathbb{N}^*)^2} F_{2,g}(F_{q,h'}(y_{j_1}, \dots, y_{j_q}), y_n) \right).
\end{aligned} \tag{2.43}$$

We now check that the right-hand side of (2.43) is the same as the right-hand side of (2.40). 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 $q \in [\![1, n-1]\!]$, $j_1 < \dots < j_q \in [\![1, n-1]\!]$ 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 (2.40). In (2.43), it is

$$\sum_{h' \in (\mathbb{N}^*)^q} \sum_{g \in (\mathbb{N}^*)^2} F_{2,g}(F_{q,h'}(y_{j_1}, \dots, y_{j_q}), y_n), \tag{2.44}$$

  where the sum is restricted to $g_1 h'_i = h_i$ and $g_2 = h_{q+1}$. Hence, both terms are equal thanks to the definition (2.42).

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

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.

2.3.6 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. [46] where the first 14 such coordinates are computed, and [24, 57] 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

$$W_1 := \mathrm{span}_{\mathbb{K}} \{\mathrm{eval}(b);\ \ b \in \mathrm{Br}(X), n_1(b) = 1\} \subset \mathcal{L}(X). \tag{2.45}$$

Lemma 2.34. *The family* $(\mathrm{ad}^k_{X_0}(X_1))_{k \in \mathbb{N}}$ *is a basis of* W_1.

Proof. From (2.45), W_1 is spanned by the evaluations in $\mathcal{L}(X)$ of the formal brackets $B \in \mathrm{Br}(X)$ involving X_1 exactly once. Let $B \in \mathrm{Br}(X)$ be such a formal bracket. We assume $\mathrm{eval}(B) \neq 0$ in $\mathcal{L}(X)$ and $B \neq X_1$. There exists a unique couple $(B', B'') \in \mathrm{Br}(X)^2$ such that $B = [B', B'']$. Then $\mathrm{eval}(B) = [\mathrm{eval}(B'), \mathrm{eval}(B'')]$ thus $\mathrm{eval}(B')$ and $\mathrm{eval}(B'')$ are non null in $\mathcal{L}(X)$. Moreover, either B' or B'' does not involve X_1 and is thus equal to X_0. Therefore $\mathrm{eval}(B) = \pm[X_0, \mathrm{eval}(\overline{B})]$ where $\overline{B} \in \mathrm{Br}(X)$ involves X_1 exactly once and $\mathrm{eval}(\overline{B}) \neq 0$. Working by induction on the number k of occurrences of X_0 in B, we obtain $\mathrm{eval}(B) = \pm \operatorname{ad}^k_{X_0}(X_1)$.

The previous argument proves that the given family spans W_1. Moreover, this family is linearly independent in $\mathcal{L}(X)$ because two different elements have different lengths. □

To express the coordinates of the first kind on this basis, we introduce (using the modern NIST sign and indexing convention) the Bernoulli numbers $(B_n)_{n\in\mathbb{N}}$, defined by the identity

$$\forall z \in \mathbb{C}, |z| < 2\pi, \qquad \frac{z}{e^z - 1} = \sum_{n=0}^{+\infty} B_n \frac{z^n}{n!} = 1 - \frac{z}{2} + \sum_{n=1}^{+\infty} B_{2n} \frac{z^{2n}}{(2n)!}. \tag{2.46}$$

Moreover, we introduce

$$W_{2^+} := \mathrm{span}_{\mathbb{K}} \{\mathrm{eval}(b);\ \ b \in \mathrm{Br}(X), n_1(b) \geq 2\} \subset \mathcal{L}(X), \tag{2.47}$$

thanks to which we can write the direct sum decomposition $\mathcal{L}(X) = \mathbb{K}X_0 \oplus W_1 \oplus W_{2^+}$.

Proposition 2.35. *Let $\mathcal{B}$ a monomial basis of $\mathcal{L}(X)$ containing X_0 and the family $(\operatorname{ad}^k_{X_0}(X_1))_{k\in\mathbb{N}}$. The associated coordinates of the first kind satisfy, for each $t > 0$, $a_0, a_1 \in L^1((0,t);\mathbb{K})$ and $k \in \mathbb{N}$,*

$$\begin{aligned} \zeta_{\operatorname{ad}^k_{X_0}(X_1)}(t, a_0, a_1) &= (-1)^k A_0(t)^k \frac{B_k}{k!} A_1(t) \\ &\quad + (-1)^k \sum_{\ell=1}^{k} A_0(t)^{k-\ell} \frac{B_{k-\ell}}{(k-\ell)!} \int_{0<\tau_\ell<\cdots<\tau_1<t} a_0(\tau_1)\cdots a_0(\tau_\ell) A_1(\tau_\ell)\,\mathrm{d}\tau, \end{aligned} \tag{2.48}$$

where $A_0(t) := \int_0^t a_0$ and $A_1(t) := \int_0^t a_1$ and the sum is empty by convention for $k = 0$.

Proof. First, the considered coordinates are well-defined independently on the exact choice of $\mathcal{B}$ (see Remark 2.30). Let x be the solution to (2.8) starting from $x^\star = 1$. To simplify the notations in this proof, we write $x(t)$, $\zeta_k(t)$ and $Z(t)$ instead of $x(t,a)$, $\zeta_{\operatorname{ad}^k_{X_0}(X_1)}(t, a_0, a_1)$ and $\mathrm{Log}_\infty\{a\}(t)$. From (2.35),

$$Z(t) = \sum_{b\in\mathcal{B}} \zeta_b(t,a) b = \zeta_{X_0}(t,a) X_0 + Z_1(t) + Z_2(t), \tag{2.49}$$

where $Z_2(t) \in W_{2^+}$ and

$$Z_1(t) := \sum_{k=0}^{+\infty} \zeta_k(t) \operatorname{ad}^k_{X_0}(X_1). \tag{2.50}$$

First, a straightforward identification in (2.24) yields $\zeta_{X_0} = A_0$ and $\zeta_{X_1} = A_1$. Let $k \in \mathbb{N}^*$. The proof consists in computing $\langle x(t), X_1 X_0^k \rangle$ in two ways: first by the ODE (2.8), then by the formula $x(t) = e^{Z(t)}$. By definition of the solution to (2.8), we have, for every word $\sigma \in I^*$ and $t > 0$

$$\langle x(t), X_\sigma X_0 \rangle = \int_0^t \langle x(\tau), X_\sigma \rangle a_0(\tau)\,\mathrm{d}\tau. \tag{2.51}$$

Taking into account that $\langle x(t), X_1 \rangle = A_1(t)$, we obtain

$$\langle x(t), X_1 X_0^k \rangle = \int_{0<\tau_k<\cdots<\tau_1<t} a_0(\tau_1)\cdots a_0(\tau_k) A_1(\tau_k)\,\mathrm{d}\tau. \tag{2.52}$$

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

$$\left\langle x(t), X_1 X_0^k \right\rangle = \left\langle Z(t), X_1 X_0^k \right\rangle + \sum_{\ell=2}^{k+1} \frac{1}{\ell!} \left\langle Z(t)^\ell, X_1 X_0^k \right\rangle \tag{2.53}$$

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

$$Z(t)^\ell = \sum_{j=0}^{\ell-1} (A_0(t)X_0)^j Z_1(t) (A_0(t)X_0)^{\ell-1-j} + Z_{2,\ell}(t), \quad \text{where} \quad Z_{2,\ell}(t) \in W_{2^+}. \tag{2.54}$$

Thus

$$\left\langle Z(t)^\ell, X_1 X_0^k \right\rangle = \left\langle Z_1(t)(A_0(t)X_0)^{\ell-1}, X_1 X_0^k \right\rangle = A_0(t)^{\ell-1} (-1)^{k-\ell+1} \zeta_{k-\ell+1}(t), \tag{2.55}$$

because the word $X_1 X_0^{k-\ell+1}$ appears in the decomposition of $\mathrm{ad}_{X_0}^n(X_1)$ iff $k - \ell + 1 = n$ and then it appears with coefficient $(-1)^n$. We deduce from (2.53) and (2.55) that

$$\left\langle x(t), X_1 X_0^k \right\rangle = (-1)^k \zeta_k(t) + \sum_{\ell=2}^{k+1} \frac{(-1)^{k+1-\ell}}{\ell!} A_0(t)^{\ell-1} \zeta_{k+1-\ell}(t). \tag{2.56}$$

Using (2.52) and the index change $j = k + 1 - \ell \in [\![0, k-1]\!]$, we obtain

$$\int_{0<\tau_k<\cdots<\tau_1<t} a_0(\tau_1) \cdots a_0(\tau_k) A_1(\tau_k) \, \mathrm{d}\tau = (-1)^k \zeta_k(t) + \sum_{j=0}^{k-1} \frac{(-1)^j A_0(t)^{k-j}}{(k+1-j)!} \zeta_j(t), \tag{2.57}$$

When $A_0(t) = 0$, this formula yields (2.48) immediately. When $A_0(t) \neq 0$, let, for $j \in \mathbb{N}$,

$$\alpha_j := \frac{\langle x(t), X_1 X_0^j \rangle}{A_0(t)^{j+1}} \qquad \text{and} \qquad \beta_j := \frac{(-1)^j \zeta_j(t)}{A_0(t)^{j+1}} \tag{2.58}$$

we deduce from (2.57) that

$$\alpha_k = \sum_{j=0}^{k} \frac{\beta_j}{(k+1-j)!}. \tag{2.59}$$

We have

$$z \left(\sum_{k \geq 0} \alpha_k z^k \right) = \sum_{k \geq 0} \sum_{j=0}^{k} \beta_j z^j \frac{z^{k+1-j}}{(k+1-j)!} = \left(\sum_{j \geq 0} \beta_j z^j \right) (e^z - 1) \tag{2.60}$$

or equivalently

$$\sum_{j \geq 0} \beta_j z^j = \frac{z}{e^z - 1} \left(\sum_{k \geq 0} \alpha_k z^k \right) = \sum_{n \geq 0} \sum_{k \geq 0} B_n \frac{z^n}{n!} \alpha_k z^k. \tag{2.61}$$

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

$$\beta_j = \sum_{k=0}^{j} \frac{B_{j-k}}{(j-k)!} \alpha_k. \tag{2.62}$$

Finally (2.58) and (2.52) give (2.48). □

In particular, we recover the following very classical formula for the partial coefficients of the CBHD formula (see e.g. [71, equation (2)] or [60, Corollary 3.24]).

Corollary 2.36. *There holds $e^{X_1}e^{X_0} = e^Z$ where $Z = X_0 + Z_1 + Z_2$, $Z_2 \in W_{2^+}$ and*

$$Z_1 := \sum_{n=0}^{+\infty} \frac{B_n}{n!} \operatorname{ad}_{X_0}^n(X_1) = X_1 - \frac{1}{2}[X_0, X_1] + \sum_{n=1}^{+\infty} \frac{B_{2n}}{(2n)!} \operatorname{ad}_{X_0}^{2n}(X_1). \tag{2.63}$$

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 (2.8) with $x^\star = 1$ satisfies $x(2) = e^{X_1}e^{X_0}$. For $\ell \in \mathbb{N}^*$ and $0 < \tau_\ell < \cdots < \tau_1 < 2$, the real number $a_0(\tau_1)\cdots a_0(\tau_\ell)A_1(\tau_\ell)$ does not vanish iff $1 < \tau_\ell < \cdots \tau_1 < 2$ and then it equals 1. Thus, for every $k \geq 2$, using (2.48) and (2.65),

$$(-1)^k \zeta_k(2) = \sum_{\ell=0}^{k} \frac{B_{k-\ell}}{(k-\ell)!}\frac{1}{\ell!} = \sum_{j=0}^{k} \frac{B_j}{j!(k-j)!} = \frac{B_k}{k!} \tag{2.64}$$

We conclude by noticing, thanks to (2.48), that $\zeta_0(2) = A_1(2) = 1$ and $\zeta_1(2) = -\frac{1}{2} = B_1$. □

The following lemma states properties about Bernoulli numbers (and thus about coordinates of the first kind) that will be used later in this article.

Lemma 2.37. *The Bernoulli numbers defined in* (2.46) *satisfy, for every $n \geq 2$*

$$\sum_{k=0}^{n-1} \binom{n}{k} B_k = 0, \tag{2.65}$$

$$\sum_{k=0}^{n} \binom{n}{k} \frac{B_k}{n+1-k} = 0. \tag{2.66}$$

Moreover, the odd Bernoulli numbers except B_1 vanish and, for every $n \geq 1$,

$$B_{2n} = (-1)^{n+1}\frac{2(2n)!}{(2\pi)^{2n}}\zeta(2n) \sim (-1)^{n+1} 2\sqrt{2\pi n}\left(\frac{n}{\pi}\right)^{2n}, \tag{2.67}$$

where ζ is the Riemann zeta function.

Proof. Both sum equalities are classical and can be proved using the generating series of the Bernoulli numbers of (2.46), respectively by identification in $z = (e^z - 1) \times (z/(e^z - 1))$ for (2.65) and in $1 = ((e^z - 1)/z) \times (z/(e^z - 1))$ for (2.66).

The relationship with the Riemann zeta function is proved in [7, 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 $\zeta(s) = \sum n^{-s}$). □

Example 2.38. *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* (2.48) *of Proposition 2.35 we obtain*

$$\begin{aligned}
\zeta_{\operatorname{ad}_{X_0}^k(X_1)}(t,a) &= (-1)^k t^k \frac{B_k}{k!}\frac{t^2}{2} + (-1)^k \sum_{\ell=1}^{k} t^{k-\ell}\frac{B_{k-\ell}}{(k-\ell)!}\frac{t^{\ell+2}}{(\ell+2)!} \\
&= (-1)^k t^{k+2} \sum_{j=0}^{k} \frac{B_{k-j}}{(k-j)!(j+2)!} \\
&= (-1)^{k+1} t^{k+2} \frac{B_{k+1}}{(k+1)!},
\end{aligned} \tag{2.68}$$

where we used that

$$\begin{aligned}\sum_{j=0}^{k}\frac{B_{k-j}}{(k-j)!(j+2)!}&=\frac{1}{(k+1)!}\sum_{\ell=0}^{k}\binom{k}{k-\ell}\frac{k+1}{k+1-\ell}\frac{B_\ell}{k-\ell+2}\\&=\frac{1}{(k+1)!}\sum_{\ell=0}^{k}\binom{k+1}{\ell}\frac{B_\ell}{(k+1)-\ell+1}=-\frac{B_{k+1}}{(k+1)!},\end{aligned}\tag{2.69}$$

thanks to (2.66).

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). It 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

$$a(t)=X_0+\sum_{i=1}^{q}a_i(t)X_i.\tag{2.70}$$

Theorem 2.39. *For $t\in\mathbb{R}_+$, $x^\star\in\widehat{\mathcal{A}}(X)$ and a of the form* (2.70)*, the solution x to* (2.8) *satisfies*

$$x(t)=x^\star\exp(tX_0)\exp\left(\mathcal{Z}_\infty(t,X,a)\right),\tag{2.71}$$

where $\mathcal{Z}_\infty(t,X,a):=\mathrm{Log}_\infty\{b_t\}(t)$ with the notation of Definition 2.19 and

$$b_t(s):=e^{-(t-s)X_0}\left(\sum_{i=1}^{q}a_i(s)X_i\right)e^{(t-s)X_0}=\sum_{i=1}^{q}\sum_{k=0}^{+\infty}\frac{(-1)^k}{k!}(t-s)^k a_i(s)\,\mathrm{ad}_{X_0}^k(X_i)\tag{2.72}$$

i.e.

$$\begin{aligned}\mathcal{Z}_\infty(t,X,a)=\sum\frac{(-1)^{m-1}}{mr}\int_{\tau\in\mathcal{T}_{\mathbf{r}}(t)}&\frac{(\tau_r-t)^{k_r}}{k_r!}\cdots\frac{(\tau_1-t)^{k_1}}{k_1!}a_{i_r}(\tau_r)\cdots a_{i_1}(\tau_1)\,\mathrm{d}\tau\\&[\cdots[\mathrm{ad}_{X_0}^{k_r}(X_{i_r}),\mathrm{ad}_{X_0}^{k_{r-1}}(X_{i_{r-1}})],\ldots\mathrm{ad}_{X_0}^{k_1}(X_{i_1})]\end{aligned}\tag{2.73}$$

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 $i_1,\ldots,i_r\in[\![1,q]\!]$.

Proof. 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 unknowns $\{Y_{k,i}\}_{k\in\mathbb{N},i\in[\![1,q]\!]}$, the solution to

$$\dot z(s)=z(s)\gamma_t(s)\quad\text{where}\quad\gamma_t(s):=\sum_{k,i}\frac{(-1)^k}{k!}(t-s)^k a_i(s)Y_{k,i},\tag{2.74}$$

with initial data $z(0) = 1$ satisfies, thanks to Theorem 2.26,

$$z(t) = \exp\left(\mathrm{Log}_\infty\{\gamma_t\}(t)\right). \tag{2.75}$$

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

$$\Theta(Y_{k,i}) := \mathrm{ad}^k_{X_0}(X_i). \tag{2.76}$$

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 unknown $y(s) := x(s)e^{(t-s)X_0}$. Then,

$$\dot{y}(s) = \dot{x}(s)e^{(t-s)X_0} - x(s)X_0e^{(t-s)X_0} = x(s)\left(\sum_{i=1}^q a_i(s)X_i\right)e^{(t-s)X_0} = y(s)b_t(s). \tag{2.77}$$

Hence

$$x(t) = y(t) = y(0)z_\Theta(t) = x^\star e^{tX_0}\exp\left(\mathrm{Log}_\infty\{b_t\}(t)\right), \tag{2.78}$$

which concludes the proof of (2.71). □

Remark 2.40. *In expansion* (2.71)*, 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 2.41. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$ 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 $t \geq 0$*

$$\mathcal{Z}_\infty(t, X, a) = \sum_{b\in\mathcal{B}} \eta_b(t, a)b \quad \text{in} \quad \widehat{\mathcal{L}}(X). \tag{2.79}$$

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

$$\begin{aligned}\mathcal{Z}^{r,\nu}_\infty(t, X, a) = \sum \frac{(-1)^{m-1}}{mr} \int_{\tau\in\mathcal{T}_{\mathbf{r}}(t)} \frac{(\tau_r - t)^{k_r}}{k_r!} \cdots \frac{(\tau_1 - t)^{k_1}}{k_1!} a_{i_r}(\tau_r)\cdots a_{i_1}(\tau_1)\,\mathrm{d}\tau \\ [\cdots[\mathrm{ad}^{k_r}_{X_0}(X_{i_r}), \mathrm{ad}^{k_{r-1}}_{X_0}(X_{i_{r-1}})], \dots \mathrm{ad}^{k_1}_{X_0}(X_{i_1})]\end{aligned} \tag{2.80}$$

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

$$[\cdots[\mathrm{ad}^{k_r}_{X_0}(X_{i_r}), \mathrm{ad}^{k_{r-1}}_{X_0}(X_{i_{r-1}})], \dots \mathrm{ad}^{k_1}_{X_0}(X_{i_1})] \tag{2.81}$$

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 $\mathcal{L}(X)$

$$\mathcal{Z}^{r,\nu}_\infty(t, X, a) = \sum_{b\in\mathcal{B}_{r,\nu}} \eta_b(t, a)b. \tag{2.82}$$

By summing these relations, we get (2.79). □

Definition 2.42 (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 (2.79) 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 2.43 (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*

$$[a, b] = \sum_{c \in \mathcal{B}} \gamma_{a,b}^{c} c, \tag{2.83}$$

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 2.44 (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 $b_1, b_2 \in \mathcal{B}$,*

$$\sum_{c \in \mathcal{B}} |\gamma_{b_1,b_2}^{c}| \leq \Gamma^{|b_1|+|b_2|}. \tag{2.84}$$

Definition 2.45 (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 $n(b_1) + n(b_2) \leq k$,*

$$\sum_{c \in \mathcal{B}} |\gamma_{b_1,b_2}^{c}| \leq \Gamma(k)^{|b_1|+|b_2|}. \tag{2.85}$$

Asymmetric geometric growth is a weaker notion than geometric growth (which can be seen as asymmetric geometric growth with a constant Γ independent on k). Up to our knowledge, it is not known if the usual bases of $\mathcal{L}(X)$ have these properties. These definitions therefore lead to the following algebraic open problem, which will be studied in a forthcoming paper, where we intend to prove, in particular, that there exist bases, useful for control theory, satisfying these properties.

Open problem 2.46. *Which monomial bases $\mathcal{B}$ of $\mathcal{L}(X)$ have (asymmetric) geometric growth?*

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 (2.80).

Lemma 2.47. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \ldots, X_q\}$, $\mathcal{B}$ 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 $i_1, \ldots, i_r \in [\![1, q]\!]$ and $k_1, \ldots, k_r \in \mathbb{N}$, for every $b \in \mathcal{B}$,*

$$\left| \left\langle [\cdots [\mathrm{ad}_{X_0}^{k_r}(X_{i_r}), \mathrm{ad}_{X_0}^{k_{r-1}}(X_{i_{r-1}})], \ldots \mathrm{ad}_{X_0}^{k_1}(X_{i_1})], b \right\rangle_{\mathcal{B}} \right| \leq C(r)^{|b|}, \tag{2.86}$$

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

Proof. Since the basis is monomial, by homogeneity, the bra-ket vanishes when $|b| \neq k_1 + \cdots + k_r + r$. Moreover, for each $j \in [\![1, q]\!]$, 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. Then, repeated nested application of (2.85) yields

$$\begin{aligned} |\langle [\cdots [b_r, b_{r-1}], \ldots b_1], b \rangle_{\mathcal{B}}| &\leq \Gamma(2)^{k_r+k_{r-1}+2} \Gamma(3)^{k_r+k_{r-1}+k_{r-2}+3} \cdots \Gamma(r)^{k_r+\cdots+k_1+r} \\ &\leq \Gamma(r)^{(r-1)(k_1+\cdots+k_r+r)}, \end{aligned} \tag{2.87}$$

since the sequence Γ is non-decreasing. This proves the desired estimate with $C(r) := \Gamma(r)^{r-1}$. □

Proposition 2.48. *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 $t \in [0,T]$,*

$$|\eta_b(t,u)| \leq \frac{C_M^{|b|}}{|b|!} t^{n_0(b)} \|u\|_{L^1(0,t)}^{n(b)}. \tag{2.88}$$

Proof. We may assume that $(C(r))_{r\in\mathbb{N}^*}$ given by Lemma 2.47 is non-decreasing. Let $M \in \mathbb{N}^*$ and $b \in \mathcal{B}$ be such that $n(b) \leq M$. We deduce from (2.73) that

$$\begin{aligned}\eta_b(t,u) = \sum \frac{(-1)^{m-1}}{mr} \int_{\tau\in\mathcal{T}_{\mathbf{r}}(t)} \frac{(\tau_r - t)^{k_r}}{k_r!} \cdots \frac{(\tau_1 - t)^{k_1}}{k_1!} u_{j_r}(\tau_r) \cdots u_{j_1}(\tau_1)\, \mathrm{d}\tau \\ \left\langle [\cdots [\mathrm{ad}_{X_0}^{k_r}(X_{j_r}), \mathrm{ad}_{X_0}^{k_{r-1}}(X_{j_{r-1}})], \dots \mathrm{ad}_{X_0}^{k_1}(X_{j_1})], b \right\rangle\end{aligned} \tag{2.89}$$

where the sum is taken over $r \in [\![1,\infty]\!]$, $m \in [\![1,r]\!]$, $\mathbf{r} \in \mathbb{N}_r^m$, $k_1, \dots, k_r \in \mathbb{N}$ and $j_1, \dots, j_r \in [\![1,q]\!]$. If the summand bra-ket in (2.89) does not vanish, then $r = n(b)$ and $k_1 + \cdots + k_r = n_0(b)$. Thus the sum in (2.89) is taken over the finite set $r = n(b)$, $m \in [\![1, n(b)]\!]$, $k_1, \dots, k_r \in \mathbb{N}$ such that $k_1 + \cdots + k_r = n_0(b)$ and $j_1, \dots, j_r \in [\![1,q]\!]$, whose cardinal is bounded by $M 2^{|b|} q^M$. Moreover, for every $r, m, k_1, \dots, k_r, j_1, \dots, j_r$ in this set, the associated term in (2.89) is bounded, thanks to Lemma 2.47, by

$$\frac{t^{k_1}}{k_1!} \cdots \frac{t^{k_r}}{k_r!} \|u\|_{L^1(0,t)}^r C(r)^{|b|} \leq t^{n_0(b)} \|u\|_{L^1}^r \left(2^r C(r)\right)^{|b|} \frac{n(b)!}{|b|!} \tag{2.90}$$

thanks to (3.2). Thus

$$|\eta_b(t,u)| \leq \frac{1}{|b|!} M! M q^M \left(2^{M+1} C(M)\right)^{|b|} t^{n_0(b)} \|u\|_{L^1}^{n(b)} \tag{2.91}$$

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 (2.8) as a product of exponentials of the members of a generalized 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 to generalized Hall bases of Sussmann's infinite product on length-compatible Hall bases [68], suggested in [45].

2.5.1 Lazard sets, Hall sets and generalized Hall bases

Definition 2.49. *A Lazard set is a subset $\mathcal{B}$ of $\mathrm{Br}(X)$, totally ordered by a relation $<$ and such that, for every $M \in \mathbb{N}^*$, the set $\mathcal{B}_{[\![1,M]\!]}$ of elements of $\mathcal{B}$ with length at most M, denoted $\mathcal{B}_{[\![1,M]\!]} = \{b_1, \dots, b_{k+1}\}$ with $k \in \mathbb{N}$ and $b_1 < \cdots < b_{k+1}$ satisfies*

$$\begin{cases} b_1 \in Y_0 := X, \\ b_2 \in Y_1 := \{\mathrm{ad}_{b_1}^j(v); j \in \mathbb{N}, v \in Y_0 \setminus \{b_1\}\}, \\ \cdots \\ b_{k+1} \in Y_k := \{\mathrm{ad}_{b_k}^j(v); j \in \mathbb{N}, v \in Y_{k-1} \setminus \{b_k\}\} \end{cases} \tag{2.92}$$

and

$$\mathcal{B}_{[\![1,M]\!]} \cap Y_k = \{b_{k+1}\}, \tag{2.93}$$

where condition (2.93) can equivalently be written

$$\mathcal{B}_{[\![1,M]\!]} \cap Y_{k+1} = \emptyset, \tag{2.94}$$

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

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 $\mathcal{B}$.

Viennot proves in [70, Proposition 1.1 and Theorem 1.1] that properties (2.92) and (2.93) ensure that $\mathrm{eval}(\mathcal{B})$ is a linearly independent and generating family of $\mathcal{L}(X)$. In particular, $\mathrm{eval} : \mathrm{Br}(X) \to \mathcal{L}(X)$ is one to one on $\mathcal{B}$, thus $\mathcal{B}$ can be regarded as a set of Lie monomials.

Definition 2.50 (Hall set)**.** *A Hall set is a subset* $\mathcal{B}$ *of* $\mathrm{Br}(X)$*, totally ordered by a relation* $<$ *and such that*

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

When $b = [b_1, [b_3, b_4]] \in \mathcal{B}$ then b_1 is "sandwiched" in between b_3 and b, since $b_3 \leq b_1 < b$.

Remark 2.51. *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* $b_1 \in \mathcal{B}$*,* $|b_1| = n - 1$*,* $b_2 \in X$ *and* $b_1 < b_2$*. Then for each bracket* $b_2 = [b_2', b_2''] \in \mathcal{B}$ *with length* $|b_2| < n$ *one adds all the* $[b_1, b_2]$ *with* $b_1 \in \mathcal{B}$ *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 [70, Corollary 1.1] that a subset $\mathcal{B}$ of $\mathrm{Br}(X)$ is a Lazard set iff it is a Hall set. With a slight abuse of naming, we call $\mathcal{B}$ a "*generalized Hall basis of* $\mathcal{L}(X)$".

Definition 2.52 (Generalized Hall basis)**.** $\mathcal{B}$ *is a generalized Hall basis of* $\mathcal{L}(X)$ *if* $\mathcal{B}$ *is a Hall set or equivalently a Lazard set.*

Remark 2.53. *Historically, Hall bases where introduced by Marshall Hall in [38], based on ideas of Philip Hall in [39]. In his historical narrower definition, the third condition in Definition 2.50 was replaced by the stronger condition: for every* $b_1, b_2 \in \mathcal{B}$*,* $b_1 < b_2 \Rightarrow |b_1| \leq |b_2|$*. To avoid confusion with the generalized definition, we name them* length-compatible Hall bases *in the sequel.*

Two famous families of generalized Hall bases of $\mathcal{L}(X)$ are the Chen-Fox-Lyndon basis (see [70, Chapter 1]) and the historical length-compatible Hall bases, for which $b_1 < b_2 \Rightarrow |b_1| \leq |b_2|$.

Example 2.54. *For instance, with* $X = \{X_1, X_2\}$*, the elements with length at most* 4 *of each generalized Hall basis* $\mathcal{B}$ *of* $\mathcal{L}(X)$ *with a length-compatible order* $<$ *such that* $X_1 < X_2$ *are:* X_1*,* X_2*,* $[X_1, X_2]$*,* $\mathrm{ad}^2_{X_1}(X_2)$*,* $[X_2, [X_1, X_2]]$*,* $\mathrm{ad}^3_{X_1}(X_2)$*,* $[X_2, \mathrm{ad}^2_{X_1}(X_2)]$*,* $\mathrm{ad}^2_{X_2}([X_1, X_2])$*. Note that* $[X_1, [X_2, [X_1, X_2]]]$ *does not belong to* $\mathcal{B}$ *because* $\lambda([X_2, [X_1, X_2]]) = X_2$ *is not smaller than* X_1*, and the following equality holds in* $\mathcal{L}(X)$

$$[X_1, [X_2, [X_1, X_2]]] = [[X_1, X_2], [X_1, X_2]] + [X_2, [X_1, [X_1, X_2]]] = [X_2, \mathrm{ad}^2_{X_1}(X_2)] \tag{2.95}$$

This illustrates how Definition 2.50 prevents elements from $\mathrm{Br}(X)$*, whose evaluations in* $\mathcal{L}(X)$ *are linked by Jacobi relations, to appear simultaneously in* $\mathcal{B}$.

2.5.2 Infinite product on a generalized Hall basis

Definition 2.55 (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

$$\overleftarrow{\prod_{j\in J}} S^j = \sum_{\sigma\in I^*} P_\sigma X_\sigma, \tag{2.96}$$

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

$$P_\sigma := \sum_{n=0}^{|\sigma|} \sum_{\substack{\sigma_1,\dots,\sigma_n\in I^*,\\ X_{\sigma_1}\cdots X_{\sigma_n}=X_\sigma}} \sum_{\substack{j_1,\dots,j_n\in J,\\ j_1>\cdots>j_n}} \langle S^{j_1}, X_{\sigma_1}\rangle \cdots \langle S^{j_n}, X_{\sigma_n}\rangle. \tag{2.97}$$

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

Lemma 2.56. *Let $\mathcal{B}$ be a generalized 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 $\sigma \in I^*$,*

$$\left\langle \overleftarrow{\prod_{b\in\mathcal{B}}} e^{\alpha_b b}, X_\sigma \right\rangle = \left\langle \overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,|\sigma|]\!]}}} e^{\alpha_b b}, X_\sigma \right\rangle \tag{2.98}$$

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

$$e^{\alpha b} - 1 = \sum_{k=1}^{+\infty} \frac{\alpha^k}{k!} b^k \tag{2.99}$$

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 (2.97), the formula (2.98) holds. □

2.5.3 Coordinates of the second kind

Definition 2.57. *Let $\mathcal{B}$ be a generalized 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 couple (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*

$$\xi_b(t;a) := \frac{1}{m!} \int_0^t \xi^m_{b_1}(\tau;a)\dot{\xi}_{b_2}(\tau;a)\,\mathrm{d}\tau. \tag{2.100}$$

Formula (2.100) indeed defines continuous functionals on L^1 and the following estimates hold.

Lemma 2.58. *Let $a_i \in L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{K})$ for $i \in I$. For every $b \in \mathcal{B}$ and $t \geq 0$,*

$$|\dot{\xi}_b(t;a)| \leq |b||a(t)|\|a\|^{|b|-1}_{L^1(0,t)}, \tag{2.101}$$

$$|\xi_b(t;a)| \leq \|a\|^{|b|}_{L^1(0,t)}. \tag{2.102}$$

Proof. Estimate (2.101) 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 (2.100). Estimate (2.102) is obtained by time-integration of (2.101) for each b. □

2.5.4 Infinite product expansion of the solution to the formal ODE

Theorem 2.59. *Let $\mathcal{B}$ be a generalized 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 ODE (2.8) satisfies, for every $t \in [0,T]$,*

$$x(t) = x^\star \overleftarrow{\prod_{b\in\mathcal{B}}} e^{\xi_b(t;a)b}. \tag{2.103}$$

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 2.56 it is sufficient to prove that, for every $t \in [0,T]$ and $\sigma \in I^*$

$$\langle x(t), X_\sigma\rangle = \left\langle \overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,|\sigma|]\!]}}} e^{\xi_b(t)b}, X_\sigma \right\rangle. \tag{2.104}$$

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 (2.92). The equality (2.104) can equivalently we written

$$\langle x(t), X_\sigma\rangle = \left\langle e^{\xi_{b_{k+1}}(t)b_{k+1}}\cdots e^{\xi_{b_1}(t)b_1}, X_\sigma \right\rangle. \tag{2.105}$$

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

$$x_j(t) := x(t)e^{-\xi_{b_1}(t)b_1}\cdots e^{-\xi_{b_j}(t)b_j}. \tag{2.106}$$

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

$$\dot{x}_j(t) = x_j(t)\left(\sum_{b\in Y_j}\dot{\xi}_b(t)b\right) \quad \text{and} \quad x_j(0) = 1. \tag{2.107}$$

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 (2.107) holds at step $j-1$. We deduce from the definition of $x_j(t)$ that

$$x_j(t) = x_{j-1}(t)e^{-\xi_{b_j}(t)b_j}. \tag{2.108}$$

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

$$\begin{aligned}
\dot{x}_j(t) &= x_{j-1}(t)\left(\sum_{b\in Y_{j-1}}\dot{\xi}_b(t)b\right)e^{-\xi_{b_j}(t)b_j} - x_{j-1}(t)\dot{\xi}_{b_j}(t)b_j e^{-\xi_{b_j}(t)b_j} \\
&= x_j(t)e^{\xi_{b_j}(t)b_j}\left(\sum_{b\in Y_{j-1}\setminus\{b_j\}}\dot{\xi}_b(t)b\right)e^{-\xi_{b_j}(t)b_j} \\
&= x_j(t)\sum_{m\in\mathbb{N}}\sum_{b\in Y_{j-1}\setminus\{b_j\}}\frac{\xi^m_{b_j}(t)}{m!}\dot{\xi}_b(t)\,\mathrm{ad}^m_{b_j}(b)
\end{aligned} \tag{2.109}$$

which ends the proof by induction.

We deduce from (2.94) and (2.107) for $j = (k+1)$ that $x_{k+1}(t)-1$ has non vanishing coefficients only on monomials $X_{\sigma'}$ with $|\sigma'| > |\sigma|$. Therefore, by (2.97),

$$\langle x(t), X_\sigma \rangle = \left\langle x_{k+1}(t) e^{\xi_{b_{k+1}}(t) b_{k+1}} \cdots e^{\xi_{b_1}(t) b_1}, X_\sigma \right\rangle = \left\langle e^{\xi_{b_{k+1}}(t) b_{k+1}} \cdots e^{\xi_{b_1}(t) b_1}, X_\sigma \right\rangle, \quad (2.110)$$

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, that are used in the sequel. For the sake of completeness, the proofs, although classical, are provided.

Throughout the whole paper, $d \in \mathbb{N}^*$ 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 $|\alpha| := \alpha^1 + \cdots + \alpha^a$, $\partial^\alpha := \partial_{x_1}^{\alpha^1} \cdots \partial_{x_a}^{\alpha^a}$ and $\alpha! := \alpha^1! \cdots \alpha^a!$.

Lemma 3.1. *The following estimates hold*

$$\forall n \in \mathbb{N}, \quad n^n e^{-n} e \leq n! \leq (n+1)^{n+1} e^{-(n+1)} e, \quad (3.1)$$

$$\forall a \in \mathbb{N}^*, \forall \alpha = (\alpha^1, \ldots, \alpha^a) \in \mathbb{N}^a, \quad 2^{-(a-1)|\alpha|} |\alpha|! \leq \alpha! \leq |\alpha|! \quad (3.2)$$

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

3.1.2 Regular functions and vector fields

Definition 3.2 (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*

$$\|f\|_{\mathcal{C}^k} := \sum_{j=1}^{b} \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \|\partial^\alpha f_j\|_{L^\infty(K)}, \quad (3.3)$$

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 $k \in \mathbb{N}$.

Definition 3.3 (Regular vector fields)**.** *Let $\delta > 0$ and $k \in \mathbb{N}$. We define $\mathcal{C}^k_\delta := \mathcal{C}^k(B_\delta; \mathbb{K}^d)$ the space of vector fields on $\mathbb{K}^d$ defined and regular in a ball of radius δ. We denote by $\mathcal{C}^\infty_\delta$ the intersection of these spaces over $k \in \mathbb{N}$.*

3.1.3 Analytic norms

Definition 3.4 (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*

$$|||f|||_r := \sum_{i=1}^{b} \sum_{\alpha\in\mathbb{N}^d} \frac{r^{|\alpha|}}{\alpha!} \|\partial^\alpha f_i\|_{L^\infty(K)}. \tag{3.4}$$

Definition 3.5 (Analytic vector fields)**.** *Let $r, \delta > 0$. We define $\mathcal{C}^{\omega,r}_\delta := \mathcal{C}^{\omega,r}(B_\delta;\mathbb{K}^d)$ the space of real-analytic vector fields on $\mathbb{K}^d$ defined in a ball of radius δ. We denote by $\mathcal{C}^\omega_\delta$ the union of these spaces over $r > 0$.*

Lemma 3.6 (Algebra property)**.** *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*

$$|||fg|||_r \le |||f|||_r \, |||g|||_r \,. \tag{3.5}$$

Proof. Using the multivariate Leibniz formula, one has

$$\begin{aligned} |||fg|||_r &= \sum_{\alpha\in\mathbb{N}^a} \frac{r^{|\alpha|}}{\alpha!} \|\partial^\alpha(fg)\|_{L^\infty(K)} \\ &\le \sum_{\alpha\in\mathbb{N}^a} \frac{r^{|\alpha|}}{\alpha!} \sum_{\beta\le\alpha} \binom{\alpha}{\beta} \|\partial^\beta f\|_{L^\infty(K)} \|\partial^{\alpha-\beta} g\|_{L^\infty(K)} = |||f|||_r \, |||g|||_r \,, \end{aligned} \tag{3.6}$$

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

Lemma 3.7 (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 $j \in [\![1, a]\!]$,*

$$|||\partial_j f|||_{r_1} \le \frac{1}{r_1} \left(e \ln \frac{r_2}{r_1} \right)^{-1} |||f|||_{r_2} \,. \tag{3.7}$$

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

$$|||\partial_j f|||_{r_1} \le \frac{1}{r_2 - r_1} \, |||f|||_{r_2} \,. \tag{3.8}$$

Proof. We start with the first estimate (3.7). One has

$$\begin{aligned} |||\partial_j f|||_{r_1} = \sum_{\alpha\in\mathbb{N}^a} \frac{r_1^{|\alpha|}}{\alpha!} \|\partial^{\alpha+e_j} f\|_{L^\infty(K)} &= \frac{1}{r_1} \sum_{\alpha\in\mathbb{N}^a} \frac{r_1^{|\alpha+e_j|}}{(\alpha+e_j)!} \frac{(\alpha+e_j)!}{\alpha!} \|\partial^{\alpha+e_j} f\|_{L^\infty(K)} \\ &\le \frac{1}{r_1} \, |||f|||_{r_2} \sup_{\alpha\in\mathbb{N}^a} \left(\frac{r_1}{r_2} \right)^{|\alpha+e_j|} \frac{(\alpha+e_j)!}{\alpha!} \\ &\le \frac{1}{r_1} \, |||f|||_{r_2} \sup_{n\ge1} n \left(\frac{r_1}{r_2} \right)^n . \end{aligned} \tag{3.9}$$

For $x \in (0,1)$, let $C(x) := \sup_{n\ge1} n x^n = \sup_{n\ge1} \exp(\ln n + n \ln x)$. Differentiating inside the exponent with respect to $n \in [1, +\infty)$ yields

$$\frac{\partial}{\partial n} (\ln n + n \ln x) = \frac{1}{n} + \ln x. \tag{3.10}$$

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

$$C(x) \le (-e \ln x)^{-1} \,. \tag{3.11}$$

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 (3.11) 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 3.8. *The first estimate* (3.7) *is classical (see e.g. [56]). The second estimate* (3.7) *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 3.7 in the sequel since we consider small radius losses.*

3.2 Well-posedness of ordinary differential equations

The nonlinear differential equations

$$\dot{x}(t) = f(t, x(t)) \quad \text{and} \quad x(0) = p \tag{3.12}$$

will be studied in the following classical frameworks.

Lemma 3.9. *Let $\delta, T > 0$ and $f \in L^1((0,T); \mathcal{C}^1_{2\delta})$ 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*

$$\forall t \in [0,T], \quad x(t; f, p) = p + \int_0^t f\left(\tau, x(\tau; f, p)\right) \mathrm{d}\tau. \tag{3.13}$$

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 (3.12) 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, for each $p \in B_\delta$ and $t \in [0,T]$,*

$$|x(t; f, p) - x(t; g, p)| \leq \|f - g\|_{L^1((0,t);\mathcal{C}^0)} \exp\left(\|f\|_{L^1((0,t);\mathcal{C}^1)}\right). \tag{3.14}$$

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

1. Define $\Theta : X \to X$ by $\Theta(x)(t) := p + \int_0^t f(\tau, x(\tau))\,\mathrm{d}\tau$ for $x \in X$. 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 $F : B_\delta \times X \to X$ by

$$\forall t \in [0,T], \quad F(p, x)(t) := x(t) - p - \int_0^t f(\tau, x(\tau))\,\mathrm{d}\tau \tag{3.15}$$

   Then F is of class $\mathcal{C}^\infty$, vanishes at $(\bar{p}, \bar{x})$ and $\frac{\partial F}{\partial x}(\bar{p}, \bar{x})$ is a bijection on X. 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 3.10. *Let $\delta, \delta_u > 0$, $q \in \mathbb{N}^*$ and $f \in \mathcal{C}^\omega(B_{2\delta} \times \overline{B}_{\mathbb{K}^q}(0, \delta_u); \mathbb{K}^d)$. 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*

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)), \\ x(0) = p, \end{cases} \tag{3.16}$$

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 3.9. 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. [21, Theorem 4.5.3]). □

3.3 Estimates for differential operators and Lie brackets

As is usual, a smooth vector field f is identified with the first-order linear differential operator $f\nabla$ acting on smooth functions and defined as $(f\nabla)\phi : p \mapsto f(p) \cdot \nabla\phi(p)$.

3.3.1 Estimates for products

Lemma 3.11. *Let $r_2 > 0$ and $r_1 \in [r_2/e, r_2)$. Let $n \in \mathbb{N}^*$ and $\delta > 0$. For every $f_1, \dots, f_n \in \mathcal{C}^{\omega,r_2}_\delta$ and $\phi \in \mathcal{C}^{\omega,r_2}_\delta$,*

$$|||(f_n \cdot \nabla) \cdots (f_1 \cdot \nabla)\phi|||_{r_1} \leq \frac{n!}{e} \left(\frac{e}{r_2 - r_1}\right)^n |||f_n|||_{r_2} \cdots |||f_1|||_{r_2} \, |||\phi|||_{r_2} \,. \tag{3.17}$$

In particular, under the same assumptions,

$$\|(f_n \cdot \nabla) \cdots (f_1 \cdot \nabla)\phi\|_{\mathcal{C}^0} \leq n! \left(\frac{5}{r_2}\right)^n |||f_n|||_{r_2} \cdots |||f_1|||_{r_2} \, |||\phi|||_{r_2} \,. \tag{3.18}$$

Proof. For $n = 1$, estimate (3.17) is a consequence of (3.4), (3.5) and (3.8). 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

$$\begin{aligned} |||(f_n \cdot \nabla) \cdots (f_1 \cdot \nabla)\phi|||_{r_1} &\leq \left(\frac{n}{r_2 - r_1}\right) |||f_n|||_{r_1} \, |||(f_{n-1} \cdot \nabla) \cdots (f_1 \cdot \nabla)\phi|||_{r_1 + \frac{r_2 - r_1}{n}} \\ &\leq \left(\frac{n}{r_2 - r_1}\right)^n |||\phi|||_{r_2} \prod_{j=1}^{n} |||f_j|||_{r_1 + (n-j)\frac{r_2 - r_1}{n}} \,, \end{aligned} \tag{3.19}$$

which concludes the proof because the norm (3.4) is non-decreasing with respect to r, and we can bound n^n using (3.1). Estimate (3.18) is a direct consequence for the particular choice $r_1 = r_2/e$, because $e^2/(e-1) \leq 5$. □

3.3.2 Lie brackets

Definition 3.12 (Lie bracket of vector fields)**.** *For smooth vector fields f, g, we define $[f, g]\nabla$ as the usual commutator of the associated operators $f\nabla$ and $g\nabla$, hence $[f, g]\nabla = [f\nabla, g\nabla] = (f\nabla)(g\nabla) - (g\nabla)(f\nabla) = ((f\nabla g) - (g\nabla f))\nabla$, where the last equality comes from Schwarz's theorem. In particular, $[f, g]\nabla$ is the operator associated to the smooth vector field $[f, g] := (f \cdot \nabla)g - (g \cdot \nabla)f$.*

Definition 3.13 (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 a formal Lie bracket $b \in \mathrm{Br}(X)$, we define $f_b = \Lambda(b)$, where $\Lambda : \mathcal{L}(X) \to \mathcal{C}^\infty(\Omega; \mathbb{K}^d)$ is the unique homomorphism of Lie algebra such that $\Lambda(X_i) = f_i$ for every $i \in I$ (see Lemma 2.7).*

The vector field f_b is obtained by replacing the indeterminates X_i with the corresponding vector fields f_i in the formal 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, build under weaker regularity assumptions, for instance $f_i \in \mathcal{C}^{|b|-1}$ and then $f_b \in \mathcal{C}^0$.

Lemma 3.14 (Finite regularity estimate)**.** *Let $k \in \mathbb{N}$ and $b \in \mathrm{Br}(X)$. Let $\delta > 0$ and $f_i \in \mathcal{C}^{k+|b|-1}_\delta$ for $i \in I$. Then,*

$$\|f_b\|_{\mathcal{C}^k} \leq 2^{|b|} \frac{(k + |b| - 1)!}{k!} \prod_{i \in I} \|f_i\|^{n_i(b)}_{\mathcal{C}^{k+|b|-1}}. \tag{3.20}$$

Proof. This estimate follows from (3.3), the algebra property that this norm satisfies and the estimate $\|\partial^j f\|_{\mathcal{C}^m} \leq (m+1)\|f\|_{\mathcal{C}^{m+1}}$ for every $j \in [\![1,d]\!]$, $m \in \mathbb{N}$ and $f \in \mathcal{C}^{m+1}_\delta$. □

Lemma 3.15 (Analytic estimate)**.** *Let $r_2 > 0$ and $r_1 \in [r_2/e, r_2)$. Let $\delta > 0$. Let $f_i \in \mathcal{C}^{\omega,r_2}_\delta$ for $i \in I$ and $b \in \mathrm{Br}(X)$. Then,*

$$|||f_b|||_{r_1} \leq \frac{(|b|-1)!}{e}\left(\frac{2e}{r_2-r_1}\right)^{|b|-1}\prod_{i\in I}|||f_i|||_{r_2}^{n_i(b)}. \tag{3.21}$$

In particular, under the same assumptions,

$$\|f_b\|_{\mathcal{C}^0} \leq (|b|-1)!\left(\frac{9}{r_2}\right)^{|b|-1}\prod_{i\in I}|||f_i|||_{r_2}^{n_i(b)}. \tag{3.22}$$

$$\|f_b\|_{\mathcal{C}^1} \leq \max\left\{1,\frac{1}{r_2}\right\}(|b|-1)!\left(\frac{9}{r_2}\right)^{|b|-1}\prod_{i\in I}|||f_i|||_{r_2}^{n_i(b)}. \tag{3.23}$$

Proof. Estimate (3.21) stems from (3.17) 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 3.11, where ϕ is one of the vector fields f_i. Estimates (3.22) and (3.23) are direct consequences of (3.21) 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 3.16. *The fact that estimate* (3.21) *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*

$$f_0(x) := e_1 \quad \textit{and} \quad f_1(x) := \frac{1}{1-x_1}e_2. \tag{3.24}$$

Using (3.4)*, one checks that these vector fields belong in particular to $\mathcal{C}^{\omega,r}_\delta$ 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*

$$\mathrm{ad}^k_{f_0}(f_1)(x) = \frac{\partial^k}{\partial x_1^k}\left(\frac{1}{1-x_1}\right)e_2 = \frac{k!}{(1-x_1)^{k+1}}e_2. \tag{3.25}$$

Moreover, since f_0 is constant and f_1 depends only on x_1 but is supported by 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.

3.3.3 Nilpotent Lie algebra of vector fields

Lemma 3.17. *Let $\mathcal{F}$ be a set of $\mathcal{C}^\infty$ vector fields on a subset Ω of $\mathbb{K}^d$. If each Lie bracket with length m of vector fields in $\mathcal{F}$ vanishes on Ω, then the Lie algebra $\mathcal{L}(\mathcal{F})$ generated by $\mathcal{F}$ is nilpotent with index at most m (see Definition 2.5).*

Proof. Each Lie bracket with length $m+1$ or more of vector fields in $\mathcal{F}$ vanishes on Ω. This can be proved by expanding the bracket into monomials and then applying Dynkin's formula (Lemma 2.21) to recover brackets with length m inside brackets with length $m+1$ or more. □

3.4 Flows, compositions and pushforwards

3.4.1 Definitions and approximations

By applying Lemma 3.9 to a time-independent vector field we obtain the following object.

Definition 3.18 (Flow of time-independent vector fields)**.** *Let $\delta > 0$. Let $f \in \mathcal{C}^1_{2\delta}$ such that $\|f\|_{\mathcal{C}^0} < \delta$. We denote by e^f the flow at time one of the vector field f,*

$$e^f : \begin{cases} B_\delta \to B_{2\delta}, \\ p \mapsto x(1; f, p), \end{cases} \tag{3.26}$$

with the notations of Section 3.2. We will write $e^f p$ instead of $e^f(p)$ to allow easier composition of flows. If $f \in \mathcal{C}^\infty_{2\delta}$, then 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 3.19. *Let $\delta > 0$ and $f \in \mathcal{C}^1_\delta$. Assume that $\delta' := \delta - \|f\|_{\mathcal{C}^0_\delta} > 0$. For each $p \in B_{\delta'}$, $e^f p$ is well-defined and $e^f p \in B_\delta$. Moreover,*

$$|e^f p - p| \leq \|f\|_{\mathcal{C}^0_\delta}, \tag{3.27}$$

and

$$\|D(e^f)\|_{\mathcal{C}^0_{\delta'}} \leq e^{\|Df\|_{\mathcal{C}^0_\delta}} \leq e^{\|f\|_{\mathcal{C}^1_\delta}}. \tag{3.28}$$

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

$$\dot{R}(t) = Df(e^{tf} p) R(t) \quad \text{and} \quad R(0) = \mathrm{Id}. \tag{3.29}$$

Thus, by Grönwall's lemma,

$$\|R(1)\| \leq \|\mathrm{Id}\| e^{\int_0^1 \|Df(e^{tf} p)\| \, \mathrm{d}t} \leq e^{\|Df\|_{\mathcal{C}^0}}, \tag{3.30}$$

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 operator $f\nabla$. It is completely legitimate in the analytic setting, as underlined by the following result.

Lemma 3.20 (Approximation of autonomous flows)**.** *Let $\delta > 0$ and $f \in \mathcal{C}^1_{2\delta}$ such that $\|f\|_{\mathcal{C}^0} < \delta$.*

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

$$\left| \left(e^f - \sum_{k=0}^{M} \frac{(f \cdot \nabla)^k}{k!} \right) (\phi)(p) \right| \leq \|f\|^{M+1}_{\mathcal{C}^M} \|\nabla \phi\|_{\mathcal{C}^M}. \tag{3.31}$$

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

$$e^{tf}(\phi)(p) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} (f \cdot \nabla)^k \phi(p) \tag{3.32}$$

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

Proof. First statement. By the first point of Lemma 3.9, $e^{tf}(p)$ 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

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k} \left[\phi(e^{tf}(p)) \right] = \big((f \cdot \nabla)^k \phi \big)(e^{tf}(p)). \tag{3.33}$$

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

$$\left(e^f - \sum_{k=0}^{M} \frac{(f \cdot \nabla)^k}{k!} \right) (\phi)(p) = \int_0^1 \frac{(1-s)^M}{M!} \big((f \cdot \nabla)^{M+1} \phi \big)(e^{sf}(p)) \, \mathrm{d}s. \tag{3.34}$$

By induction on $k \in \mathbb{N}$, one checks that $(f \cdot \nabla)^k \phi$ is a sum of at most $k!$ terms of the form

$$(\nabla^{j_1} f) \cdots (\nabla^{j_k} f)(\nabla^{j_0} \phi), \tag{3.35}$$

where $j_0 + j_1 + \ldots + j_k = k$ and $j_0 \geq 1$. This concludes the proof of (3.31) with a constant 1 thanks to the integration in (3.34).

Second statement. Let $r > 0$ be such that $f \in \mathcal{C}^{\omega,r}_{2\delta}$ and $\phi \in \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K})$. Let $r' \in [r/e, r)$. By (3.17), for each $k \in \mathbb{N}$,

$$\left| \left| \left| \frac{t^k}{k!} (f \cdot \nabla)^k (\phi) \right| \right| \right|_{r'} \leq \frac{|t|^k}{k!} \frac{k!}{e} \left(\frac{e}{r - r'} \right)^k |||f|||_r^k \, |||\phi|||_r \,, \tag{3.36}$$

so that the sum converges absolutely in $\mathcal{C}^{\omega,r'}$ when $|t| e \, |||f|||_r < r - r'$. Moreover, by (3.34) with $f \leftarrow tf$ and (3.17),

$$\left\| \left(e^{tf} - \sum_{k=0}^{M} \frac{t^k}{k!} (f \cdot \nabla)^k \right) (\phi) \right\|_{\mathcal{C}^0} \leq \frac{|t|^{M+1}}{(M+1)!} \| (f \cdot \nabla)^{M+1} (\phi) \|_{\mathcal{C}^0}, \tag{3.37}$$

where, using (3.36), 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 3.21 (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*

$$(\theta_* f)(q) := (D\theta)_{|\theta^{-1}(q)} f(\theta^{-1}(q)) = (D\theta^{-1}(q))^{-1} f(\theta^{-1}(q)). \tag{3.38}$$

Lemma 3.22 (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'',*

$$\theta'_*(\theta_* f) = (\theta' \circ \theta)_* f. \tag{3.39}$$

Proof. This is a consequence of the chain rule for differentiation, see e.g. [51, Problem 12-10]. □

Lemma 3.23 (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',*

$$[\theta_* f, \theta_* g] = \theta_* [f, g]. \tag{3.40}$$

Proof. This is a consequence of the chain rule for differentiation, see e.g. [51, Corollary 8.31]. □

3.4.3 Composition of vector fields with flows

Lemma 3.24. *Let $\delta > 0$, $f_0 \in \mathcal{C}^1_{2\delta}$ 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 $f_0, f_1 \in \mathcal{C}^{M+1}_{2\delta}$, then, for each $p \in B_\delta$,*

$$\left| \partial_p \Phi_0(t, p)^{-1} f_1 \left(\Phi_0(t, p) \right) - \sum_{k=0}^{M-1} \frac{t^k}{k!} \operatorname{ad}^k_{f_0}(f_1)(p) \right| \leq \frac{t^M}{M!} \left\| \operatorname{ad}^M_{f_0}(f_1) \right\|_{\mathcal{C}^0}. \tag{3.41}$$

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

$$(\Phi_0(-t)_* f_1)(p) = \partial_p \Phi_0(t,p)^{-1} f_1\left(\Phi_0(t,p)\right) = \sum_{k=0}^{M-1} \frac{t^k}{k!} \mathrm{ad}^k_{f_0}(f_1)(p). \tag{3.42}$$

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

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

$$(\Phi_0(-t)_* f_1)(p) = \partial_p \Phi_0(t,p)^{-1} f_1\left(\Phi_0(t,p)\right) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \mathrm{ad}^k_{f_0}(f_1)(p), \tag{3.43}$$

where, for every $r' \in [r/e, r)$ the series converges in $\mathcal{C}^{\omega,r'}_{2\delta}$ 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*

$$\left\| e^{H_0} H_1 e^{-H_0} - \sum_{k=0}^{M-1} \frac{1}{k!} \mathrm{ad}^k_{H_0}(H_1) \right\| \leq \frac{(2\|H_0\|)^M}{M!} \|H_1\| e^{2\|H_0\|} \tag{3.44}$$

and

$$e^{H_0} H_1 e^{-H_0} = \sum_{k=0}^{+\infty} \frac{1}{k!} \mathrm{ad}^k_{H_0}(H_1), \tag{3.45}$$

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

Proof. We proceed step by stem.

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

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\tau}\left[\partial_p \Phi_0(\tau,p)^{-1}\right] &= -\partial_p \Phi_0(\tau,p)^{-1} \frac{\mathrm{d}}{\mathrm{d}\tau}\left[\partial_p \Phi_0(\tau,p)\right] \partial_p \Phi_0(\tau,p)^{-1} \\ &= -\partial_p \Phi_0(\tau,p)^{-1} f_0'\left(\Phi_0(\tau,p)\right), \end{aligned} \tag{3.46}$$

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

$$\frac{\mathrm{d}^k}{\mathrm{d}\tau^k}\left[\partial_p \Phi_0(\tau,p)^{-1} f_1\left(\Phi_0(\tau,p)\right)\right] = \partial_p \Phi_0(\tau,p)^{-1} \mathrm{ad}^k_{f_0}(f_1)\left(\Phi_0(\tau,p)\right). \tag{3.47}$$

The Taylor formula

$$\begin{aligned} \partial_p \Phi_0(t,p)^{-1} f_1\left(\Phi_0(t,p)\right) - \sum_{k=0}^{M-1} \frac{t^k}{k!} \mathrm{ad}^k_{f_0}(f_1)(p) \\ = \int_0^t \frac{(t-s)^{M-1}}{(M-1)!} \partial_p \Phi_0(s,p)^{-1} \mathrm{ad}^M_{f_0}(f_1)\left(\Phi_0(s,p)\right) \mathrm{d}s \end{aligned} \tag{3.48}$$

proves the first statement.

2. Equation (3.48) yields the conclusion.

3. Let $r' \in [r/e, r)$. Thanks to (3.21),

$$\left|\!\left|\!\left| \frac{t^k}{k!} \mathrm{ad}^k_{f_0}(f_1) \right|\!\right|\!\right|_{r'} \leq \frac{|t|^k}{k!} \frac{k!}{e} \left(\frac{2e}{r-r'}\right)^k |||f_0|||_r^k \, |||f_1|||_r, \tag{3.49}$$

so the series converges absolutely in $\mathcal{C}^{\omega,r'}_{2\delta}$ 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 (3.48) and (3.49),

$$\begin{aligned}\left|(\Phi_0(-t)_*f_1)(p) - \sum_{k=0}^{M-1} \frac{t^k}{k!}\operatorname{ad}^k_{f_0}(f_1)(p)\right| &\le \frac{|t|^M}{M!}\|\operatorname{ad}^M_{f_0}(f_1)\|_{\mathcal{C}^0} \sup_{s\in[0,t]} \|(\partial_p\Phi_0(s,\cdot))^{-1}\|_{\mathcal{C}^0} \\ &\le A_0\,|||f_1|||_r \left(\frac{2e|t|\,|||f_0|||_r}{r-r'}\right)^M,\end{aligned} \tag{3.50}$$

where A_0 denotes the supremum in the right-hand side of (3.50) which is finite. So the sum converges towards the pushforward under the same smallness assumption on time.

4. The last statement can be proved similarly, by considering the function $t \mapsto e^{tH_0}H_1e^{-tH_0}$.

□

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 nilpotent assumption.

Lemma 3.25. *Let J an open 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 $\lambda \in J$, $\operatorname{ad}^M_{f(\lambda_0)}(f(\lambda)) \equiv 0$. Then, for each $p \in B_\delta$,*

$$\frac{\partial}{\partial\lambda}\left(e^{f(\lambda)}p\right)_{|\lambda=\lambda_0} = \sum_{k=0}^{M-1} \frac{(-1)^k}{(k+1)!}\operatorname{ad}^k_{f(\lambda_0)}\left(\partial_\lambda f(\lambda_0)\right)\left(e^{f(\lambda_0)}p\right). \tag{3.51}$$

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

$$\dot z(t) = \partial_x f(\lambda_0, x_0(t))z(t) + \partial_\lambda f(\lambda_0, x_0(t)). \tag{3.52}$$

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

$$\partial_t R(t,s) = \partial_x f(\lambda_0, x_0(t))R(t,s), \tag{3.53}$$

i.e. $R(t,s) = \partial_p\Theta(t-s,\lambda_0,x_0(s))$. Then by the Duhamel formula

$$\begin{aligned} z(1) &= \int_0^1 R(\tau,1)^{-1}\partial_\lambda f(\lambda_0,x_0(\tau))\,\mathrm{d}\tau \\ &= \int_0^1 \partial_p\Theta(\tau-1,\lambda_0,x_0(1))^{-1}\partial_\lambda f(\lambda_0,\Theta(\tau-1,\lambda_0,x_0(1)))\,\mathrm{d}\tau. \end{aligned} \tag{3.54}$$

By (3.42) of Lemma 3.24 with $t \leftarrow \tau - 1$, $f_0 \leftarrow f(\lambda_0,\cdot)$, $f_1 \leftarrow \partial_\lambda f(\lambda_0,\cdot)$ and $p \leftarrow x_0(1)$,

$$z(1) = \int_0^1 \sum_{k=0}^{M-1} \frac{(\tau-1)^k}{k!}\operatorname{ad}^k_{f(\lambda_0)}\left(\partial_\lambda f(\lambda_0)\right)(x_0(1))\,\mathrm{d}\tau, \tag{3.55}$$

which gives the conclusion. □

4 Error estimates in time for nonlinear vector fields

Using a classical linearization trick for smooth vector fields f, 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 trick for smooth fields

We explain how, by identifying vector fields with first-order differential operators and points on the manifold with the operator of evaluation at this point, one recasts a nonlinear ODE driven by smooth vector fields to a linear equation set on a larger space of operators on smooth functions. This well-known method is notably used in [2] and [65].

4.1.1 Definition of an operator acting on smooth functions

When $T > 0$ and $f \in \mathcal{C}^\infty_c([0,T] \times \mathbb{K}^d)$ satisfies $\|f\|_{L^1_T(\mathcal{C}^0)} < 1$, we take the nonlinear ODE (3.12) back to a linear framework by considering, for every $t \in [0,T]$ the linear operator $L(t)$ on $\mathcal{C}^\infty_c(\mathbb{K}^d;\mathbb{K})$ defined, for $\varphi \in \mathcal{C}^\infty_c(\mathbb{K}^d;\mathbb{K})$, by

$$L(t)\varphi : p \mapsto \varphi\left(x(t;f,p)\right). \tag{4.1}$$

$L(t)\varphi$ is of class $\mathcal{C}^\infty$ as a composition of $\mathcal{C}^\infty$ functions, by the third statement of Lemma 3.9. $L(t)\varphi$ 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 3.9 (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 $t \in [0,T]$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(L(t)\varphi\big)(p) = D\varphi\Big(x(t;f,p)\Big)f\Big(t,x(t;f,p)\Big) = \Big(L(t)(f(t)\cdot\nabla)\varphi\Big)(p). \tag{4.2}$$

Thus, L solves the following linear equation

$$\frac{\mathrm{d}}{\mathrm{d}t}L(t) = L(t)(f(t)\cdot\nabla) \tag{4.3}$$

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

$$\forall \varphi \in \mathcal{C}^\infty_c(\mathbb{K}^d;\mathbb{K}), \forall p \in \mathbb{K}^d, \quad \big(L(t)\varphi\big)(p) = \varphi(p) + \int_0^t \Big(L(\tau)\big(f(\tau)\cdot\nabla\varphi\big)\Big)(p)\,\mathrm{d}\tau, \tag{4.4}$$

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:

$$L(t) = \mathrm{Id} + \int_0^t L(\tau)(f(\tau)\cdot\nabla)\,\mathrm{d}\tau. \tag{4.5}$$

In the sequel, all integral equalities between operators on $\mathcal{C}^\infty_c(\mathbb{K}^d;\mathbb{K})$ should be understood in this weak sense (after evaluation on a test function and at a point). Here the right-hand side refers to the composition of two operators on $\mathcal{C}^\infty_c(\mathbb{K}^d;\mathbb{K})$: $L(t)$ and $\varphi \mapsto (f(t,\cdot)\cdot\nabla)\varphi$, i.e. we identify each vector field with a first-order differential operator on smooth functions.

Equation (4.3) 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 $j\in\mathbb{N}$,

$$L_{j+1}(t) := \int_0^t L_j(\tau)(f(\tau)\cdot\nabla)\,\mathrm{d}\tau, \tag{4.6}$$

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

$$L_j(t) = \int_{0<\tau_j<\cdots<\tau_1<t} (f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\,\mathrm{d}\tau = \int_{\mathcal{T}_{(j)}(t)} (f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\,\mathrm{d}\tau, \tag{4.7}$$

where the integration domain is defined in (2.14). 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 following sense.

Lemma 4.1. *For each $M\in\mathbb{N}$, there exists $C_M>0$ such that, for each $T>0$, $f\in\mathcal{C}_c^\infty([0,T]\times\mathbb{K}^d)$ and $\varphi\in\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$, for each $t\in[0,T]$,*

$$\left\| \left(L(t) - \sum_{j=0}^{M} L_j(t) \right) \varphi \right\|_{\mathcal{C}^0} \leq C_M \|f\|_{L_t^1(\mathcal{C}^M)}^{M+1} \|\varphi\|_{\mathcal{C}^{M+1}}. \tag{4.8}$$

Proof. Let $p\in\mathbb{K}^d$. Thanks to Lemma 3.9 (for δ large enough), $x(\tau;f,p)$ 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 $\tau\in[0,T]$,

$$\varphi(x(\tau;f,p)) = \varphi(p) + \int_0^\tau \big(f(\tau_1)\cdot\nabla\big)(\varphi)(x(\tau_1;f,p))\,\mathrm{d}\tau_1. \tag{4.9}$$

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

$$\begin{aligned} &\varphi\left(x(t;f,p)\right) - \varphi(p) - \sum_{j=1}^{M} \int_{\mathcal{T}_{(j)}(t)} \Big((f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\Big)(\varphi)(p)\,\mathrm{d}\tau \\ &\qquad = \int_{\mathcal{T}_{(M+1)}(t)} \Big((f(\tau_{M+1})\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\Big)(\varphi)\left(x(\tau_{M+1};f,p)\right)\,\mathrm{d}\tau, \end{aligned} \tag{4.10}$$

which concludes the proof. □

4.2 Iterated Duhamel or Chen-Fliess expansion

The approximating sequence for the operator $L(t)$ yields the following straight-forward estimate for the iterated Duhamel or Chen-Fliess expansion of the state.

Proposition 4.2. *For every $M\in\mathbb{N}$, there exists $C_M>0$ such that, for every $\delta>0$, $T>0$, $f\in L^1((0,T);\mathcal{C}_{2\delta}^M\cap\mathcal{C}_{2\delta}^1)$, with $\|f\|_{L_T^1(\mathcal{C}^0)}<\delta$, $p\in B_\delta$ and $\varphi\in\mathcal{C}^{M+1}(B_{2\delta};\mathbb{K})$, for each $t\in[0,T]$,*

$$\left| \varphi\left(x(t;f,p)\right) - \sum_{j=0}^{M} \int_{\mathcal{T}_{(j)}(t)} \Big((f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\Big)(\varphi)(p)\,\mathrm{d}\tau \right| \leq C_M \|f\|_{L_t^1(\mathcal{C}^M)}^{M+1} \|\varphi\|_{\mathcal{C}^{M+1}}. \tag{4.11}$$

In particular

$$\left| x(t;f,p) - \sum_{j=0}^{M} \int_{\mathcal{T}_{(j)}(t)} \Big((f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\Big)(\mathrm{Id}_d)(p)\,\mathrm{d}\tau \right| \leq C_M \|f\|_{L_t^1(\mathcal{C}^M)}^{M+1}. \tag{4.12}$$

Hence, if $f\in L^\infty((0,T);\mathcal{C}_{2\delta}^M)$, both estimates correspond to a bound scaling like t^{M+1}.

Proof. Estimate (4.11) is a direct consequence of Lemma 4.1. Indeed, by density of $\mathcal{C}_c^\infty([0,T]\times\mathbb{K}^d)$ in $L^1((0,T);\mathcal{C}^M_{2\delta})$, one can apply Lemma 4.1 to a sequence f_n of regularized extended versions of f. By the fourth point of Lemma 3.9, the solutions x_n converge towards x. Estimate (4.12) follows by applying (4.11) 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. [58, Proposition 4.3]). It bears a lot of similarity with [29, Theorem 1.32], but is both easier to state and to prove in our flat setting since [29] 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.

Proposition 4.3. *For every $M\in\mathbb{N}$, there exists $\delta_M, C_M>0$ such that, for every $\delta>0$, $T>0$, $f\in L^1((0,T);\mathcal{C}^{M^2}_{2\delta}\cap\mathcal{C}^1_{2\delta})$ with $\|f\|_{L^1_T(\mathcal{C}^{M^2})}\le\delta_M\min\{1;\delta\}$, $p\in B_\delta$ and $t\in[0,T]$ then*

$$\left|x(t;f,p)-e^{Z_M(t,f)}p\right|\le C_M\|f\|^{M+1}_{L^1_t(\mathcal{C}^{M^2})}, \tag{4.13}$$

where $Z_M(t,f):=\mathrm{Log}_M\{f\}(t)$ is the vector field introduced in Definition 2.19.

Hence, if $f\in L^\infty((0,T),\mathcal{C}^{M^2}_{2\delta})$ 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 $f_i\in\mathcal{C}^{M^2}_{2\delta}\cap\mathcal{C}^1_{2\delta}$, then, for each monomial basis $\mathcal{B}$ of $\mathcal{L}(X)$,

$$Z_M(t,f)=\sum_{b\in\mathcal{B}_{[\![1,M]\!]}}\zeta_b(t,u)f_b \tag{4.14}$$

where the functionals ζ_b are the associated coordinates of the first kind and f_b are the evaluated Lie brackets (see Definitions 2.27, 2.29 and 3.13).

Proof. For $M=0$, $Z_0(t,f)=0$ thus (4.13) holds with $C_0=1$ because $|x(t;f,p)-p|\le\|f\|_{L^1(\mathcal{C}^0)}$. From now on $M\in\mathbb{N}^*$ is fixed. By Definition 2.19, there exists $C'_M>0$ such that, for every $\delta>0$, $T>0$, $f\in L^1((0,T);\mathcal{C}^{M-1}_{2\delta})$ with $\|f\|_{L^1(\mathcal{C}^{M-1})}\le1$ and $t\in[0,T]$

$$\|\mathrm{Log}_M\{f\}(t)\|_{\mathcal{C}^0_{2\delta}}\le C'_M\|f\|_{L^1((0,t),\mathcal{C}^{M-1}_{2\delta})}. \tag{4.15}$$

In particular, for every $\delta>0$, $T>0$, $f\in L^1((0,T);\mathcal{C}^{M-1}_{2\delta})$ with $\|f\|_{L^1(\mathcal{C}^{M-1})}\le\min\{1;\delta;\delta/C'_M\}$, 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 $\|f\|_{L^1(\mathcal{C}^{M-1})}\le\delta_M\min\{1;\delta\}$ with $\delta_M:=\min\{1;1/C'_M\}$.

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

In order to use the operators $L(t)$ defined in Section 4.1, we assume moreover that $f \in \mathcal{C}_c^\infty([0,T]\times\mathbb{K}^d)$. This is not restrictive because $\mathcal{C}_c^\infty([0,T]\times\mathbb{K}^d)$ is dense in $L^1((0,T);\mathcal{C}^{M^2}_{2\delta})$ and both sides of (4.13) are continuous for the $L^1((0,T);\mathcal{C}^{M^2}_{2\delta})$ topology on f.

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 series (recall the formal series for $\log(1+x)$):

$$\log\, L(t) = \sum_{m\in\mathbb{N}^*} \frac{(-1)^{m-1}}{m}\left(\sum_{j\in\mathbb{N}^*} L_j(t)\right)^m \tag{4.16}$$

i.e. we define

$$Z_M(t,f) := \sum_{r=1}^{M}\sum_{m=1}^{M} \frac{(-1)^{m-1}}{m} \sum_{\mathbf{r}\in\mathbb{N}_r^m} L_{\mathbf{r}_m}(t)\cdots L_{\mathbf{r}_1}(t), \tag{4.17}$$

where $\mathbb{N}_r^m$ is defined in (2.13). For instance,

$$Z_3 = L_1 + \left(L_2 - \frac{1}{2}L_1^2\right) + \left(L_3 - \frac{1}{2}\left(L_1L_2 + L_2L_1\right) + \frac{1}{3}L_1^3\right). \tag{4.18}$$

Then, by (4.7),

$$Z_M(t,f) = \sum_{r=1}^{M}\sum_{m=1}^{M} \frac{(-1)^{m-1}}{m} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \int_{\mathcal{T}_{\mathbf{r}}(t)} (f(\tau_r)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\,\mathrm{d}\tau, \tag{4.19}$$

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 4.6 (see below in the next paragraph) and Definition 2.19, $Z_M(t,f) = \mathrm{Log}_M\{f\}(t)$ and satisfies (4.14). Thus $Z_M(t,f)$ is both a smooth vector field on $\mathbb{K}^d$ and a first-order differential operator, which we identify.

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 $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$,

$$\left|\left(L(t) - e^{Z_M(t,f)}\right)(\varphi)(p)\right| \le C_M \|f\|^{M+1}_{L^1(\mathcal{C}^{M^2})} \|\varphi\|_{\mathcal{C}^{M^2+1}}. \tag{4.20}$$

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 (4.20), we will decompose the difference in three terms

$$L - e^{Z_M} = \left(L - \sum_{j=0}^{M} L_j\right) + \left(\sum_{j=0}^{M} L_j - \sum_{k=0}^{M} \frac{Z_M^k}{k!}\right) + \left(\sum_{k=0}^{M} \frac{Z_M^k}{k!} - e^{Z_M}\right). \tag{4.21}$$

The first term is estimated in Proposition 4.2.

Step 3: Bound for $\sum L_j - \sum \frac{Z_M^k}{k!}$. By (4.17), this operator is a (finite) linear combination of terms of the form $L_{j_1}(t)\cdots L_{j_p}(t)$ where $p \in \mathbb{N}^*$, $j_1,\dots,j_p \in [\![1,M]\!]$ and $M+1 \le j_1+\ldots+j_p \le M^2$. Indeed, $Z_M(t,f)$ is also the finite sum of terms "of order at most M with respect to f" in the formal series (4.16). Thus, there exists $C''_M > 0$ (independent of δ, T, f) such that, for every $p \in B_\delta$, $t \in [0,T]$ and $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$,

$$\left|\left(\sum_{j=0}^{M} L_j(t) - \sum_{k=0}^{M} \frac{Z_M(t,f)^k}{k!}\right)(\varphi)(p)\right| \le C''_M \|f\|^{M+1}_{L^1(\mathcal{C}^{M^2-1})} \|\varphi\|_{\mathcal{C}^{M^2}}. \tag{4.22}$$

Step 4: Bound for $\sum \frac{Z_M^k}{k!} - e^{Z_M}$. Using Lemma 3.20 for the time-independent vector field $Z_M(t,f)$ (where $t \in [0,T]$ has been fixed), estimate (3.31) yields for every $p \in B_\delta$, $t \in [0,T]$ and $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$

$$\left| \left(e^{Z_M(t,f)} - \sum_{k=0}^{M} \frac{Z_M(t,f)^k}{k!} \right)(\varphi)(p) \right| \leq \|Z_M(t,f)\|_{\mathcal{C}_{2\delta}^M}^{M+1} \|\nabla \varphi\|_{\mathcal{C}_{2\delta}^M}. \tag{4.23}$$

We deduce from (2.24) the existence of $C_M''' > 0$ (independent of δ, T, f) such that for every $t \in [0,T]$

$$\|Z_M(t,f)\|_{\mathcal{C}_{2\delta}^M} \leq C_M''' \|f\|_{L^1((0,t),\mathcal{C}_{2\delta}^{2M-1})}. \tag{4.24}$$

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

$$\left| \left(e^{Z_M(t,f)} - \sum_{k=0}^{M} \frac{Z_M(t,f)^k}{k!} \right)(\varphi)(p) \right| \leq (C_M''')^{M+1} \|f\|_{L^1(\mathcal{C}^{2M-1})}^{M+1} \|\nabla \varphi\|_{\mathcal{C}^M}. \tag{4.25}$$

Gathering (4.8), (4.22) and (4.25) concludes the proof of (4.13). □

4.3.2 Campbell Baker Hausdorff Dynkin formula

We deduce from Proposition 4.3 the following estimate for the classical CBHD formula with q time-independent vector fields.

Corollary 4.4. *For every $M \in \mathbb{N}$, there exists $\delta_M, C_M > 0$ such that, for every $\delta > 0$, $q \in \mathbb{N}^*$, $f_1, \dots, f_q \in \mathcal{C}_{2\delta}^{M^2} \cap \mathcal{C}_{2\delta}^1$ with $\sum_{1 \leq j \leq q} \|f_j\|_{\mathcal{C}^{M^2}} \leq \delta_M \min\{1;\delta\}$,*

$$\left\| e^{f_q} \cdots e^{f_1} - e^{\mathrm{CBHD}_M(f_1,\dots,f_q)} \right\|_{\mathcal{C}_\delta^0} \leq C_M \|f\|^{M+1} \tag{4.26}$$

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 $\|f\| := \|f\|_{L^1(\mathcal{C}^{M^2})} = \sum_{1 \leq j \leq q} \|f_j\|_{\mathcal{C}^{M^2}}$.

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

$$\mathrm{CBHD}_M(f_1,\dots,f_q) = \sum_{b \in \mathcal{B}_{[\![1,M]\!]}} \alpha_b f_b \tag{4.27}$$

where $(\alpha_b)_{b \in \mathcal{B}} \subset \mathbb{K}^{\mathcal{B}}$ is given by Corollary 2.32.

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 (4.26) is an application of Proposition 4.3. 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

$$\mathrm{CBHD}_M(f_1,\dots,f_q) = \Lambda(\mathrm{CBHD}_M(X_1,\dots,X_q)) = \Lambda\left(\sum_{b \in \mathcal{B}_{[\![1,M]\!]}} \alpha_b b \right) = \sum_{b \in \mathcal{B}_{[\![1,M]\!]}} \alpha_b \Lambda(b), \tag{4.28}$$

which proves (4.27). □

4.3.3 Replacing products with brackets in logarithm integrals

The goal of this section is to prove Lemma 4.6, which is a key point in the proof of Proposition 4.3, 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 2.26 in algebras. Indeed, Theorem 2.26 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 4.5. *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*

$$\begin{aligned}\sum_{m=1}^{r} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\mathcal{T}_{\mathbf{r}}(t)} \mathrm{a}(\tau_r)\mathrm{a}(\tau_{r-1})\cdots\mathrm{a}(\tau_1)\,\mathrm{d}\tau = \\ \frac{1}{r}\sum_{m=1}^{r} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\mathcal{T}_{\mathbf{r}}(t)} [\cdots[\mathrm{a}(\tau_r),\mathrm{a}(\tau_{r-1})],\ldots\mathrm{a}(\tau_1)]\,\mathrm{d}\tau,\end{aligned} \tag{4.29}$$

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 $\mathcal{L}_r(X)$,

$$\frac{1}{r}\sum_{m=1}^{r} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\mathcal{T}_{\mathbf{r}}(t)} [\cdots[\mathrm{a}(\tau_r),\mathrm{a}(\tau_{r-1})],\ldots\mathrm{a}(\tau_1)] = \sum_{b\in\mathcal{B}_r} \zeta_b(t,\alpha)\mathrm{y}_b, \tag{4.30}$$

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 algebra such that $\Upsilon(X_i) = \mathrm{y}_i$ (see Definition 2.29 and Lemma 2.7).

Proof. Let $q \in \mathbb{N}^*$ be the dimension of A_1 (as a linear subspace) and $\mathrm{y}_1,\ldots\mathrm{y}_q$ 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 $\mathrm{y}_1,\ldots\mathrm{y}_q$, 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 (2.33) and (2.34), one obtains that (4.29) holds for $a(\cdot)$. Applying the homomorphism of algebra Υ to both sides proves (4.29) for $\mathrm{a}(\cdot)$. The same strategy proves (4.30). □

Lemma 4.6. *For every $r \in \mathbb{N}^*$, $t > 0$ and $f \in L^1([0,t];\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K}^d))$*

$$\begin{aligned}\sum_{m=1}^{r} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\mathcal{T}_{\mathbf{r}}(t)} (f(\tau_r)\cdot\nabla)(f(\tau_{r-1})\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\,\mathrm{d}\tau = \\ \frac{1}{r}\sum_{m=1}^{r} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\mathcal{T}_{\mathbf{r}}(t)} [\cdots[f(\tau_r)\cdot\nabla, f(\tau_{r-1})\cdot\nabla],\ldots f(\tau_1)\cdot\nabla]\,\mathrm{d}\tau,\end{aligned} \tag{4.31}$$

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 real numbers.

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

$$\begin{aligned}\frac{1}{r}\sum_{m=1}^{r} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\mathcal{T}_{\mathbf{r}}(t)} [\cdots[f(\tau_r)\cdot\nabla, f(\tau_{r-1})\cdot\nabla],\ldots f(\tau_1)\cdot\nabla]\,\mathrm{d}\tau \\ = \sum_{b\in\mathcal{B}_r} \zeta_b(t,u)f_b,\end{aligned} \tag{4.32}$$

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 2.4, 2.29 and 3.13).

Proof. Let $(f_n)_{n\in\mathbb{N}^*}$ be a sequence of functions in $L^1([0,t];\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K}^d))$ 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 4.5 with $A = \mathrm{Op}(\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K}))$ and $A_1 = \mathrm{span}\{f_i\}_{i\in I}$ (a vector field g is identified with the first order operator $g\cdot\nabla$ on $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$) proves (4.31) 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

$$\begin{aligned}\sum_{m=1}^{r}\sum_{\mathbf{r}\in\mathbb{N}_r^m}\frac{(-1)^{m-1}}{m}\int_{\mathcal{T}_{\mathbf{r}}(t)}\left((f_n(\tau_r)\cdot\nabla)\cdots(f_n(\tau_1)\cdot\nabla)\varphi\right)(p)\,\mathrm{d}\tau = \\ \frac{1}{r}\sum_{m=1}^{r}\sum_{\mathbf{r}\in\mathbb{N}_r^m}\frac{(-1)^{m-1}}{m}\int_{\mathcal{T}_{\mathbf{r}}(t)}\left([\cdots[f_n(\tau_r)\cdot\nabla, f_n(\tau_{r-1})\cdot\nabla],\ldots f_n(\tau_1)\cdot\nabla]\varphi\right)(p)\,\mathrm{d}\tau.\end{aligned} \tag{4.33}$$

For each fixed φ and p, both sides converge as $n \to +\infty$ towards the same quantities for f. This proves that (4.31) holds as an equality between linear operators. Applying (4.30) gives (4.32). □

Remark 4.7. *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 4.6, 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 [61, 62], 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

$$\dot{x}(t) = f_0(x) + f_1(t,x) \tag{4.34}$$

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_1, 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. [15, Section 3.6] or [16, Section 2]).

4.4.1 Error bound

Proposition 4.8. *For every $M \in \mathbb{N}$, there exists $C_M > 0$ and $\Theta_M \in \mathcal{C}^0(\mathbb{R}_+^2;\mathbb{R}_+^*)$ such that, for every $\delta > 0$, $T > 0$, $f_0 \in \mathcal{C}_{5\delta}^{M^2+1}$ with $T\|f_0\|_{\mathcal{C}^0} < \delta$, $f_1 \in L^1((0,T);\mathcal{C}_{5\delta}^{M^2})$ with*

$$\|f_1\|_{L^1(\mathcal{C}^{M^2})} \le \Theta_M(T, \|f_0\|_{\mathcal{C}^{M^2+1}})\min\{1;\delta\}, \tag{4.35}$$

$p \in B_\delta$ and $t \in [0,T]$ then

$$\left|x(t;f_0+f_1,p) - e^{\mathcal{Z}_M(t,f_0,f_1)}e^{tf_0}p\right| \le C_M\|g_t\|_{L_i^1(\mathcal{C}^{M^2})}^{M+1} \tag{4.36}$$

where $\mathcal{Z}_M(t,f_0,f_1) = \mathrm{Log}_M\{g_t\}(t)$ in the sense of Definition 2.19, $g_t : [0,t]\times B_{4\delta} \mapsto \mathbb{K}^d$ is defined by

$$g_t(\tau,y) := (\Phi_0(t-\tau)_* f_1(\tau))(y) = \partial_p\Phi_0(\tau-t,y)^{-1} f_1\big(\tau,\Phi_0(\tau-t,y)\big) \tag{4.37}$$

and $\Phi_0 : [0,T]\times B_{4\delta} \to B_{5\delta}$ is the flow associated with f_0 i.e. $\Phi_0(\tau;p) = e^{\tau f_0}(p)$.

Hence, if $f_1 \in L^\infty((0,T);\mathcal{C}_{5\delta}^{M^2})$, estimate (4.36) corresponds to a bound scaling like t^{M+1}.

Proof. Let $M \in \mathbb{N}$, $\delta > 0$, $T > 0$, $f_0 \in \mathcal{C}_{5\delta}^{M^2+1}$ with $T\|f_0\|_{\mathcal{C}^0} < \delta$ and $f_1 \in L^1((0,T);\mathcal{C}_{5\delta}^{M^2})$ such that $\|f_1\|_{L^1((0,T);\mathcal{C}^0)} \le \delta$. Then, for every $p \in B_\delta$ and $\tau \in [0,T]$, $x(\tau;f_0+f_1,p)$ is well defined and

belongs to $B_{3\delta}$. To simplify the notations in this proof, we write $x(\tau)$ instead of $x(\tau; f_0 + f_1, p)$. Let $t \in [0, T]$. The function $y : [0, t] \to \mathbb{K}^d$ defined by

$$y(\tau) := \Phi_0\big(t - \tau; x(\tau)\big) \tag{4.38}$$

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

$$\dot{y}(\tau) = g_t\left(\tau, y(\tau)\right). \tag{4.39}$$

By (4.37), there exists $\Phi_M \in \mathcal{C}^0(\mathbb{R}_+^2; \mathbb{R}_+^*)$ (independent of δ, T, f_0, f_1) such that

$$\|g_t\|_{L^1((0,t);\mathcal{C}^{M^2}_{4\delta})} \leq \Phi_M\left(T, \|f_0\|_{\mathcal{C}^{M^2+1}_{5\delta}}\right) \|f_1\|_{L^1((0,t);\mathcal{C}^{M^2}_{5\delta})}. \tag{4.40}$$

Let us assume that (4.35) holds with $\Theta_M(T, a) := \min\{1; \delta_M/\Phi_M(T, a)\}$ and δ_M as in Proposition 4.3. This implies $\|f_1\|_{L^1((0,T);\mathcal{C}^0)} \leq \delta$. Moreover, one has $\|g_t\|_{L^1((0,t);\mathcal{C}^{M^2})} \leq \delta_M \min\{1; 4\delta\}$. Thus, by Proposition 4.3

$$|y(t) - e^{\mathcal{Z}_M(t,f_0,f_1)} y(0)| \leq C_M \|g_t\|^{M+1}_{L^1((0,t);\mathcal{C}^{M^2})}, \tag{4.41}$$

which is exactly (4.36) because $y(t) = x(t)$ and $y(0) = e^{tf_0}p$. □

4.4.2 Expansions of $\mathcal{Z}_M$

Proposition 4.9. *Let $r > 0$ and $M, \delta, T, f_0, f_1, g_t, \mathcal{Z}_M$ be as in Proposition 4.8. If $f_0 \in \mathcal{C}^{\omega,r}_{5\delta}$ and $f_1 \in \mathcal{C}^0([0, T]; \mathcal{C}^{\omega,r}_{5\delta})$ then, for $0 \leq \tau \leq t \leq \min\{T; \frac{r}{9|||f_0|||_r}\}$*

$$g_t(\tau, \cdot) = e^{(\tau - t)\,\mathrm{ad}_{f_0}}(f_1(\tau)) = \sum_{k=0}^{+\infty} \frac{(\tau - t)^k}{k!}\,\mathrm{ad}^k_{f_0}(f_1(\tau)) \tag{4.42}$$

and

$$\begin{aligned}\mathcal{Z}_M(t, f_0, f_1) = \sum \frac{(-1)^{m-1}}{rm} \int_{\mathcal{T}_{\mathbf{r}}(t)} \frac{(\tau_1 - t)^{k_1}}{k_1!} \cdots \frac{(\tau_r - t)^{k_r}}{k_r!} \\ \Big[\cdots\Big[\mathrm{ad}^{k_r}_{f_0}(f_1(\tau_r)), \mathrm{ad}^{k_{r-1}}_{f_0}(f_1(\tau_{r-1}))\Big], \ldots, \mathrm{ad}^{k_1}_{f_0}(f_1(\tau_1))\Big]\,\mathrm{d}\tau\end{aligned} \tag{4.43}$$

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 (4.42) and (4.43) converge absolutely in $\mathcal{C}^{\omega,r'}_{5\delta}$.

Proof. We apply the third statement of Lemma 3.24 to f_0 and $f_1(\tau)$ to get (4.42). The absolute convergence in this series allows to interchange the sums and the integrals. □

When the perturbation $f_1(t, x)$ is affine, i.e. of the form $\sum_{i=1}^q u_i(t) f_i(x)$, by analogy with Theorem 2.39, 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 4.10. *Let $r > 0$ and $M, \delta, T, f_0, f_1, g_t, \mathcal{Z}_M$ be as in Proposition 4.8. We assume $f_0 \in \mathcal{C}^{\omega,r}_{5\delta}$ and $f_1(t, x) = \sum_{i=1}^q u_i(t) f_i(x)$ where $u_i \in L^1(0, T)$ and $f_i \in \mathcal{C}^{\omega,r}_{5\delta}$. Then*

$$\mathcal{Z}_M(t, f, u) = \lim_{N \to \infty} \sum_{\substack{b \in \mathcal{B} \\ n(b) \leq M \\ n_0(b) \leq N}} \eta_b(t, u) f_b \tag{4.44}$$

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

Proof. Let $X = \{X_0, X_1, \dots X_q\}$ and $\Lambda : \mathcal{L}(X) \to \mathcal{C}^{\omega,r}_{5\delta}$ be the homomorphism of Lie algebra such that $\Lambda(X_i) = f_i$ for $i \in [\![0, q]\!]$ (see Lemma 2.7). By applying Λ to each term in the equality (2.82) (where $\mathcal{Z}^{r,\nu}_\infty(t, X, a)$ is the finite sum defined in (2.80)), we obtain for every $r \in \mathbb{N}^*$ and $\nu \in \mathbb{N}$

$$\mathcal{Z}^{r,\nu}_\infty(t, f, u) = \sum_{b \in \mathcal{B}_{r,\nu}} \eta_b(t, u) f_b. \tag{4.45}$$

By Proposition 4.9

$$\mathcal{Z}_M(t, f, u) = \lim_{N\to\infty} \sum_{\nu=0}^{N} \sum_{r=1}^{M} \mathcal{Z}^{r,\nu}_\infty(t, f, u) \tag{4.46}$$

where for every $r' \in [r/e, r]$ the limit holds in $\mathcal{C}^{\omega,r'}_{5\delta}$ when $0 \le t \le \min\{T; \frac{r-r'}{6|||f_0|||_r}\}$. This proves (4.44). □

Remark 4.11. *Although the family $\eta_b(t, u)f_b$ for $b \in \mathcal{B} \cap S_M$ is not proved to be absolutely summable, equality (4.44) gives a sense to the expression*

$$\mathcal{Z}_M(t, f, u) = \sum_{b \in \mathcal{B} \cap S_M} \eta_b(t, u) f_b. \tag{4.47}$$

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 $b \in \mathcal{B} \cap S_M$.

Proposition 4.12. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$ and $\mathcal{B}$ be a monomial basis of $\mathcal{L}(X)$ with geometric growth with respect to X_0 (see Definition 2.45). Let $M \in \mathbb{N}$, $r > 0$ and $r' \in [r/e, r)$. There exists $T^* = T^*(M, q, r, r') > 0$ such that, for every $\delta, T, f_0, f_1, \dots, f_q, u_1, \dots, u_q$ as in Proposition 4.10 (in particular $f_0, \dots, f_q \in \mathcal{C}^{\omega,r}_{5\delta}$) and $t \in [0, T^*]$*

$$\mathcal{Z}_M(t, f, u) = \sum_{b \in \mathcal{B} \cap S_M} \eta_b(t, u) f_b \tag{4.48}$$

where the series converges absolutely in $\mathcal{C}^{\omega,r'}_{5\delta}$.

Proof. By (2.88) of Proposition 2.48 and (3.21), for every $b \in \mathcal{B} \cap S_M$ and $t \in [0, T]$

$$|\eta_b(t; u)| \, |||f_b|||_{r'} \le \frac{r - r'}{2e^2} \left(\frac{2eC_M t}{r - r'}\right)^{n_0(b)} \left(\frac{2eC_M}{r - r'} \|u\|_{L^1(0,t)} \, |||f|||_r\right)^{n(b)} \tag{4.49}$$

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

$$\sum_{b \in \mathcal{B} \cap S_M} (2(q + 1))^{-n_0(b)} \le \sum_{n=1}^{M} \sum_{n_0=0}^{+\infty} (q + 1)^{n_0+n} (2(q + 1))^{-n_0} \le M(q + 1)^M. \tag{4.50}$$

□

4.5 Sussmann's infinite product expansion

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

$$\dot{x}(t) = \sum_{i \in I} u_i(t) f_i(x(t)) \quad \text{and} \quad x(0) = p, \tag{4.51}$$

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 4.13. *Let $\mathcal{B}$ be a generalized 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 $T, \delta > 0$, $f_i \in \mathcal{C}^{2M}_{3\delta}$ and $u_i \in L^1((0,T);\mathbb{K})$ for $i \in I$. Assume that*

$$\|u\|_{L^1_T} \|f\|_{\mathcal{C}^M} \leq \eta_M \min\{1, \delta\}. \tag{4.52}$$

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

$$\left| x(t; f, u, p) - \overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t;u) f_b} p \right| \leq C_M \|u\|^{M+1}_{L^1_t} \|f\|^{M+1}_{\mathcal{C}^{2M}} \left(1 + \|f\|^{M-1}_{\mathcal{C}^{2M}}\right), \tag{4.53}$$

where the arrow above the product symbol designates the order for the product, i.e. with the notations of Definition 2.49

$$\overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t;u) f_b} = e^{\xi_{b_1}(t;u) f_{b_1}} \cdots e^{\xi_{b_{k+1}}(t;u) f_{b_{k+1}}}. \tag{4.54}$$

Proof. Let $M \in \mathbb{N}^*$. We adopt the notations $b_1, \ldots, b_{k+1}$ and $Y_0, \ldots, Y_{k+1}$ of Definition 2.49. For $j \in [\![1, k+1]\!]$, we denote by Φ_j the flow associated with f_{b_j}, i.e. $\Phi_j(t,p) := e^{t f_{b_j}}(p)$. 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 (4.52),

$$\left\| \sum_{i\in I} u_i f_i \right\|_{L^1_T(\mathcal{C}^0)} \leq \eta_M \min\{1,\delta\} \leq \delta. \tag{4.55}$$

Thus, for $t \in [0,T]$, $x(t)$ is well-defined and $x(t) \in B_{2\delta}$. For $b \in \mathcal{B}$, using (2.102), Lemma 3.14 and the crude estimate $|\mathcal{B}_\ell| \leq |I|^\ell$, we obtain, for each $t \in [0,T]$,

$$\sum_{b\in\mathcal{B}_{[\![1,M]\!]}} \|\xi_b(t) f_b\|_{\mathcal{C}^1} \leq \sum_{\ell=1}^{M} |I|^\ell \|u\|^\ell_{L^1_t} 2^\ell \ell! \|f\|^\ell_{\mathcal{C}^\ell} \leq \frac{2M! |I| \|u\|_{L^1_T} \|f\|_{\mathcal{C}^M}}{1 - 2|I|\|u\|_{L^1_T}\|f\|_{\mathcal{C}^M}} \leq \min(\delta, 1). \tag{4.56}$$

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

$$x_j(t) := e^{-\xi_{b_j}(t) f_{b_j}} \cdots e^{-\xi_{b_1}(t) f_{b_1}}(x(t)) \tag{4.57}$$

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

$$(\mathcal{H}_j): \begin{cases} \dot{x}_j(t) = \sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_j} \dot{\xi}_b(t) f_b(x_j(t)) + \varepsilon_j(t), \\ x_j(0) = p, \end{cases} \tag{4.58}$$

where

$$|\varepsilon_j(t)| \leq C_j |u(t)| \|u\|^M_{L^1_t} F^{M+1}(1 + F^{M-1}). \tag{4.59}$$

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

$$x_j(t) = e^{-\xi_{b_j}(t) f_{b_j}}(x_{j-1}(t)) = \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) \tag{4.60}$$

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

$$\dot{x}_j(t) = -\dot{\xi}_{b_j}(t) f_{b_j}(x_j(t)) + \sum_{b \in \mathcal{B}_{[\![1,M]\!]} \cap Y_{j-1}} \dot{\xi}_b(t) \partial_p \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) f_b(x_{j-1}(t)) + \widetilde{\varepsilon}_j(t) \tag{4.61}$$

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 (4.56), $\|\xi_{b_j}(t) f_{b_j}\|_{\mathcal{C}^1} \leq 1$, so, using (3.28),

$$|\widetilde{\varepsilon}_{j-1}(t)| \leq e|\varepsilon_{j-1}(t)|. \tag{4.62}$$

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

$$\partial_p \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) f_b(x_{j-1}(t)) = \left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)), \tag{4.63}$$

thus,

$$\dot{x}_j(t) = \sum_{b \in \mathcal{B}_{[\![1,M]\!]} \cap Y_{j-1} \setminus \{b_j\}} \dot{\xi}_b(t) \left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)) + \widetilde{\varepsilon}_j(t). \tag{4.64}$$

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

$$|b| + (h(b) - 1)|b_j| \leq M. \tag{4.65}$$

Then, by the first statement of Lemma 3.24 and Definition 2.57

$$\begin{aligned} \dot{\xi}_b(t) \left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)) &= \sum_{m=1}^{h(b)-1} \frac{\xi_{b_j}^m(t)}{m!} \dot{\xi}_b(t) f_{\mathrm{ad}_{b_j}^m(b)}(x_j(t)) + \overline{\varepsilon}_b^j(t) \\ &= \sum_{m=1}^{h(b)-1} \dot{\xi}_{\mathrm{ad}_{b_j}^m(b)}(t) f_{\mathrm{ad}_{b_j}^m(b)}(x_j(t)) + \overline{\varepsilon}_b^j(t) \end{aligned} \tag{4.66}$$

where

$$|\overline{\varepsilon}_b^j(t)| \leq |\dot{\xi}_b(t)| \frac{|\xi_{b_j}(t)|^{h(b)}}{h(b)!} \|f_{\mathrm{ad}_{b_j}^{h(b)}(b)}\|_{\mathcal{C}^0}. \tag{4.67}$$

By definition of $h(b)$ we have $M + 1 \leq |b| + h(b)|b_j| \leq M + |b_j| \leq 2M$. Thus, using Lemma 3.14, (2.101) and (2.102), we get

$$\begin{aligned} |\overline{\varepsilon}_b^j(t)| &\leq |u(t)| \|u\|_{L_t^1}^{|b|+h(b)|b_j|-1} \frac{|b|}{h(b)!} 2^{2M} (2M-1)! F^{M+1}(1 + F^{M-1}) \\ &\leq |u(t)| \|u\|_{L_t^1}^{M} M 2^{2M} (2M-1)! F^{M+1}(1 + F^{M-1}). \end{aligned} \tag{4.68}$$

By definition of Y_j in Definition 2.49, we obtain $(\mathcal{H}_j)$ with

$$\varepsilon_j(t) := \widetilde{\varepsilon}_{j-1}(t) + \sum_{b \in \mathcal{B}_{[\![1,M]\!]} \cap Y_{j-1} \setminus \{b_j\}} \overline{\varepsilon}_b^j(t). \tag{4.69}$$

that satisfies (4.59) with, for instance $C_{j+1} := eC_j + |I|^{M+1} M 2^{2M} (2M-1)!$.

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 $|x_{k+1}(t) - p| \leq C_{k+1} \|u\|_{L_t^1}^{M+1} F^{M+1}(1 + F^{M-1})$, i.e.

$$\left| \overleftarrow{\prod_{b \in \mathcal{B}_{[\![1,M]\!]}}} e^{-\xi_b(t;u) f_b} x(t) - p \right| \leq C_{k+1} \|u\|_{L_t^1}^{M+1} F^{M+1}(1 + F^{M-1}). \tag{4.70}$$

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 (4.53) holds. Note that (4.56) and (3.28) 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 2.5 and Lemma 3.17). 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
Iterated Duhamel or 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 Iterated Duhamel or Chen-Fliess expansion

5.1.1 Counter-example for nilpotent vector fields

As already discussed in Remark 2.16, the iterated Duhamel or 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-nilpotent 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 5.1. *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 (3.12) with $f(t,x):=f_0(x)$ satisfies*

$$\lim_{N\to+\infty}\left|x(t;f,0)-\sum_{n=0}^{N}\frac{t^n}{n!}\left((f_0\cdot\nabla)^n\mathrm{Id}_1\right)(0)\right|=+\infty. \tag{5.1}$$

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

$$\forall n \geq 2, \quad ((f_\alpha \cdot \nabla)^n \mathrm{Id}_1)(0) = \alpha_n. \tag{5.2}$$

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

$$((f_\alpha \cdot \nabla)^n \mathrm{Id}_1)(0) = f_\alpha^{(n-1)}(0) + P_n\left(f_\alpha(0), \dots, f_\alpha^{(n-2)}(0)\right), \tag{5.3}$$

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 (5.2). 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 (5.2), for each $t > 0$

$$\sum_{n=0}^{N} \frac{t^n}{n!} ((f_0 \cdot \nabla)^n \mathrm{Id}_1)(0) = \sum_{n=0}^{N} n! t^n \to +\infty, \tag{5.4}$$

which proves (5.1). □

Remark 5.2. *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

$$\dot{x}(t) = A(t)x(t) \quad \text{and} \quad x(0) = p, \tag{5.5}$$

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

Proposition 5.3. *Let $T > 0$ and $A \in L^1((0,T);\mathcal{M}_d(\mathbb{K}))$. For each $t \in [0,T]$ and $p \in \mathbb{K}^d$,*

$$x(t; A, p) = p + \sum_{j=1}^{+\infty} \int_{0<\tau_1<\cdots<\tau_j<t} A(\tau_j) \cdots A(\tau_1) p \, \mathrm{d}\tau, \tag{5.6}$$

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

$$x(\tau) = p + \int_0^\tau A(\tau')x(\tau') \, \mathrm{d}\tau' \tag{5.7}$$

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

$$\begin{aligned} \left| x(t) - p - \sum_{j=1}^{M-1} \int_{0<\tau_M<\cdots<\tau_1<t} A(\tau_j) \cdots A(\tau_1) p \, \mathrm{d}\tau \right| &= \left| \int_{0<\tau_1<\cdots<\tau_M<t} A(\tau_M) \cdots A(\tau_1) x(\tau) \, \mathrm{d}\tau \right| \\ \leq \int_{0<\tau_1<\cdots<\tau_M<t} \|A(\tau_M)\| \cdots \|A(\tau_1)\| \, \mathrm{d}\tau |p| e^{\|A\|_{L^1(0,t)}} &= \frac{\|A\|_{L^1(0,t)}^M}{M!} |p| e^{\|A\|_{L^1(0,t)}} \end{aligned} \tag{5.8}$$

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 iterated Duhamel or Chen-Fliess series converges locally in time (see e.g. [67, Proposition 4.3] for slightly different assumptions).

Proposition 5.4. *Let $T, \delta, r > 0$. There exists $\eta > 0$ such that, for each $f \in L^1([0,T]; \mathcal{C}^{\omega,r}_{2\delta})$ with $\|f\|_{L^1_t(\mathcal{C}^{\omega,r})} \leq \eta$, for each $\varphi \in \mathcal{C}^{\omega,r}_{2\delta}$, $t \in [0,T]$ and $p \in B_\delta$,*

$$\varphi\left(x(t;f,p)\right) = \varphi(p) + \sum_{j=1}^{+\infty} \int_{0<\tau_1<\cdots<\tau_j<t} \left((f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\right)(\varphi)(p)\,\mathrm{d}\tau, \tag{5.9}$$

where the sum converges absolutely. In particular,

$$x(t;f,p) = p + \sum_{j=1}^{+\infty} \int_{0<\tau_1<\cdots<\tau_j<t} \left((f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\right)(\mathrm{Id}_d)(p)\,\mathrm{d}\tau. \tag{5.10}$$

Proof. Let $\eta := \min\{\delta/2, r/10\}$. By Lemma 3.9, $x(t;f,p)$ is well defined for $t \in [0,T]$, $p \in B_\delta$ and belongs to $B_{2\delta}$. Moreover, by Lemma 3.11, we have, for every $j \in \mathbb{N}^*$

$$\int_{0<\tau_1<\cdots<\tau_j<t} \left|\left((f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\right)(\varphi)(p)\right|\,\mathrm{d}\tau \leq j!\left(\frac{5}{r}\right)^j \cdot \frac{\|f\|^j}{j!}\,|||\varphi|||_r\,, \tag{5.11}$$

where $\|f\| = \|f\|_{L^1_t(\mathcal{C}^{\omega,r})}$, which proves the absolute convergence because the right-hand side is bounded by $2^{-j}\,|||\varphi|||_r$. Eventually, we deduce from (4.10) and Lemma 3.11 that

$$\left|\varphi\left(x(t;f,p)\right) - \sum_{j=0}^{M-1} \int_{0<\tau_1<\cdots<\tau_j<t} \left((f(\tau_j)\cdot\nabla)\cdots(f(\tau_1)\cdot\nabla)\right)(\varphi)(p)\,\mathrm{d}\tau\right| \leq 2^{-M}\,|||\varphi|||_r\,, \tag{5.12}$$

which proves (5.9). □

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 5.6).

If the vector fields are analytic in space, a simple proof can be given (see e.g. [43, 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 nilpotent assumption, $Z_M = Z_{M'}$ for every $M' \geq M$ and estimate (4.13) 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 [3, 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 2.7).

To that end, we start with the following technical result about formal series.

Lemma 5.5. *The following statements hold.*

1. *Let $T > 0$ and $z \in \mathcal{C}^1([0,T]; \mathcal{L}(X))$. Then the following equality holds in $\widehat{\mathcal{L}}(X)$*

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(z(t)) = \exp(z(t)) \sum_{n=0}^{+\infty} \frac{(-1)^n}{(n+1)!}\,\mathrm{ad}^n_{z(t)}(\dot{z}(t)). \tag{5.13}$$

2. *Let a be given by (2.7) and $Z_\infty(t) := \mathrm{Log}_\infty\{a\}(t)$ with the notation of Definition 2.19. Then for every $t \in \mathbb{R}$, the following equality holds in $\widehat{\mathcal{L}}(X)$*

$$\sum_{n=0}^{+\infty} \frac{(-1)^n}{(n+1)!} \mathrm{ad}^n_{Z_\infty(t)}(\dot{Z}_\infty(t)) = a(t). \tag{5.14}$$

3. *Let a be given by (2.7), $M \in \mathbb{N}^*$, $Z_M(t) := \mathrm{Log}_M\{a\}(t)$ with the notation of Definition 2.19. Then for every $t \in \mathbb{R}$, the following equality holds in $\mathcal{N}_{M+1}(X)$*

$$\sum_{n=0}^{M-1} \frac{(-1)^n}{(n+1)!} \mathrm{ad}^n_{Z_M(t)}(\dot{Z}_M(t)) = a(t), \tag{5.15}$$

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

Proof. We prove each claim separately.

1. We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\exp(z(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=0}^{+\infty}\frac{z^k(t)}{k!}\right) = \sum_{k=0}^{+\infty}\frac{1}{(k+1)!}\sum_{j=0}^{k} z^j(t)\dot{z}(t)z^{k-j}(t) \\ &= \exp(z(t))\left(\sum_{l=0}^{+\infty}\frac{(-1)^l}{l!}z^l(t)\right)\left(\sum_{k=0}^{+\infty}\frac{1}{(k+1)!}\sum_{j=0}^{k} z^j(t)\dot{z}(t)z^{k-j}(t)\right).\end{aligned} \tag{5.16}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(z(t)) = \exp(z(t))\sum_{n=0}^{+\infty}\frac{1}{(n+1)!}\sum_{i=0}^{n} z^i(t)\dot{z}(t)z^{n-i}(t)\sum_{l=0}^{i}(-1)^l\binom{n+1}{l} \tag{5.17}$$

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

$$\mathrm{ad}^n_z(y) = (-1)^n\sum_{i=0}^{n}(-1)^i\binom{n}{i}z^i y z^{n-i}, \tag{5.18}$$

$$\sum_{l=0}^{i}(-1)^l\binom{n+1}{l} = (-1)^i\binom{n}{i} \tag{5.19}$$

   give the conclusion.

2. By Theorem 2.26, the solution $x(t)$ of the formal ODE $\dot{x}(t) = x(t)a(t)$ is $x(t) = e^{Z_\infty(t)}$. We conclude by identifying the two expressions for $\frac{\mathrm{d}}{\mathrm{d}t}\left[e^{Z_\infty}(t)\right]$.

3. The canonical surjection $\sigma_{M+1} : \mathcal{L}(X) \to \mathcal{N}_{M+1}(X)$ is a Lie-algebra homomorphism. Applying to (5.14) gives (5.15).

□

This technical lemma leads to the main result of this section.

Proposition 5.6. *Let $M \in \mathbb{N}^*$. There exists $\eta_M > 0$ such that, for every $T, \delta > 0$ and $f : [0,T] \to \mathcal{C}^\infty_{4\delta}$ such that $\mathcal{L}(f([0,T]))$ is nilpotent with index at most $M+1$ and $f \in L^1((0,T);\mathcal{C}^{M+1}_{4\delta})$ with*

$$\|f\|_{L^1(\mathcal{C}^M)} + \|f\|^M_{L^1(\mathcal{C}^M)} \leq \eta_M\delta, \tag{5.20}$$

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 2.19.

Proof. Let $M \in \mathbb{N}^*$. By Definition 2.19, there exists $\eta_M > 0$ such that, for every t, f,

$$\|Z_M(t,f)\|_{\mathcal{C}^0} \leq \frac{1}{\eta_M}\left(\|f\|_{L^1(\mathcal{C}^M)} + \|f\|^M_{L^1(\mathcal{C}^M)}\right). \tag{5.21}$$

Step 1: Proof for $f(t,x) = \sum_{j=1}^q a_j(t) f_j(x)$ with $q \in \mathbb{N}^*$, $a_j \in L^1([0,T];\mathbb{K})$ 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 $p \in B_\delta$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{Z_M(t,f)}(p)) = f\left(t, e^{Z_M(t,f)}(p)\right). \tag{5.22}$$

By Definition 2.19, (5.20) and (5.21), the map $(t,p) \mapsto Z_M(t,f)(p)$ belongs to $\mathcal{C}^\infty([0,T]\times B_{4\delta};\mathbb{K}^d)$ and $\|Z_M(\cdot,f)\|_{\mathcal{C}^0} \leq \delta$. Thanks to the nilpotent 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 3.25 yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{Z_M(t,f)}(p)\right) = \sum_{k=0}^M \frac{(-1)^k}{(k+1)!}\,\mathrm{ad}^k_{Z_M(t,f)}\left(\dot{Z}_M(t,f)\right)\left(e^{Z_M(t,f)}(p)\right). \tag{5.23}$$

Let $\Lambda : \mathcal{N}_{M+1}(X) \to \mathcal{L}(f_1,\dots,f_q)$ be the homomorphism of nilpotent Lie algebra such that $\Lambda(X_j) = f_j$ for $j = 1,\dots,q$. By applying Λ to the equality (5.15), 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 $f([0,T])$, uniformly bounded in $L^1((0,T);\mathcal{C}^{M+1}_{4\delta})$ and such that $f_n \to f$ in $L^1(\mathcal{C}^0)$. We get the conclusion by passing to the limit in both sides. □

5.2.2 Convergence for linear systems

In this paragraph, we consider linear systems of the form (5.5). 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 [65], $T\|A\|_{L^\infty(0,T)} \leq 1$ is shown to be a sufficient condition for convergence on $[0,T]$ thanks to a careful estimate of the combinatorial terms. In [55], $\|A\|_{L^1(0,T)} < \pi$ 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 useful later in the article. Let $\|\cdot\|$ be a sub-multiplicative norm on $\mathcal{M}_d(\mathbb{K})$.

Proposition 5.7. *Let $T > 0$ and $A \in L^1((0,T);\mathcal{M}_d(\mathbb{K}))$ such that $\|A\|_{L^1_T} < \frac{1}{4}$. For each $t \in [0,T]$,*

$$Z_\infty(t) := \sum_{r=1}^{+\infty} \frac{1}{r} \sum_{m=1}^{r} \frac{(-1)^{m-1}}{m} \sum_{\mathbf{r}\in\mathbb{N}^m_r} \int_{\tau\in\mathcal{T}_{\mathbf{r}}(t)} [\cdots[A(\tau_r), A(\tau_{r-1})], \dots A(\tau_1)]\,\mathrm{d}\tau \tag{5.24}$$

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 $\mathbf{r} \in \mathbb{N}^m_r$,

$$\begin{aligned}&\int_{\tau\in\mathcal{T}_{\mathbf{r}}(t)} \|[\cdots[A(\tau_r), A(\tau_{r-1})], \dots A(\tau_1)]\|\,\mathrm{d}\tau \\ &\qquad \leq \int_{\tau\in\mathcal{T}_{\mathbf{r}}(t)} 2^r \|A(\tau_1)\|\dots\|A(\tau_{r-1})\|\|A(\tau_r)\|\,\mathrm{d}\tau \leq 2^r\left(\int_0^t \|A(\tau)\|\,\mathrm{d}\tau\right)^r.\end{aligned} \tag{5.25}$$

Moreover, recalling the definition of (2.13), $|\mathbb{N}_r^m| = \binom{r-1}{m-1}$ and $\sum_{m=1}^r \binom{r-1}{m-1} = 2^{r-1}$. Thus,

$$\sum_{r=1}^{+\infty} \frac{1}{r} \sum_{m=1}^{r} \frac{1}{m} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \int_{\tau\in\mathcal{T}_{\mathbf{r}}(t)} \|[\cdots[A(\tau_r), A(\tau_{r-1})], \dots A(\tau_1)]\| \,\mathrm{d}\tau \leq \sum_{r=1}^{+\infty} (4\|A\|_{L^1})^r < \infty. \tag{5.26}$$

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

$$\begin{cases} L'(\tau) = L(\tau)A(\tau) \\ L(0) = \mathrm{Id}_d. \end{cases} \tag{5.27}$$

By working as in the proof of Proposition 5.3, we obtain

$$L(t) = \mathrm{Id}_d + \sum_{r=1}^{+\infty} \int_{0<\tau_r<\cdots<\tau_1<t} A(\tau_r)\cdots A(\tau_1)\,\mathrm{d}\tau \tag{5.28}$$

where the series converges absolutely. Moreover, we have

$$\left\| \sum_{r=1}^{+\infty} \int_{0<\tau_r<\cdots<\tau_1} A(\tau_r)\cdots A(\tau_1)\,\mathrm{d}\tau \right\| \leq \sum_{r=1}^{+\infty} \frac{\|A\|_{L^1}^r}{r!} < e^{\frac{1}{4}} - 1 < 1. \tag{5.29}$$

Thus

$$\log\left(L(t)\right) = \sum_{m=1}^{+\infty} \frac{(-1)^m}{m} \left(\sum_{r=1}^{+\infty} \int_{0<\tau_r<\cdots<\tau_1} A(\tau_r)\cdots A(\tau_1)\,\mathrm{d}\tau \right)^m \tag{5.30}$$

is well defined in $\mathcal{M}_d(\mathbb{K})$ and $L(t) = e^{\log(L(t))}$. By applying Corollary 4.5 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 5.8. *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 $e^X e^Y = e^{Z_\infty}$. Much work has been devoted to proving optimal convergence domains 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 [13], for instance, $\|X\| + \|Y\| < \frac{\ln 2}{2}$. We refer to [12] for a nice survey of the convergence questions regarding the CBHD formula.*

Remark 5.9. *The smallness assumption (on time or on the matrices) is in general necessary, both for the CBHD formula (see [12, Example 2.3] or [71, Section II]) and for the Magnus expansion (see [55], 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. [3, p.1671] or [65, 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 [65, p.335] and [47, 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. [12, 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 5.10. *There exists $\delta > 0$ and $f_0, f_1 \in \mathcal{C}^\omega_\delta$ such that,*

$$\forall M \in \mathbb{N}, \exists C_M, \varepsilon_M > 0, \forall \varepsilon \in [0, \varepsilon_M], \quad \left| e^{\varepsilon f_0} e^{\varepsilon f_1}(0) - e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(0) \right| \leq C_M \varepsilon^{M+1}, \tag{5.31}$$

where $\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)$ is defined in Corollary 4.4, but, simultaneously, for every $\varepsilon > 0$,

$$\lim_{M \to +\infty} |\,\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)(0)| = +\infty \tag{5.32}$$

and

$$\lim_{M \to +\infty} \left| e^{\varepsilon f_0} e^{\varepsilon f_1}(0) - e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(0) \right| = +\infty. \tag{5.33}$$

Proof. Let f_0, f_1 as in Remark 3.16. For these vector fields, estimate (5.31) comes from Corollary 4.4. Due to their structure, the only non vanishing brackets are those containing f_1 at most once. Therefore, formula (2.63) of Corollary 2.36 yields, for $M \geq 1$,

$$\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0) = \varepsilon f_0 + \sum_{k=0}^{M-1} \frac{B_k}{k!} \varepsilon^{k+1} \,\mathrm{ad}^k_{f_0}(f_1). \tag{5.34}$$

Hence, using (3.25),

$$\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)(x) = \varepsilon e_1 + \varepsilon \Theta^\varepsilon_M(x_1) e_2, \tag{5.35}$$

where we introduce, for $q \in \mathbb{R}$,

$$\Theta^\varepsilon_M(q) := \sum_{k=0}^{M-1} B_k \varepsilon^k (1-q)^{-k-1}. \tag{5.36}$$

In particular,

$$|\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)(0)| \geq |\varepsilon \Theta^\varepsilon_M(0)|\,. \tag{5.37}$$

Since the odd Bernoulli numbers except B_1 are zero, when $M = 2M' + 2$ with $M' \geq 1$, $\Theta^\varepsilon_{2M'+2} = \Theta^\varepsilon_{2M'+1}$. Then,

$$\Theta^\varepsilon_{2M'+1}(q) = \frac{1}{1-q} - \frac{\varepsilon}{2(1-q)^2} + \sum_{k=1}^{M'} B_{2k} \varepsilon^{2k} (1-q)^{-2k-1}. \tag{5.38}$$

In particular, using (2.67),

$$\Theta^\varepsilon_{2M'+1}(0) = 1 - \frac{\varepsilon}{2} + \sum_{k=1}^{M'} B_{2k} \varepsilon^{2k} = 1 - \frac{\varepsilon}{2} + \sum_{k=1}^{M'} (-1)^{k+1} \frac{2(2k)!}{(2\pi)^{2k}} \zeta(2k) \varepsilon^{2k}. \tag{5.39}$$

Thus, for every fixed $\varepsilon > 0$, $|\Theta^\varepsilon_M(0)| \to +\infty$ 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 (5.37), this proves (5.32).

For $p \in \mathbb{R}^2$ close enough to the origin, one can also compute the flow $e^{\mathrm{CBHD}_M(\epsilon f_1, \epsilon f_0)}(p)$, which is $y(1)$ where y is the solution to the ODE $y(0) = p$ and

$$\dot{y}_1(s) = \varepsilon \quad \text{and} \quad \dot{y}_2(s) = \varepsilon \Theta^\varepsilon_M\left(y_1(s)\right). \tag{5.40}$$

Solving successively for y_1 then y_2 yields $y_1(s) = p_1 + s\varepsilon$ and

$$y_2(s) = p_2 + \int_{y_1(0)}^{y_1(s)} \Theta^\varepsilon_M(h)\,\mathrm{d}h. \tag{5.41}$$

Thus,

$$e^{\mathrm{CBHD}_M(\varepsilon f_1,\varepsilon f_0)}(p) = (p_1+\varepsilon)e_1 + \left(p_2 + \int_{p_1}^{p_1+\varepsilon} \Theta_M^\varepsilon(h)\,\mathrm{d}h\right) e_2. \tag{5.42}$$

In particular,

$$e^{\mathrm{CBHD}_M(\varepsilon f_1,\varepsilon f_0)}(0) = \varepsilon e_1 + \left(\int_0^{\varepsilon} \Theta_M^\varepsilon(h)\,\mathrm{d}h\right) e_2. \tag{5.43}$$

When $M = 2M'+2$ with $M' \geq 1$, using (5.38), we get

$$e^{\mathrm{CBHD}_M(\varepsilon f_1,\varepsilon f_0)}(0) = \varepsilon e_1 + \left(-\ln(1-\varepsilon) - \frac{\varepsilon}{2}\left(\frac{1}{1-\varepsilon}-1\right) + \sum_{k=1}^{M'} \frac{B_{2k}}{2k}\varepsilon^{2k}\left(\frac{1}{(1-\varepsilon)^{2k}}-1\right)\right) e_2. \tag{5.44}$$

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 (5.33). □

Remark 5.11. *If one sees (x_1, x_2) as (q, p) in an Hamiltonian setting, one checks that the vector fields defined in* (3.24) *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 5.12. *Does 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 5.10?*

The counter-example of Proposition 5.10 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 5.13. *There exist $T, \delta > 0$ and $f \in \mathcal{C}^\omega([0,T] \times B_\delta)$ such that, for every $\varepsilon > 0$ and $t \in (0,T]$,*

$$\lim_{M\to+\infty} |Z_M(t,\varepsilon f)(0)| = +\infty \tag{5.45}$$

and

$$\lim_{M\to+\infty} \left|x(t;\varepsilon f,0) - e^{Z_M(t,\varepsilon f)}(0)\right| = +\infty, \tag{5.46}$$

where x is the solution to $\dot{x}(t) = \varepsilon f(t,x(t))$ with $x(0) = 0$ and $Z_M(t,\epsilon f) = \mathrm{Log}_M\{\varepsilon f\}(t)$.

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 3.16. 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 2.38. Hence, recalling (3.25), we have

$$Z_M(t,\varepsilon f) = \varepsilon t e_1 + \sum_{k=0}^{M-1} \varepsilon^{k+1} \frac{(-1)^{k+1} t^{k+2}}{(k+1)!} B_{k+1} \frac{k!}{(1-x_1)^{k+1}} e_2. \tag{5.47}$$

Proceeding along the same lines as in the proof of Proposition 5.10 allows to conclude that both $Z_M(t,\epsilon f)(0)$ and $e^{Z_M(t,\epsilon)}(0)$ diverge when $M \to +\infty$. □

5.3 Magnus expansion in the interaction picture

5.3.1 Nilpotent systems

For ODEs of the form (4.34), 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

$$x(t; f_0 + f_1, p) = e^{\mathcal{Z}_M(t,f_0,f_1)} e^{tf_0} p, \tag{5.48}$$

where $\mathcal{Z}_M(t, f_0, f_1)$ is defined in Proposition 4.8. 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 analiticity assumption cannot be removed, which is quite surprising but stems from the mixing induced by pushforwards.

We therefore start with the following definition.

Definition 5.14 (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.*

Proposition 5.15. *Let $T, \delta > 0$. Let $M \in \mathbb{N}$. Let $f_0 \in \mathcal{C}^\infty_{4\delta}$ with $T\|f_0\|_{\mathcal{C}^0} \leq \delta$. There exists $\eta > 0$ such that, for every $f_1 : [0,T] \to \mathcal{C}^\infty_{4\delta}$ with $f_1 \in L^1([0,T]; \mathcal{C}^{M+1}_{4\delta})$ and $\|f_1\|_{L^1(\mathcal{C}^M)} \leq \eta$, the following family is well-defined*

$$\mathcal{G} := \{\Phi_0(-t)_* f_1(t);\ \ t \in [0,T]\} \subset \mathcal{C}^\infty_\delta. \tag{5.49}$$

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 (4.34) *satisfies* (5.48)*.*

Proof. Let $t > 0$. As in the proof of Proposition 4.8, we introduce the new unknown $y(s) := \Phi_0(t-s, x(s))$. Then $\dot{y}(s) = g_t(s, y(s))$, where g_t is defined in (4.37). Thanks to Lemma 3.22, $g_t(s) = \Phi_0(t)_* \Phi_0(-s)_* f_1(s)$. Thanks to the assumption and to Lemma 3.23, the family $\{g_t(s);\ s \in [0,t]\}$ is nilpotent of index $M+1$. Thus, by Proposition 5.6, $y(t) = e^{\mathcal{Z}_M(t,f_0,f_1)} y(0)$. Since $x(t) = y(t)$ and $y(0) = \Phi_0(t,p)$, this concludes the proof of (5.48). □

Lemma 5.16. *Let $T, \delta > 0$, $\mathcal{F} \subset \mathcal{C}^\infty_{4\delta}$, $f_0 \in \mathcal{C}^\infty_{4\delta}$ such that $T\|f_0\|_{\mathcal{C}^0} \leq \delta$. The following family is well-defined*

$$\mathcal{G} := \{\Phi_0(-t)_* f;\ \ t \in [0,T], f \in \mathcal{F}\} \subset \mathcal{C}^\infty_\delta. \tag{5.50}$$

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 $\mathcal{F} \cup \{f_0\} \subset \mathcal{C}^{\omega,r}_\delta$. Then $\mathcal{G}$ is nilpotent of index M.

Proof. For $t \in [0,T]$ and $f \in \mathcal{F}$, equation (3.43) of Lemma 3.24 implies that

$$\Phi_0(-t)_* f = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \operatorname{ad}^k_{f_0}(f) \tag{5.51}$$

and that the series converges absolutely in $\mathcal{C}^M_\delta$ (in particular). Hence, if $t_1, \dots, t_M \in [0,T]$ and $f_1, \dots f_M \in \mathcal{F}$, the bracket

$$\begin{aligned}[\Phi_0(-t_M)_* f_M, &[\cdots [\Phi_0(-t_2)_* f_2, \Phi_0(-t_1)_* f_1] \cdots]] \\ &= \sum_{k_1, \dots k_M \in \mathbb{N}} \frac{t_1^{k_1} \cdots t_M^{k_M}}{k_1! \cdots k_M!} [\operatorname{ad}^{k_M}_{f_0}(f_M), [\cdots [\operatorname{ad}^{k_2}_{f_0}(f_2), \operatorname{ad}^{k_1}_{f_0}(f_1)] \cdots]]\end{aligned} \tag{5.52}$$

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 5.17. *Let $T, \delta, r > 0$. Let $f_0 \in \mathcal{C}^{\omega,r}_{4\delta}$ with $T\|f_0\|_{\mathcal{C}^0} \leq \delta$ and $f_1 \in L^1([0,T];\mathcal{C}^{\omega,r}_{4\delta})$. 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 (4.34) satisfies (5.48), where $\mathcal{Z}_M(t,f_0,f_1)$ is defined in Proposition 4.8.*

Proof. This corollary is a direct consequence of Proposition 5.15 and Lemma 5.16. □

The analyticity assumption in Lemma 5.16 is necessary, as illustrated by the following counter-example for smooth functions.

Example 5.18. *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 and $\mathcal{F} := \{f_1, f_2\}$ where*

$$f_0(x) := e_2, \tag{5.53}$$
$$f_1(x) := \chi(x_2)x_1e_3, \tag{5.54}$$
$$f_2(x) := \chi(-x_2)e_1. \tag{5.55}$$

Heuristically, f_1 and f_2 commute because they have disjoint (touching) supports, but the flow of f_0 involved in (5.50) 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 $j \in \mathbb{N}$,

$$\mathrm{ad}^j_{f_0}(f_1)(x) = \chi^{(j)}(x_2)x_1e_3, \tag{5.56}$$
$$\mathrm{ad}^j_{f_0}(f_2)(x) = (-1)^j\chi^{(j)}(-x_2)e_1. \tag{5.57}$$

Thus, for $j,k \in \mathbb{N}$, $[\mathrm{ad}^j_{f_0}(f_1), \mathrm{ad}^k_{f_0}(f_1)]$ (resp. $[\mathrm{ad}^j_{f_0}(f_2), \mathrm{ad}^k_{f_0}(f_2)]$) vanishes because both vector fields are supported by e_3 but independent on x_3 (resp. supported by e_1 but independent on x_1). Moreover,

$$[\mathrm{ad}^j_{f_0}(f_1), \mathrm{ad}^k_{f_0}(f_2)](x) = -(-1)^k\chi^{(k)}(-x_2)\chi^{(j)}(x_2)e_3 = 0, \tag{5.58}$$

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 (5.50) is not nilpotent of index 2. Indeed, for $t \geq 0$ 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,

$$[f_2, (\Phi_0(-t)_*f_1)](x) = \chi(-x_2)\chi(x_2 + t)e_3. \tag{5.59}$$

In particular, for every $\varepsilon > 0$, $[f_2, (\Phi_0(-2\varepsilon)_*f_1)](-\varepsilon e_2) = \chi(\varepsilon)^2e_3 \neq 0$, which prevents the family $\mathcal{G}$ from being nilpotent of index 2 (even locally in time and space).

The analyticity assumption in Corollary 5.17 is also necessary, as illustrated by the following counter-example for smooth functions, inspired by the previous one.

Example 5.19. *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 on time) with*

$$f_1(x) := 2\chi'(x_2)x_1e_3 + \chi'(-x_2)e_1. \tag{5.60}$$

For $j \in \mathbb{N}$, one has

$$\mathrm{ad}^j_{f_0}(f_1)(x) = \partial_2^j f_1(x) = 2\chi^{(j+1)}(x_2)x_1e_3 + (-1)^j\chi^{(j+1)}(-x_2)e_1. \tag{5.61}$$

Thus, for every $j_1, j_2 \in \mathbb{N}$,

$$\begin{aligned}[\mathrm{ad}^{j_1}_{f_0}(f_1), \mathrm{ad}^{j_2}_{f_0}(f_1)](x) = {}& 2(-1)^{j_1}\chi^{(j_1+1)}(-x_2)\chi^{(j_2+1)}(x_2)e_3 \\ & - 2(-1)^{j_2}\chi^{(j_2+1)}(-x_2)\chi^{(j_1+1)}(x_2)e_3 = 0\end{aligned} \tag{5.62}$$

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 (5.48) 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*

$$x_1(t) = p_1 + \chi(-p_2) - \chi(-p_2 - t), \tag{5.63}$$
$$x_2(t) = p_2 + t, \tag{5.64}$$
$$x_3(t) = p_3 + 2\left(\chi(p_2 + t) - \chi(p_2)\right)(p_1 + \chi(-p_2)). \tag{5.65}$$

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*

$$\begin{aligned}\mathcal{Z}_1(t, f_0, f_1)(q) &= \int_0^t f_1(q + (s-t)e_2)\,\mathrm{d}s \\ &= 2q_1(\chi(q_2) - \chi(q_2 - t))e_3 + (\chi(-q_2 + t) - \chi(-q_2))e_1.\end{aligned} \tag{5.66}$$

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

$$y_1(s) = p_1 + s(\chi(-p_2) - \chi(-p_2 - t)), \tag{5.67}$$
$$y_2(s) = p_2 + t, \tag{5.68}$$
$$y_3(s) = p_3 + (\chi(p_2 + t) - \chi(p_2))\left(2p_1 s + s^2(\chi(-p_2) - \chi(-p_2 - t))\right). \tag{5.69}$$

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 $\varepsilon > 0$,

$$\left|x(2\varepsilon, f_0 + f_1, -\varepsilon e_2) - e^{\mathcal{Z}_1(2\varepsilon,f_0,f_1)}e^{2\varepsilon f_0}(-\varepsilon e_2)\right| = \chi^2(\varepsilon) > 0. \tag{5.70}$$

5.3.2 Convergence for linear systems

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

$$\dot{x}(t) = (H_0 + H_1(t))\,x(t) \quad \text{and} \quad x(0) = p, \tag{5.71}$$

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 5.20. *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 $\|H_1\|_{L^1_T} < \frac{e^{-2T\|H_0\|}}{8}$. Then, for each $t \in [0,T]$ and $p \in \mathbb{K}^d$ the solution to (5.71) satisfies $x(t) = e^{-Z_\infty(t)}e^{tH_0}p$ where $Z_\infty(t)$ is defined by (5.24) with*

$$A_t(\tau) = e^{(t-\tau)H_0}H_1(\tau)e^{(\tau-t)H_0} = \sum_{k=0}^{+\infty} \frac{(t-\tau)^k}{k!}\,\mathrm{ad}^k_{H_0}(H_1). \tag{5.72}$$

Proof. The function $y : \tau \in [0,t] \mapsto e^{(t-\tau)H_0}x(\tau)$ solves $y'(\tau) = A(\tau)$, $y(0) = e^{tH_0}p$. Thus, by Proposition 5.7, $y(t) = e^{-Z_\infty(t)}e^{tH_0}p$, which gives the conclusion because $y(t) = x(t)$. □

Remark 5.21. *The Magnus expansion in the usual setting (Proposition 5.7), when applied directly to $A(t) = H_0 + H_1(t)$ requires a smallness assumption on $T\|H_0\|$ (through the condition $\|A\|_{L^1_T} < \frac{1}{8}$), even for small perturbations H_1. On the contrary, the Magnus expansion in the interaction picture (Proposition 5.20) holds even when $T\|H_0\|$ is large, provided that the perturbation H_1 is small enough.*

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 $a \in [\![1, d]\!]$, $f_0(x)$ only depends on $x_1, \dots, x_a$ and is supported by $e_1, \dots, e_a$ and $f_1(t, x)$ only depends on $x_{a+1}, \dots x_d$ and is supported along $e_{a+1}, \dots e_d$, then the vector field $g_t(\tau) = \Phi_0(t-\tau)_* f_1(\tau)$ defined in (4.37) 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.

5.4 Sussmann's infinite product expansion

5.4.1 Equality for nilpotent systems

In this section, we study affine systems of the form (4.51).

Proposition 5.22. *Let $\mathcal{B}$ be a generalized 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 $T, \delta > 0$, $f_i \in \mathcal{C}^\infty_{3\delta}$ 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 (4.52), for each $t \in [0,T]$ and $p \in B_\delta$,*

$$x(t; f, u, p) = \overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t;u) f_b} p. \tag{5.73}$$

Proof. The proof strategy is the same as for Proposition 4.13. We apply the second statement of Lemma 3.24 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 Linear dynamics in Banach algebras

Convergence for small matrices. Let $T > 0$. In this paragraph, we study affine linear systems of the form

$$\dot{x}(t) = \left(\sum_{i\in I} u_i(t) A_i\right) x(t) \quad \text{and} \quad x(0) = p \tag{5.74}$$

where $A_i \in \mathcal{M}_d(\mathbb{K})$ and $u_i \in L^1((0,T);\mathbb{K})$. When well-defined, its solution is denoted $x(t; A, u, p)$. The main goal of this section is to prove Proposition 5.25 which asserts that Sussmann's infinite product expansion for system (5.74) converges locally (i.e. for small matrices or small controls).

Before proving this result, we need a definition for an ordered infinite product (given in Definition 5.23 below) and a sufficient condition for its convergence (given in Lemma 5.24 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 generalized Hall bases, 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. This problem already appears for a product which would be indexed by $\mathbb{N}^2$ with the lexicographic order

$$(0,0) < (0,1) < (0,2) < \cdots < (1,0) < (1,1) < (1,2) < \cdots < (2,0) < \cdots \tag{5.75}$$

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 5.23. *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*

$$\left\| M - \overleftarrow{\prod_{j\in J_1}} e^{A_j} \right\| \leq \varepsilon. \tag{5.76}$$

When such an M exists, it is unique and we write

$$M = \overleftarrow{\prod_{j\in J}} e^{A_j}. \tag{5.77}$$

Lemma 5.24. *Let J be a totally ordered set and $(A_j)_{j\in J}$ matrices of $\mathcal{M}_d(\mathbb{K})$ such that*

$$\sum_{j\in J} \|A_j\| < +\infty. \tag{5.78}$$

Then the ordered product of the e^{A_j} over J converges in the sense of Definition 5.23.

Proof. Let α be the left-hand side of (5.78).

Step 1: Basic claims. We start with straightforward claims. First, for every $j \in J$, one has

$$\|e^{A_j} - \mathrm{Id}\| \leq e^{\|A_j\|} - 1 \leq \|A_j\| e^{\|A_j\|} \leq \|A_j\| e^{\alpha}. \tag{5.79}$$

Second, for every finite part $J' \subset J$, one has

$$\left\| \overleftarrow{\prod_{j\in J'}} e^{A_j} \right\| \leq \prod_{j\in J'} e^{\|A_j\|} \leq e^{\alpha}. \tag{5.80}$$

Third, for every finite parts $J_0 \subset J_1 \subset J$, one has

$$\left\| \overleftarrow{\prod_{j\in J_1}} e^{A_j} - \overleftarrow{\prod_{j\in J_0}} e^{A_j} \right\| \leq e^{3\alpha} \sum_{j\in J_1\setminus J_0} \|A_j\|. \tag{5.81}$$

Indeed, writing $J_1 \setminus J_0 = \{j_1 > \cdots > j_n\}$, we have the following telescopic decomposition

$$\overleftarrow{\prod_{j\in J_1}} e^{A_j} - \overleftarrow{\prod_{j\in J_0}} e^{A_j} = \sum_{k=1}^{n} \overleftarrow{\prod_{\substack{j\in J_0 \\ j>j_k}}} e^{A_j} \left(e^{A_{j_k}} - \mathrm{Id}\right) \overleftarrow{\prod_{\substack{j\in J_1 \\ j<j_k}}} e^{A_j}, \tag{5.82}$$

which, together with the two first claims, proves estimate (5.81).

Step 2: Construction of a limit. We construct a possible limit. For each $n \geq 2$, let

$$J_n := \left\{ j \in J,\ \|A_j\| > \frac{1}{n} \right\}. \tag{5.83}$$

Thanks to assumption (5.78), the sets J_n are finite and, moreover,

$$\varepsilon_n := \sum_{j\in J\setminus J_n} \|A_j\| \to 0. \tag{5.84}$$

Now, for each $n \geq 2$, we define the matrix

$$M_n := \overleftarrow{\prod_{j \in J_n}} e^{A_j}. \tag{5.85}$$

This defines a Cauchy sequence in the complete space $\mathcal{M}_d(\mathbb{K})$. Indeed, for every $n < p$, thanks to estimate (5.81), one has

$$\|M_n - M_p\| \leq e^{3\alpha}\varepsilon_n. \tag{5.86}$$

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 $n \geq 2$

$$\|M_n - M\| \leq e^{3\alpha}\varepsilon_n. \tag{5.87}$$

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 5.23. 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 (5.76) holds thanks to (5.87) and (5.81). □

Proposition 5.25. *Let $\mathcal{B}$ be a generalized 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 algebra such that $\Lambda(X_i) = A_i$ for $i \in I$ (see Lemma 2.7). Let $T > 0$ and $u_i \in L^1((0,T);\mathbb{K})$ for $i \in I$. Assume that*

$$\|u\|_{L^1_T}\|A\| \leq \eta. \tag{5.88}$$

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 $p \in \mathbb{K}^d$,

$$x(t;A,u,p) = \overrightarrow{\prod_{b\in\mathcal{B}}} e^{\xi_b(t;u)A_b} p. \tag{5.89}$$

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 $b \in \mathcal{B}$, $\|A_b\| \leq (2\|A\|)^{|b|}$. Thus, recalling (2.102),

$$\|\xi_b(t)A_b\| \leq \left(2\|A\|\|u\|_{L^1(0,t)}\right)^{|b|}. \tag{5.90}$$

Taking into account that $|\mathcal{B}_\ell| \leq |I|^\ell$, we obtain, using (5.88),

$$\sum_{b\in\mathcal{B}} \|\xi_b(t)A_b\| \leq \sum_{\ell=1}^{+\infty} \left(2|I|\|A\|\|u\|_{L^1(0,t)}\right)^\ell \leq 1 \tag{5.91}$$

and Lemma 5.24 gives the conclusion.

Step 2: Estimates along a Lazard elimination in $\mathcal{B}_{[\![1,M]\!]}$. Let $M \in \mathbb{N}^*$. We adopt the notations $b_1, \ldots, b_{k+1}$ and $Y_0, \ldots, Y_{k+1}$ of Definition 2.49 and we define $x_0(t) := x(t)$ and, for $j \in [\![1, k+1]\!]$

$$x_j(t) := e^{-\xi_{b_j}(t;u)A_{b_j}} \cdots e^{-\xi_{b_1}(t;u)A_{b_1}} x(t). \tag{5.92}$$

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

$$(\mathcal{H}_j) : \begin{cases} \dot{x}_j(t) = \left(\sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_j} \dot{\xi}_b(t)A_b + \varepsilon_j(t)\right) x_j(t), \\ x_j(0) = p, \end{cases} \tag{5.93}$$

where $\varepsilon_0 = 0$ and

$$\|\varepsilon_j(t)\| \leq \Big(M|I|\|A\||u(t)|(4|I|\|A\|\|u\|_{L^1_t})^M + \|\varepsilon_{j-1}(t)\|\Big)\, e^{2\|\xi_{b_j}(t)A_{b_j}\|}. \tag{5.94}$$

First, $(\mathcal{H}_0)$ holds with $\epsilon_0 = 0$ 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

$$x_j(t) = e^{-\xi_{b_j}(t)A_{b_j}}\, x_{j-1}(t) \tag{5.95}$$

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

$$\begin{aligned}\dot x_j(t) &= -\dot\xi_{b_j}(t)A_{b_j}x_j(t) + e^{-\xi_{b_j}(t)A_{b_j}} \left(\sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_{j-1}} \dot\xi_b(t)A_b + \varepsilon_{j-1}(t)\right) e^{\xi_{b_j}(t)A_{b_j}}\, x_j(t)\\ &= \left(\sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_{j-1}\setminus\{b_j\}} \dot\xi_b(t)e^{-\xi_{b_j}(t)A_{b_j}}A_b e^{\xi_{b_j}(t)A_{b_j}} + \widetilde\varepsilon_{j-1}(t)\right) x_j(t)\end{aligned} \tag{5.96}$$

where $\widetilde\varepsilon_{j-1}(t) := e^{-\xi_{b_j}(t)A_{b_j}}\epsilon_{j-1}(t)e^{\xi_{b_j}(t)A_{b_j}}$ satisfies,

$$\|\widetilde\varepsilon_{j-1}(t)\| \leq \|\varepsilon_{j-1}(t)\|e^{2\|\xi_{b_j}(t)A_{b_j}\|}. \tag{5.97}$$

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 (4.65) holds and

$$\bar\varepsilon^j_b(t) := \dot\xi_b(t)e^{-\xi_{b_j}(t)A_{b_j}}A_b e^{\xi_{b_j}(t)A_{b_j}} - \sum_{m=0}^{h(b)-1} \dot\xi_b(t)\frac{\xi^m_{b_j}(t)}{m!}A_{\mathrm{ad}^m_{b_j}(b)}. \tag{5.98}$$

Then, by definition of Y_j, $(\mathcal{H}_j)$ holds with ε_j defined by (4.69). Using the fourth statement of Lemma 3.24, (2.101), (5.90), we obtain

$$\begin{aligned}\|\bar\varepsilon^j_b(t)\| &\leq |\dot\xi_b(t)|\frac{(2\|\xi_{b_j}(t)A_{b_j}\|)^{h(b)}}{h(b)!}\|A_b\|e^{2\|\xi_{b_j}(t)A_{b_j}\|}\\ &\leq |b||u(t)|\|u\|_{L^1_t}^{|b|-1}\cdot\frac{(2\|A\|\|u\|_{L^1_t})^{h(b)|b_j|}2^{h(b)}}{h(b)!}\cdot(2\|A\|)^{|b|}\\ &\leq M|u(t)|(4\|A\|\|u\|_{L^1_t})^M\|A\|,\end{aligned} \tag{5.99}$$

taking into account $M+1 \leq |b| + h(b)|b_j| \leq 2M$ and $\|A\|\|u\|_{L^1_t} \leq 1$.

We deduce from (4.69), (5.97), (5.99) and the relation $|\mathcal{B}_{[\![1,M]\!]}| \leq |I|^{M+1}$ that (5.94) 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 (5.94), (5.91) and the relation $k+1 = |\mathcal{B}_{[\![1,M]\!]}| \leq |I|^{M+1}$ that

$$\|\varepsilon_{k+1}(t)\| \leq eM\|A\||I|^2|u(t)|(4|I|^2\|A\|\|u\|_{L^1_t})^M. \tag{5.100}$$

Hence, using (5.88),

$$\|\varepsilon_{k+1}\|_{L^1_t} \leq \frac{e}{4}(4|I|^2\|A\|\|u\|_{L^1_t})^{M+1} \leq 2^{-M}. \tag{5.101}$$

We deduce from $(\mathcal{H}_{k+1})$, (2.94) and Grönwall's lemma that

$$\left|\overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{-\xi_b(t;u)A_b}x(t) - p\right| = |x_{k+1}(t) - p| \leq \int_0^t |\varepsilon_{k+1}(\tau)x_{k+1}(\tau)|d\tau \leq 2^{-M}e|p| \tag{5.102}$$

Multiplying both sides by the finite product $\overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t;u)A_b}$ gives

$$\left|x(t) - \overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t;u)A_b}p\right| \leq e^2 2^{-M}|p| \tag{5.103}$$

Passing to the limit $[M \to \infty]$ in the previous estimate gives (5.89). □

Divergence for matrices in large time. The goal of this paragraph is to illustrate that the smallness assumption (5.88) in Proposition 5.25 is necessary because the equality does not hold globally.

Proposition 5.26. *Let $\overline{u} : t \in \mathbb{R}_+ \mapsto (1,1) \in \mathbb{R}^2$.*

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*

$$\exists \gamma > 0, \forall k \in \mathbb{N}, t \geq 0, \quad \xi_{b_k}(t;\overline{u}) \geq \left(\frac{t}{\gamma}\right)^{|b_k|} \tag{5.104}$$

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 [68, 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

$$b_0^1 = X_1, \qquad b_0^2 = X_2, \qquad b_{k+1}^1 := [b_k^2, [b_k^1, b_k^2]], \qquad b_{k+1}^2 := [b_k^1, [b_k^1, b_k^2]]. \tag{5.105}$$

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 2.51), provided that one chooses an arbitrary length-compatible order on all other brackets.

To simplify notations in this proof, 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 $b \in \mathcal{B}$, $\xi_b(t) = t^{|b|-1}/\alpha_b$, 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 $k \in \mathbb{N}$,

$$\alpha_{b_{k+1}^1} = \alpha_{b_k^1}\alpha_{b_k^2}^2 |b_k^2||b_k^1| = \alpha_{b_k^1}\alpha_{b_k^2}^2 3^{2k}, \qquad \alpha_{b_{k+1}^2} = 2\alpha_{b_k^1}^2 \alpha_{b_k^2} |b_k^1|^2 = 2\alpha_{b_k^1}^2 \alpha_{b_k^2} 3^{2k}. \tag{5.106}$$

Let $\beta_k = \max\{\alpha_{b_k^1}, \alpha_{b_k^2}\}$. Then, $\beta_0 = 1$ and, by the previous relations,

$$\beta_{k+1} \leq 3^{2k+1}\beta_k^3. \tag{5.107}$$

Thus $\theta_k := 3^{-k}\ln(\beta_k)$ satisfies $\theta_0 = 0$ and

$$\theta_{k+1} \leq \theta_k + (2k+1)3^{-(k+1)}\ln(3), \tag{5.108}$$

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

$$\left|\xi_{b_k^j}(t)\right| \geq \frac{1}{|b_k^j|}\left(\frac{t}{\gamma'}\right)^{|b_k^j|}. \tag{5.109}$$

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 (5.104) 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

$$[F_2, [F_1, F_2]] = F_1, \quad [F_1, [F_1, F_2]] = -F_2. \tag{5.110}$$

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 $t \in \mathbb{R}$

$$e^{\xi_{b_k}(t)A_{b_k}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cosh(\xi_{b_k}(t)) & i(-1)^k \sinh(\xi_{b_k}(t)) \\ 0 & i(-1)^{k+1}\sinh(\xi_{b_k}(t)) & \cosh(\xi_{b_k}(t)) \end{pmatrix} \tag{5.111}$$

By (5.104), this sequence of matrices diverges for every $t > \gamma$. □

5.4.3 Investigation for analytic vector fields

In this paragraph, we study affine systems (4.51). 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 5.27) and a sufficient condition for the convergence (Lemma 5.28), 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 5.27. *Let J be a totally ordered set, $\delta > 0$ and $(f_j)_{j\in J}$ a family of $\mathcal{C}^1_{2\delta}$. We say that the ordered product of the e^{f_j} over J converges uniformly on B_δ if there exists $g \in \mathcal{C}^0_\delta$ 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*

$$\left\| g(p) - \overleftarrow{\prod_{j\in J_1}} e^{f_j} p \right\| \leq \varepsilon. \tag{5.112}$$

When such a g exists, it is unique and we write

$$g = \overleftarrow{\prod_{j\in J}} e^{f_j}. \tag{5.113}$$

Lemma 5.28. *Let J be a totally ordered set, $\delta > 0$ and $(f_j)_{j\in J}$ a family of $\mathcal{C}^1_{2\delta}$ such that*

$$\sum_{j\in J} \|f_j\|_{\mathcal{C}^0} < \delta \quad \text{and} \quad \alpha := \sum_{j\in J} \|f_j\|_{\mathcal{C}^1} < \infty. \tag{5.114}$$

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 5.24.

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

$$\overleftarrow{\prod_{j\in J'}} e^{f_j} p \in B_{2\delta} \qquad \text{and} \qquad \left\| \partial_p \left[\overleftarrow{\prod_{j\in J'}} e^{f_j} p \right] \right\| \leq \prod_{j\in J'} e^{\|f_j'\|_{\mathcal{C}^0}} \leq e^\alpha \tag{5.115}$$

because of Lemma 3.19 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

$$\left\| \overleftarrow{\prod_{j\in J_1}} e^{f_j} p - \overleftarrow{\prod_{j\in J_0}} e^{f_j} p \right\| \leq e^\alpha \sum_{j\in J_1\setminus J_0} \|f_j\|_{\mathcal{C}^0}. \tag{5.116}$$

Indeed, writing $J_1 \setminus J_0 = \{j_1 > \cdots > j_n\}$, we have the following telescopic decomposition

$$\overleftarrow{\prod_{j\in J_1}} e^{f_j} p - \overleftarrow{\prod_{j\in J_0}} e^{f_j} p = \sum_{k=1}^{n} \left\{ \left(\overleftarrow{\prod_{\substack{j\in J_0 \\ j>j_k}}} e^{f_j} \right) e^{f_{j_k}} \left(\overleftarrow{\prod_{\substack{j\in J_1 \\ j<j_k}}} e^{f_j} \right) p - \left(\overleftarrow{\prod_{\substack{j\in J_0 \\ j>j_k}}} e^{f_j} \right) \left(\overleftarrow{\prod_{\substack{j\in J_1 \\ j<j_k}}} e^{f_j} \right) p \right\}. \tag{5.117}$$

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 (5.115) and (3.27), the term with index k in the previous sum is bounded by

$$\left| \left(\overleftarrow{\prod_{\substack{j\in J_0\\ j>j_k}}} e^{f_j} \right) e^{f_{j_k}} x_k - \left(\overleftarrow{\prod_{\substack{j\in J_0\\ j>j_k}}} e^{f_j} \right) x_k \right| \leq e^{\alpha} \left| e^{f_{j_k}} x_k - x_k \right| \leq e^{\alpha} \|f_{j_k}\|_{\mathcal{C}^0}. \tag{5.118}$$

which, together with (5.117) proves (5.116).

Step 2: Construction of a limit. We construct a possible limit. For each $n \geq 2$, let

$$J_n := \left\{ j \in J,\ \|f_j\|_{\mathcal{C}^1} > \frac{1}{n} \right\}. \tag{5.119}$$

Thanks to assumption (5.114), the sets J_n are finite and, moreover,

$$\varepsilon_n := \sum_{j\in J\setminus J_n} \|f_j\|_{\mathcal{C}^1} \to 0. \tag{5.120}$$

Now, for each $n \geq 2$, we define $g_n \in \mathcal{C}^0_\delta$ by

$$g_n(p) := \overleftarrow{\prod_{j\in J_n}} e^{f_j} p. \tag{5.121}$$

This defines a Cauchy sequence in the complete space $\mathcal{C}^0_\delta$. Indeed, for every $n < n'$ and $p \in B_\delta$, thanks to estimate (5.116), one has

$$\|g_n(p) - g_{n'}(p)\| \leq e^{\alpha} \varepsilon_n. \tag{5.122}$$

Hence, there exists $g \in \mathcal{C}^0_\delta$ towards which the sequence $(g_n)_{n\geq 2}$ uniformly converges on B_δ. By (5.115), 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 $p \in B_\delta$

$$\|g_n(p) - g(p)\| \leq e^{\alpha} \varepsilon_n. \tag{5.123}$$

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 5.27. 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 (5.112) holds thanks to (5.123) and (5.116). □

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 generalized 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 5.26, 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 $f_i \in \mathcal{C}^{\omega,r}_\delta$ for $i \in I$. By (3.23) with $r_1 \leftarrow r$ and $r_2 \leftarrow r/e$ for every $b \in \mathcal{B}$,

$$\|f_b\|_{\mathcal{C}^1} \leq \left(1 + \frac{e}{r}\right) (|b|-1)! \left(\frac{9}{r}\right)^{|b|-1} F^{|b|}, \tag{5.124}$$

where $F := \max_{i\in I} |||f_i|||_r$. However, by Remark 3.16, 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

$$|\xi_b(t;u)|||f_b||_{\mathcal{C}^1} \leq (Ct)^{|b|}|b|!. \tag{5.125}$$

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 5.28 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 5.29. *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 5.28 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

$$\dot{x}(t) = f_0(x(t)) + \sum_{i=1}^{q} u_i(t) f_i(x(t)) \quad \text{and} \quad x(0) = p, \tag{6.1}$$

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 iterated Duhamel 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 Iterated Duhamel or Chen-Fliess expansion

The convergence of the Chen-Fliess series, for control affine systems (6.1) with analytic vector fields, under a smallness assumption on t and a uniform bound on u, is a classical result, see for instance [28, Proposition 3.37] or [67, Proposition 4.3]. In this section we prove the convergence of the Chen-Fliess expansion, (Proposition 6.1) 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 6.2), 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 $f_\sigma = (f_{\sigma_1} \cdot \nabla) \cdots (f_{\sigma_\ell} \cdot \nabla)$. 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 (2.11), with $u_0 = 1$.

Proposition 6.1. *Let $\delta, r > 0$ and $f_0, f_1, \dots, f_q \in \mathcal{C}^{\omega,r}_{2\delta}$. 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*

$$\varphi(x(t; f, u, p)) = \sum_{\sigma \in I^*} \left(\int_0^t u_\sigma \right) (f_\sigma \varphi)(p) \tag{6.2}$$

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*

$$\left|\varphi(x(t;f,u,p))-\sum_{n(\sigma)\leq M}\left(\int_0^t u_\sigma\right)(f_\sigma\varphi)(p)\right|\leq\left(C\|u\|_{L^1}\right)^{M+1}, \tag{6.3}$$

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 $u\in L^1((0,t);\mathbb{K}^q)$,

$$\left|\left(\int_0^t u_\sigma\right)\right|\leq\frac{\|u\|_{L^1(0,t)}^{n(\sigma)}}{n(\sigma)!}\frac{t^{n_0(\sigma)}}{n_0(\sigma)!}. \tag{6.4}$$

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 (6.4) and (3.18), we get

$$\left|\left(\int_0^t u_\sigma\right)(f_\sigma\varphi)(p)\right|\leq\|u\|_{L^1(0,t)}^{n(\sigma)}t^{n_0(\sigma)}\left(\frac{10}{r}\|f\|\right)^\ell|||\varphi|||_r \tag{6.5}$$

which proves the absolute convergence of the sum in (6.2), uniformly with respect to (t,u,p)

The proof of the equality in (6.2) consists in applying (5.9) to $f(t,x)=f_0(x)+\sum_{i=1}^q u_i(t)f_i(x)$. In particular the sum involved in (6.3) is the Taylor expansion of order M of $u\mapsto\varphi(x(t;f,u,p))$ at $u=0$. By adapting Lemma 3.10 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 (6.3). □

The last statement of this section focuses on nonlinear control systems with scalar input

$$\dot{x}=f(x,u) \tag{6.6}$$

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

$$\int_0^t u^k:=\int_{\mathcal{T}_{(n)}(t)}u(\tau_n)^{k_n}\cdots u(\tau_1)^{k_1}\,\mathrm{d}\tau \tag{6.7}$$

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 6.2. *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* $p\in B_\delta$

$$\varphi\left(x(t;f,u,p)\right)=\sum_{\substack{n\in\mathbb{N}\\k\in\mathbb{N}^n}}\left(\int_0^t u^k\right)\Big((f_{k_n}\cdot\nabla)\cdots(f_{k_1}\cdot\nabla)\Big)\Big(\varphi\Big)(p) \tag{6.8}$$

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* $p\in B_\delta$

$$\left|\varphi\left(x(t;f,u,p)\right)-\sum_{\substack{n\in\mathbb{N}\\k\in\mathbb{N}^n,|k|\leq M}}\left(\int_0^t u^k\right)\Big((f_{k_n}\cdot\nabla)\cdots(f_{k_1}\cdot\nabla)\Big)\Big(\varphi\Big)(p)\right|\leq\left(C\|u\|_{L^\infty}\right)^{M+1} \tag{6.9}$$

where the sum is taken over $n \in \mathbb{N}$ *and* $k = (k_1, \dots, k_n) \in \mathbb{N}^n$ *such that* $k_1 + \cdots + k_n \leq M$.

Proof. We define $r' = r/e$,

$$T^* := \min\left\{ \frac{r'}{10\,|||f|||_r}, \frac{\delta}{\|f\|_{\mathcal{C}^0}} \right\}, \qquad \eta := \min\left\{ \delta_u, \frac{r}{10} \right\}. \tag{6.10}$$

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 (6.8). Using the iterated version of (3.8) and (3.1), we get, for every $k \in \mathbb{N}$,

$$|||f_k|||_{r'} \leq \frac{1}{k!}\left(\frac{k}{r-r'}\right)^k |||f|||_r \leq \left(\frac{e}{r-r'}\right)^k |||f|||_r \leq \left(\frac{5}{r}\right)^k |||f|||_r . \tag{6.11}$$

For every $n \in \mathbb{N}^*$ and $k_1, \dots, k_n \in \mathbb{N}$, we have, using (3.18) and (6.11)

$$\begin{aligned} \Big|\Big((f_{k_n} \cdot \nabla) \cdots (f_{k_1} \cdot \nabla)\Big)\Big(\varphi\Big)(p)\Big| &\leq n! \left(\frac{5}{r'}\right)^n |||f_{k_n}|||_{r'} \cdots |||f_{k_1}|||_{r'}\, |||\varphi|||_{r'} \\ &\leq n! \left(\frac{5}{r'}\right)^n \left(\frac{5}{r}\right)^{k_1+\cdots+k_n} |||f|||_r^n\, |||\varphi|||_{r'} \end{aligned} \tag{6.12}$$

and

$$\left|\int_0^t u^k\right| = \left|\int_{0<\tau_1<\cdots<\tau_n<t} u(\tau_n)^{k_n} \cdots u(\tau_1)^{k_1}\, \mathrm{d}\tau\right| \leq \frac{t^n}{n!}\|u\|_{L^\infty}^{k_1+\cdots+k_n}. \tag{6.13}$$

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 (6.8) and error formula (6.9). We have $f(\cdot, u) = \sum_{j=0}^{+\infty} u^j f_j$ with convergence in $\mathcal{C}^{\omega,r'}_{2\delta}$ uniformly with respect to $u \in B_{\mathbb{K}^q}(0, \eta)$. Thus, the equality (6.8) is a consequence of Fubini theorem and (5.9) applied to $(t, x) \mapsto f(x, u(t))$. In particular the finite sum involved in (6.9) is the Talyor expansion of order M of $u \mapsto \varphi(x(t; f, u, p))$ at $u = 0$. By Lemma 3.10 $u \mapsto \varphi(x(t; f, u, p))$ 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 (6.9). □

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 6.3. *Let* $d := 2$. *There exists* $T, \delta > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,\delta}_\delta$ *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*

$$F_n(t) := \sum_{k=0}^{n} \zeta_{\mathrm{ad}^k_{X_0}(X_1)}(t, u)\, \mathrm{ad}^k_{f_0}(f_1), \tag{6.14}$$

then, for each $\delta^* \in (0, \delta)$ *and* $t \in (0, T)$, *the sequence* $F_n(t) \in \mathcal{C}^\infty_\delta$ *does not converge in* $\mathcal{C}^0_{\delta^*}$.

Proof. We define the following vector fields for $x \in \mathbb{R}^2$ with $|x| < 1$,

$$f_0(x) := x_2 e_1 \quad \text{and} \quad f_1(x) := \frac{1}{1 - x_1} e_2. \tag{6.15}$$

Then,

$$\mathrm{ad}^k_{f_0}(f_1)(x) = x_2^k \frac{\partial^k}{\partial x_1^k}\left(\frac{1}{1-x_1}\right) e_2 = \frac{k! x_2^k}{(1-x_1)^{k+1}} e_2. \tag{6.16}$$

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 (2.68) from Example 2.38 for the coordinates of the first kind along the brackets $\mathrm{ad}^k_{X_0}(X_1)$ for this particular control, we obtain, for $t \in (0,T)$,

$$F_n(t)(x) = \sum_{k=0}^{n} (-1)^{k+1} t^{k+2} \frac{B_{k+1}}{k+1} \frac{x_2^k}{(1-x_1)^{k+1}}. \tag{6.17}$$

Thus, for each $t, \delta^* > 0$, the sequence of vector fields $F_n(t)$ does not converge in $\mathcal{C}^0_{\delta^*}$, since for every $x_2 \neq 0$, the general term of the series does not tend to zero because of the asymptotic (2.67) for Bernoulli numbers. □

6.3 Magnus expansion in the interaction picture

Proposition 6.4. *For every $M \in \mathbb{N}$, there exists $\widetilde{\Theta}_M \in \mathcal{C}^0(\mathbb{R}_+^* \times \mathbb{R}_+^3; \mathbb{R}_+)$ and $\Phi_M \in \mathcal{C}^0(\mathbb{R}_+^{2+q}; \mathbb{R}_+)$ such that, for every $\delta > 0$, $T > 0$, $f_0 \in \mathcal{C}^{M^2+1}_{5\delta}$ with $T\|f_0\|_{\mathcal{C}^0} \leq \delta$, $f_1, \ldots, f_q \in \mathcal{C}^{M^2}_{5\delta}$, $u_1, \ldots, u_q \in L^1((0,T);\mathbb{K})$ with*

$$\|u\|_{L^1} \leq \widetilde{\Theta}_M(\delta, T, \|f_0\|_{\mathcal{C}^{M^2+1}}, \sum_{j=1}^{q} \|f_j\|_{\mathcal{C}^{M^2}}) \tag{6.18}$$

$p \in B_\delta$ and $t \in [0,T]$ then

$$\left| x(t; f, u, p) - e^{\mathcal{Z}_M(t,f,u)} e^{tf_0} p \right| \leq \|u\|^{M+1}_{L^1(0,t)} \Phi_M(T, \|f_0\|_{\mathcal{C}^{M^2+1}}, \|f_1\|_{\mathcal{C}^{M^2}}, \ldots, \|f_q\|_{\mathcal{C}^{M^2}}) \tag{6.19}$$

where $\mathcal{Z}_M(t, f_0, f_1) = \mathrm{Log}_M\{g_t\}(t)$ in the sense of Definition 2.19, $g_t : [0,t] \times B_{4\delta} \mapsto \mathbb{K}^d$ is defined by

$$g_t(\tau, y) = \sum_{i=1}^{q} u_i(t)(\Phi_0(t-\tau)_* f_i(\tau))(y) \tag{6.20}$$

and $\Phi_0 : [0,T] \times B_{4\delta} \to B_{5\delta}$ is the flow associated with f_0 i.e. $\Phi_0(t;p) = e^{tf_0}(p)$.

Proof. We define

$$\widetilde{\Theta}_M(\delta, T, \|f_0\|_{\mathcal{C}^{M^2+1}}, \sum_{i=1}^{q} \|f_i\|_{\mathcal{C}^{M^2}}) = \Theta_M(T, \|f_0\|_{\mathcal{C}^{M^2+1}}) \min\{1; \delta\} / \sum_{i=1}^{q} \|f_i\|_{\mathcal{C}^{M^2}} \tag{6.21}$$

with Θ_M as in Proposition 4.8. Then the assumptions of Proposition 6.4 imply (4.35) for $f_1(t,x) \leftarrow \sum_{i=1}^q u_i(t) f_i(x)$. Then (4.36) gives (6.19) because for every $i \in [\![1,q]\!]$ and $\tau \in [0,t]$, $\|(\Phi_0(t-\tau)_* f_i(\tau)\|_{\mathcal{C}^{M^2}}$ is bounded by a continuous function of T, $\|f_0\|_{\mathcal{C}^{M^2+1}}$ and $\|f_i\|_{\mathcal{C}^{M^2}}$. □

6.4 Sussmann's infinite product expansion

The goal of this section is to prove Proposition 6.7 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 (6.1).

We start with an elementary remark (Lemma 6.5) on the structure of brackets of a generalized Hall basis which allows to prove nice asymmetric estimates on the associated coordinates of the second kind (see Lemma 6.6).

Lemma 6.5. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$ and $\mathcal{B}$ a generalized Hall basis of $\mathcal{L}(X)$. For each $b \in \mathcal{B}$, there exist $m, \overline{m} \in \mathbb{N}$ such that*

$$b = \mathrm{ad}_{X_0}^m \, \overline{\mathrm{ad}}_{X_0}^{\overline{m}}(b^*), \tag{6.22}$$

where $\overline{\mathrm{ad}}_{X_0}^{\overline{m}}$ denotes the iterated right bracketing $\overline{m}$ times by X_0 and $b^* \in \mathcal{B}$ 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 2.50, for each $b \in \mathcal{B} \setminus X$, $\lambda(b) < b$. Let $b \in \mathcal{B}$. We disjunct cases.

- If $b \in X$ or ($\lambda(b) \neq X_0$ and $\mu(b) \neq X_0$), then (6.22) 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 $\mu(\tilde{b}) \neq X_0$, (6.22) holds with $\overline{m} = 0$ and $b^* = \tilde{b}$.
  - Otherwise, there exists a unique $\overline{m} \in \mathbb{N}^*$ and $b^* \in \mathcal{B}$ such that $\tilde{b} = \overline{\mathrm{ad}}_{X_0}^{\overline{m}}(b^*)$ where $b^* \in X$ or $\mu(b^*) \neq X_0$.
    - If $b^* \in X$, (6.22) holds.
    - Else $\mu(b^*) \neq X_0$. one has $\lambda(b^*) < b^*$ as recalled. Moreover, since $\overline{m} \geq 1$, $[b^*, X_0] \in \mathcal{B}$ so $b^* < X_0$ (by the second point of Definition 2.50). Hence $\lambda(b^*) < X_0$. So we also have $\lambda(b^*) \neq X_0$ and (6.22) holds.
- If $\mu(b) = X_0$, there exists a unique $\overline{m} \in \mathbb{N}^*$ and $\tilde{b} \in \mathcal{B}$ such that $b = \overline{\mathrm{ad}}_{X_0}^{\overline{m}}(\tilde{b})$ where $\tilde{b} \in X$ or $\mu(\tilde{b}) \neq X_0$.
  - If $\tilde{b} \in X$, (6.22) holds with $m = 0$ and $b^* = \tilde{b}$.
  - Else $\mu(\tilde{b}) \neq X_0$. Since $\overline{m} \geq 1$, $[\tilde{b}, X_0] \in \mathcal{B}$, so $\tilde{b} < X_0$. Since $\lambda(\tilde{b}) < \tilde{b}$, this proves $\lambda(\tilde{b}) \neq X_0$. So (6.22) holds with $m = 0$ and $b^* = \tilde{b}$.

Hence, the decomposition (6.22) always holds. □

We now turn to asymmetric estimates for the coordinates of the second kind, which, contrary to Lemma 2.58, isolate the role of X_0 associated with the implicit control $u_0 = 1$.

Lemma 6.6. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$, $\mathcal{B}$ a generalized Hall basis of $\mathcal{L}(X)$ and $(\xi_b)_{b \in \mathcal{B}}$ the associated coordinates of the second kind. For every $k \in \mathbb{N}^*$, there exists $c_k \geq 1$ such that, for each $b \in \mathcal{B}$ with $n(b) = k$, $T > 0$, $u \in L^1((0,T); \mathbb{K}^q)$ and $t \in [0,T]$,*

$$|\xi_b(t; 1, u)| \leq \|u\|_{L_t^1}^k \frac{(c_k t)^{n_0(b)}}{n_0(b)!} \tag{6.23}$$

and

$$|\dot{\xi}_b(t; 1, u)| \leq \begin{cases} k|u(t)|\|u\|_{L_t^1}^{k-1} & \text{when } n_0(b) = 0, \\ \|u\|_{L_t^1}^{k-1} \Big(kt|u(t)| + n_0(b)\|u\|_{L^1(0,t)} \Big) \frac{c_k (c_k t)^{n_0(b)-1}}{n_0(b)!} & \text{when } n_0(b) > 0. \end{cases} \tag{6.24}$$

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 (6.24) holds on $[0,T]$, then so does (6.23) 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 $b \in \mathcal{B}$ such that $n(b) = k$. We assume that (6.23) 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

$$|\dot{\xi}_{\tilde{b}}(t)| = |\xi_b(t)\dot{\xi}_{X_0}(t)| \leq \|u\|^k_{L^1(0,t)} \frac{(c_k t)^{n_0(b)}}{n_0(b)!}. \tag{6.25}$$

Hence $\tilde{b}$ satisfies (6.24) (and (6.23) 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 $b \in \mathcal{B}$ with $n(b) = k$. We assume that (6.24) 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 (6.24) with $c_k \leftarrow 1$. If $n_0(b) = 1$, we have

$$\begin{aligned}
|\dot{\xi}_{\tilde{b}}(t)| &= \frac{1}{m!}|\xi^m_{X_0}(t)\dot{\xi}_b(t)| \\
&\leq \frac{t^m}{m!}\|u\|^{k-1}_{L^1_t}\Big(kt|u(t)| + n_0(b)\|u\|_{L^1_t}\Big)\frac{c_k(c_k t)^{n_0(b)-1}}{n_0(b)!} \\
&\leq \|u\|^{k-1}_{L^1_t}\Big(kt|u(t)| + (m+n_0(b))\|u\|_{L^1_t}\Big)2^{m+n_0(b)}c_k^{n_0(b)}\frac{t^{m+n_0(b)-1}}{(m+n_0(b))!} \\
&\leq \|u\|^{k-1}_{L^1_t}\Big(kt|u(t)| + n_0(\tilde{b})\|u\|_{L^1_t}\Big)(2c_k)^{n_0(\tilde{b})}\frac{t^{n_0(\tilde{b})-1}}{n_0(\tilde{b})!}
\end{aligned} \tag{6.26}$$

because $n_0(\tilde{b}) = m + n_0(b)$ and $c_k \geq 1$. So $\tilde{b}$ satisfies (6.24) 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 $b \in \{X_1, \ldots, X_q\}$. By Lemma 6.5, Step 1 and Step 2, we deduce that (6.23) and (6.24) 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 6.5, Step 1 and Step 2, we can assume that $b = \mathrm{ad}^m_{b_1}(b_2)$ 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 (3.2) with $a \leftarrow (m+1)$, we obtain, when $n_0(b_2) > 0$,

$$\begin{aligned}
\Big|\dot{\xi}_b(t)\Big| &= \left|\frac{1}{m!}\xi^m_{b_1}(t)\dot{\xi}_{b_2}(t)\right| \\
&\leq \frac{1}{m!}\left(\|u\|^{k_1}_{L^1_t}\frac{(c_{k_1}t)^{n_0(b_1)}}{n_0(b_1)!}\right)^m \|u\|^{k_2-1}_{L^1_t}\Big(k_2 t|u(t)| + n_0(b_2)\|u\|_{L^1_t}\Big)\frac{c_{k_2}(c_{k_2}t)^{n_0(b_2)-1}}{n_0(b_2)!} \\
&\leq \|u\|^{k-1}_{L^1_t}\Big(kt|u(t)| + n_0(b)\|u\|_{L^1_t}\Big)2^{mn_0(b)}c_{k_1}^{mn_0(b_1)}c_{k_2}^{n_0(b_2)}\frac{t^{n_0(b)-1}}{n_0(b)!}.
\end{aligned} \tag{6.27}$$

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. □

These estimates allow to prove the main result of this section.

Proposition 6.7. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$, $\mathcal{B}$ a generalized Hall basis of $\mathcal{L}(X)$ and $(\xi_b)_{b\in\mathcal{B}}$ the associated coordinates of the second kind. Let $M \in \mathbb{N}$, $r, \delta > 0$, $f_0, \dots, f_q \in \mathcal{C}^{\omega,r}_{4\delta}$. There exists $\eta, C_M > 0$ such that, for every $u \in L^1((0,T);\mathbb{K}^q)$ with $T \le \eta$ and $\|u\|_{L^1(0,T)} \le \eta$, the ordered product of the $e^{\xi_b(t;1,u)f_b}$ over the infinite set $\mathcal{B}\cap S_M$ converges uniformly on B_δ and, for each $t \in [0,T]$ and $p \in B_\delta$,*

$$\left| x(t;f,u,p) - \overrightarrow{\prod_{b\in\mathcal{B}\cap S_M}} e^{\xi_b(t;1,u)f_b} p \right| \le C_M \|u\|^{M+1}_{L^1(0,t)}. \tag{6.28}$$

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 6.6. We define

$$C_* := \frac{18\,|||f|||_r}{r} \max_{k\in[\![1,2M]\!]} c_k, \tag{6.29}$$

$$\eta := \min\left\{ \frac{\delta}{2\|f\|_{\mathcal{C}^1}}, \frac{\min\{1,\delta\}}{2C_*(q+1)M!(1+r)} \right\} \tag{6.30}$$

$$C_M := e^{2\delta}(1+r)(2M)!(q+1)^{M+1}C_*^{M+1}. \tag{6.31}$$

For $t \in [0,T]$ and $u \in L^1((0,T);\mathbb{K}^q)$ with $T \le \eta$ and $\|u\| \le \eta$, using (6.30),

$$t\|f_0\|_{\mathcal{C}^0} + \sum_{i=1}^{q} \|u_i\|_{L^1(0,t)}\|f_i\|_{\mathcal{C}^0} \le \eta\|f\|_{\mathcal{C}^0} \le \delta. \tag{6.32}$$

Hence, for each $p \in B_\delta$, $x(t;f,u,p) \in B_{2\delta}$.

Strategy. Since the product involved in (6.28) 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, $L \ge M+1$ is fixed. In Step 5, we take the limit $L \to +\infty$.

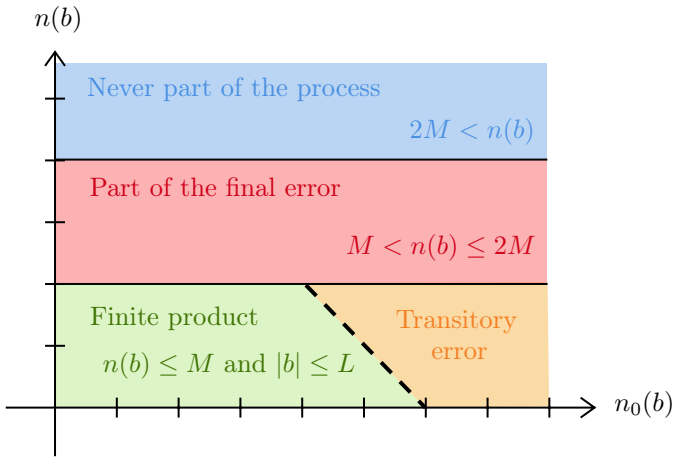

Figure 1: Decomposition of $\mathcal{B}$ along the Lazard elimination process for the product on $\mathcal{B}\cap S_M$.

Step 0: Preliminary estimates. First, using estimate (6.24) from Lemma 6.6, for each $b \in \mathcal{B}$ with $n(b) = k$, one has in particular

$$\|\dot{\xi}_b\|_{L^1} \leq \|u\|^k \frac{(c_k t)^{n_0(b)}}{n_0(b)!}. \tag{6.33}$$

Taking into account that for every $m \in \mathbb{N}^*$, $|\mathcal{B}_m| \leq (q+1)^m$ and using the analytic estimate (3.23), we obtain the following estimate for the terms which can be part of the final error

$$\begin{aligned}
&\sum_{b \in \mathcal{B} \cap (S_{2M} \setminus S_M)} \|\dot{\xi}_b\|_{L^1} \|f_b\|_{\mathcal{C}^1} \\
&\qquad \leq \sum_{k=M+1}^{2M} \sum_{n_0=0}^{+\infty} |\mathcal{B}_{k+n_0}| \|u\|^k \frac{(c_k t)^{n_0}}{n_0!} (1+r) \left(\frac{9 \|\!|\!| f |\!|\!|_r}{r} \right)^{k+n_0} (k+n_0-1)! \\
&\qquad \leq (1+r)(2M-1)! \sum_{k=M+1}^{2M} ((q+1) C_* \|u\|)^k \sum_{n_0=0}^{+\infty} ((q+1) C_* T)^{n_0} \\
&\qquad \leq (1+r)(2M)! (q+1)^{M+1} C_*^{M+1} \|u\|^{M+1},
\end{aligned} \tag{6.34}$$

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

$$\begin{aligned}
&\sum_{b \in \mathcal{B} \cap S_M} \|\dot{\xi}_b\|_{L^1} \|f_b\|_{\mathcal{C}^1} \\
&\qquad \leq T \|f_0\|_{\mathcal{C}^1} + \sum_{k=1}^{M} \sum_{n_0=0}^{+\infty} |\mathcal{B}_{k+n_0}| \|u\|^k \frac{(c_k t)^{n_0}}{n_0!} (1+r) \left(\frac{9 \|\!|\!| f |\!|\!|_r}{r} \right)^{k+n_0} (k+n_0-1)! \\
&\qquad \leq T \|f_0\|_{\mathcal{C}^1} + (1+r)(M-1)! \sum_{k=1}^{M} ((q+1) C_* \|u\|)^k \sum_{n_0=0}^{+\infty} ((q+1) C_* T)^{n_0} \\
&\qquad \leq T \|f_0\|_{\mathcal{C}^1} + (1+r) M! (q+1) C_* \|u\| \leq \delta.
\end{aligned} \tag{6.35}$$

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 (6.35), we have

$$\sum_{b \in \mathcal{B} \cap S_M} |\xi_b(t)| \, \|f_b\|_{\mathcal{C}^1} \leq \sum_{b \in \mathcal{B} \cap S_M} \|\dot{\xi}_b\|_{L^1} \|f_b\|_{\mathcal{C}^1} \leq \delta \tag{6.36}$$

and Lemma 5.28 gives the conclusion of Step 1.

Step 2: Lazard structure on $\mathcal{B}_{[\![1,L]\!]} \cap S_M$. We use the notations of Definition 2.49 to describe $\mathcal{B}_{[\![1,L]\!]}$. There exists $m \in \mathbb{N}$ and an extraction ϕ such that

$$\mathcal{B}_{[\![1,L]\!]} \cap S_M = \{b_{\phi(1)} < \cdots < b_{\phi(m+1)}\}. \tag{6.37}$$

Let $i \in [\![1, m+1]\!]$ and $n = \phi(i)$. By Definition 2.49, there exists a unique factorization

$$b_{\phi(i)} = b_n = \mathrm{ad}_{b_{n-1}}^{j_{n-1}} \cdots \mathrm{ad}_{b_1}^{j_1}(b_0) \tag{6.38}$$

where $b_0 \in X$, $j_1, \ldots, j_{n-1} \in \mathbb{N}$ (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 $X_1, \ldots, X_q$, thus it cannot be involved in the factorization of b_n. This proves that

$$\begin{cases}
b_{\phi(1)} \in \widetilde{Y}_0 := X, \\
b_{\phi(2)} \in \widetilde{Y}_1 := \{\mathrm{ad}_{b_{\phi(1)}}^j(v); j \in \mathbb{N}, v \in \widetilde{Y}_0 \setminus \{b_{\phi(1)}\}\}, \\
\cdots \\
b_{\phi(m+1)} \in \widetilde{Y}_m := \{\mathrm{ad}_{b_{\phi(m)}}^j(v); j \in \mathbb{N}, v \in \widetilde{Y}_{m-1} \setminus \{b_{\phi(m)}\}\},
\end{cases} \tag{6.39}$$

$$\mathcal{B}_{[\![1,L]\!]} \cap S_M \cap \widetilde{Y}_{m+1} = \emptyset, \tag{6.40}$$

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 (6.39) and (6.40) with $\phi = \mathrm{Id}$. Let $x_0(t) := x(t)$. By (6.36), for every $j \in [\![1, m+1]\!]$,

$$x_j(t) := e^{-\xi_{b_j}(t) f_{b_j}} \cdots e^{-\xi_{b_1}(t) f_{b_1}} x(t) \tag{6.41}$$

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

$$(\mathcal{H}_j) : \begin{cases} \dot{x}_j(t) = \sum_{b \in \mathcal{B}_{[\![1,L]\!]} \cap S_M \cap \widetilde{Y}_j} \dot{\xi}_b(t) f_b(x_j(t)) + \varepsilon_j(t), \\ x_j(0) = p, \end{cases} \tag{6.42}$$

where

$$\|\varepsilon_j\|_{L^1} \leq e^{|\xi_{b_j}(t)| \|f_{b_j}\|_{\mathcal{C}^1}} \|\varepsilon_{j-1}\|_{L^1} + \sum_{\tilde{b} \in Z_j} \|\dot{\xi}_{\tilde{b}}\|_{L^1} \|f_{\tilde{b}}\|_{\mathcal{C}^0}, \tag{6.43}$$

where $Z_j \subset (\mathcal{B} \cap S_{2M}) \setminus (\mathcal{B}_{[\![1,L]\!]} \cap S_M)$ is defined in (6.50).

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

$$x_j(t) = e^{-\xi_{b_j}(t) f_{b_j}}(x_{j-1}(t)) = \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) \tag{6.44}$$

and thus that

$$\dot{x}_j(t) = \sum_{b \in \mathcal{B}_{[\![1,L]\!]} \cap S_M \cap \widetilde{Y}_{j-1} \setminus \{b_j\}} \dot{\xi}_b(t) \left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)) + \widetilde{\varepsilon}_{j-1}(t), \tag{6.45}$$

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

$$\varepsilon_j(t) := \sum_{b \in \mathcal{B}_{[\![1,L]\!]} \cap S_M \cap \widetilde{Y}_{j-1} \setminus \{b_j\}} \overline{\varepsilon}^j_b(t) + \widetilde{\varepsilon}_{j-1}(t) \tag{6.46}$$

where, for every $b \in \mathcal{B}_{[\![1,L]\!]} \cap S_M \cap \widetilde{Y}_{j-1} \setminus \{b_j\}$,

$$\overline{\varepsilon}^j_b(t) := \dot{\xi}_b(t) \left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)) - \sum_{k=0}^{h(b)-1} \dot{\xi}_b(t) \frac{\xi^k_{b_j}(t)}{k!} f_{\mathrm{ad}^k_{b_j}(b)}(x_j(t)) \tag{6.47}$$

where $h(b) \in \mathbb{N}^*$ is the maximal integer such that

$$n(b) + (h(b)-1)\, n(b_j) \leq M \qquad \text{and} \qquad |b| + (h(b)-1)|b_j| \leq L. \tag{6.48}$$

By (3.41),

$$|\overline{\varepsilon}^j_b(t)| \leq |\dot{\xi}_b(t)| \frac{|\xi_{b_j}(t)|^{h(b)}}{h(b)!} \|f_{\mathrm{ad}^{h(b)}_{b_j}(b)}\|_{\mathcal{C}^0} = |\dot{\xi}_{\tilde{b}}(t)| \|f_{\tilde{b}}\|_{\mathcal{C}^0}, \tag{6.49}$$

for $\tilde{b} := \mathrm{ad}^{h(b)}_{b_j}(b)$. Hence, (6.43) holds with

$$Z_j := \{\mathrm{ad}^{h(b)}_{b_j}(b);\ \ b \in \mathcal{B}_{[\![1,L]\!]} \cap S_M \cap \widetilde{Y}_{j-1} \setminus \{b_j\}\}. \tag{6.50}$$

This yields $Z_j \subset (\mathcal{B} \cap S_{2M}) \setminus (\mathcal{B}_{[\![1,L]\!]} \cap S_M)$ thanks to (6.48).

Step 4: Proof of an estimate on the finite product over $\mathcal{B}_{[\![1,L]\!]} \cap S_M$. By (6.43), (6.36) and (6.34), we have

$$\begin{aligned}\|\varepsilon_{m+1}\|_{L^1} &\leq e^{\delta} \sum_{b\in\mathcal{B}\cap(S_{2M}\setminus S_M)} \|\dot\xi_b\|_{L^1}\|f_b\|_{\mathcal{C}^0} + e^{\delta} \sum_{b\in(\mathcal{B}\cap S_M)\setminus\mathcal{B}_{[\![1,L]\!]}} \|\dot\xi_b\|_{L^1}\|f_b\|_{\mathcal{C}^0} \\ &\leq e^{-\delta} C_M \|u\|^{M+1} + o_{L\to+\infty}(1),\end{aligned} \tag{6.51}$$

because the series in (6.35) converges. We deduce from (6.42) and (6.40) that

$$\left| \overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M}} e^{-\xi_b(t;1,u)f_b} x(t) - p \right| = |x_{m+1}(t) - p| \leq e^{-\delta} C_M \|u\|^{M+1} + o_{L\to+\infty}(1) \tag{6.52}$$

By (6.36), 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 (6.52),

$$\left| x(t) - \overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M}} e^{-\xi_b(t;1,u)f_b} p \right| \leq C_M \|u\|^{M+1} + o_{L\to+\infty}(1) \tag{6.53}$$

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 (6.53), we obtain the conclusion of Proposition 6.7. □

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

$$\dot x(t) = f_0(x(t)) + u(t) f_1(x(t)) \quad \text{and} \quad x(0) = p, \tag{7.1}$$

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. [10]) and might also be useful for applications to stochastic ODEs where the input is a noise with low regularity (see e.g. [11]).

Definition 7.1 (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 7.2. *Let $\delta>0$, $f_0, f_1 \in \mathcal{C}^\omega_{3\delta}$ and $\eta^*>0$ small enough so that the two following maps are well defined and (globally) analytic*

$$\Phi_1 : \begin{cases} [-\eta^*,\eta^*]\times B_{2\delta} & \to B_{3\delta} \\ (\tau, q) & \mapsto e^{\tau f_1}(q) \end{cases} \quad \text{and} \quad F : \begin{cases} B_{2\delta}\times[-\eta^*,\eta^*] & \to \mathbb{K}^d \\ (q,\tau) & \mapsto (\Phi_1(-\tau)_* f_0)(q). \end{cases} \tag{7.2}$$

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*

$$\begin{cases} \dot{x}_1(t) = F(x_1(t), U(t)), \\ x_1(0) = p, \end{cases} \tag{7.3}$$

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 $\|U\|_{L^\infty} \leq \eta^*$,*

$$x(t; f, u, p) = \Phi_1\big(U(t); x_1(t; F, U, p)\big). \tag{7.4}$$

Proof. The existence of η^* such that Φ_1 and F are well defined and globally analytic results from the third statement of Lemma 3.24. The analytic dependence of x_1 with respect to (p, U) is given by Lemma 3.10. By definition of x_1, the right-hand side of (7.4) solves 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 (7.1).

Proposition 7.3. *Let $\delta, r > 0$ and $f_0, f_1 \in \mathcal{C}^{\omega,r}_{3\delta}$. 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 $p \in B_\delta$,*

$$\varphi(x(t; f, u, p)) = \sum_{\substack{\ell \in \mathbb{N}, n \in \mathbb{N} \\ k \in \mathbb{N}^n}} \frac{U(t)^\ell}{\ell! k!} \left(\int_0^t U^k \right) \Big((f_1 \cdot \nabla)^\ell (\mathrm{ad}^{k_n}_{f_1}(f_0) \cdot \nabla) \cdots (\mathrm{ad}^{k_1}_{f_1}(f_0) \cdot \nabla) \Big) \Big(\varphi \Big)(p) \tag{7.5}$$

with the notation (6.7), 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 $p \in B_\delta$,

$$\begin{aligned} &\left| \varphi(x(t; f, u, p)) - \sum_{\substack{\ell \in \mathbb{N}, n \in \mathbb{N} \\ \ell + |k| \leq M}} \frac{U(t)^\ell}{\ell! k!} \left(\int_0^t U^k \right) \Big((f_1 \cdot \nabla)^\ell (\mathrm{ad}^{k_n}_{f_1}(f_0) \cdot \nabla) \cdots (\mathrm{ad}^{k_1}_{f_1}(f_0) \cdot \nabla) \Big) \Big(\varphi \Big)(p) \right| \\ &\leq C^{M+1} \left(|U(t)|^{M+1} + \int_0^t |U|^{M+1} \right) \end{aligned} \tag{7.6}$$

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 7.2, $\|f\| := |||f_0|||_r + |||f_1|||_r$ and

$$\eta := \min\left\{ T, \eta^*, \frac{\delta}{\|f\|}, \frac{r}{28\|f\|} \right\}. \tag{7.7}$$

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 (7.4) and (7.7), $x(t; f, u, p) \in B_{3\delta}$.

Step 1: Proof of the absolute convergence in (7.5) uniformly with respect to $p \in B_\delta$. Let $r' := r/e$. Then, by Lemma 3.15, for every $k \in \mathbb{N}$, $\mathrm{ad}^k_{f_1}(f_0) \in \mathcal{C}^{\omega,r'}_{3\delta}$ and

$$\left| \left| \left| \mathrm{ad}^k_{f_1}(f_0) \right| \right| \right|_{r'} \leq \frac{k!}{e} \left(\frac{9}{r} \right)^k \|f\|^{k+1}. \tag{7.8}$$

Thus, by (3.18),

$$\begin{aligned}
&\left|\Big((f_1\cdot\nabla)^\ell(\mathrm{ad}_{f_1}^{k_n}(f_0)\cdot\nabla)\cdots(\mathrm{ad}_{f_1}^{k_1}(f_0)\cdot\nabla)\Big)\Big(\varphi\Big)(p)\right| \\
&\qquad\leq (n+\ell)!\left(\frac{5}{r'}\right)^{n+\ell}|||f_1|||_{r'}^{\ell}\left|\left|\left|\mathrm{ad}_{f_1}^{k_n}(f_0)\right|\right|\right|_{r'}\cdots\left|\left|\left|\mathrm{ad}_{f_1}^{k_1}(f_0)\right|\right|\right|_{r'} \\
&\qquad\leq (n+\ell)!\left(\frac{14}{r}\right)^{n+\ell}\|f\|^{\ell}\frac{k_n!}{e}\left(\frac{9}{r}\right)^{k_n}\|f\|^{k_n+1}\cdots\frac{k_1!}{e}\left(\frac{9}{r}\right)^{k_1}\|f\|^{k_1+1} \\
&\qquad\leq e^{-n}(n+\ell)!k_n!\cdots k_1!\left(\frac{14\|f\|}{r}\right)^{n+\ell+k_1+\cdots+k_n}.
\end{aligned}\tag{7.9}$$

Moreover, recalling notation (6.7),

$$\left|\frac{U(t)^\ell}{\ell!k!}\int_0^t U^k\right| = \left|\frac{U(t)^\ell}{\ell!}\int_{\mathcal{T}_{(n)}(t)}\frac{U(\tau_n)^{k_n}\cdots U(\tau_1)^{k_1}}{k_n!\cdots k_1!}\,\mathrm{d}\tau\right| \leq \|U\|_{L^\infty}^{\ell+k_1+\cdots+k_n}\frac{t^n}{n!}\frac{1}{\ell!k_n!\cdots k_1!}. \tag{7.10}$$

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

$$\begin{aligned}
&\left(\frac{t}{e}\right)^n\frac{(n+\ell)!}{n!\ell!}\left(\frac{14\|f\|}{r}\right)^{n+\ell+k_1+\cdots+k_n}\|U\|_{L^\infty}^{\ell+k_1+\cdots+k_n} \\
&\qquad\leq\left(\frac{t}{e}\right)^n 2^{n+\ell}\left(\frac{14\|f\|}{r}\right)^{n+\ell+k_1+\cdots+k_n}\|U\|_{L^\infty}^{\ell+k_1+\cdots+k_n} \\
&\qquad\leq\left(\frac{28t\|f\|}{er}\right)^n\left(\frac{28\|f\|}{r}\|U\|_{L^\infty}\right)^{\ell+k_1+\cdots+k_n}
\end{aligned}\tag{7.11}$$

which is ensured by (7.7).

Step 2: Proof of (7.5) and (7.6). Applying Lemma 3.20 and Proposition 6.2 we get

$$\begin{aligned}
\varphi(x(t;f,u,p)) &= \varphi\left(e^{U(t)f_1}x_1(t;F,U,p)\right) \\
&= \sum_{\ell=0}^{+\infty}\frac{U(t)^\ell}{\ell!}(f_1\cdot\nabla)^\ell\varphi(x_1(t;F,U,p)) \\
&= \sum_{\ell=0}^{+\infty}\frac{U(t)^\ell}{\ell!}(f_1\cdot\nabla)^\ell\sum_{\substack{n\in\mathbb{N}\\k\in\mathbb{N}^n}}\frac{1}{k!}\left(\int_0^t U^k\right)\Big((\mathrm{ad}_{f_1}^{k_n}(f_0)\cdot\nabla)\cdots(\mathrm{ad}_{f_1}^{k_1}(f_0)\cdot\nabla)\Big)\Big(\varphi\Big)(p)
\end{aligned}\tag{7.12}$$

The bound proved in Step 1 allows to exchange the differential operator $(f_1\cdot\nabla)^\ell$ and the second sum, which proves (7.5). To prove (7.6), one bounds the queue of the series thanks to (7.9) and the following consequence of Hölder's inequality, valid when $\ell+|k|\geq(M+1)$

$$\left|U(t)^\ell\int_{\mathcal{T}_{(n)}(t)}U(\tau_n)^{k_n}\cdots U(\tau_1)^{k_1}\,\mathrm{d}\tau\right| \leq C(\eta)\left(|U(t)|^{M+1}+\int_0^t|U|^{M+1}\right). \tag{7.13}$$

□

Remark 7.4. *The bound* (7.6) *between the exact solution and the truncated Chen-Fliess series (in its' original formulation) is used by Stefani in [64, Lemma 3.1 and Corollary 3.1]. Our proof is both different and shorter.*

Remark 7.5. *Equality* (7.5) *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 6.1 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 7.6. *Let $\delta > 0$ and $f_0, f_1 \in \mathcal{C}^\omega_{3\delta}$. 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 $p \in B_\delta$,*

$$\left| x(t; f, u, p) - e^{\mathcal{Z}_M(t,f,u)} e^{tf_0} p \right| \leq C_M \left(|U(t)|^{M+1} + \int_0^t |U|^{M+1} \right). \tag{7.14}$$

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 (7.3). Since, by (7.4), $x(t) = e^{U(t)f_1}(x_1(t))$, the desired estimate then relies on the following decomposition

$$\begin{aligned} x(t; f, u, p) - e^{\mathcal{Z}_M(t,f,u)} e^{tf_0} p = x(t; f, u, p) - e^{U(t)f_1} e^{\mathcal{Y}_M(t,f,U)} e^{tf_0} p \\ + e^{U(t)f_1} e^{\mathcal{Y}_M(t,f,U)} e^{tf_0} p - e^{\mathcal{Z}_M(t,f,u)} e^{tf_0} p. \end{aligned} \tag{7.15}$$

Using Proposition 7.7 and Proposition 7.11 (see further) for the first and second lines, we get

$$\left| x(t; f, u, p) - e^{\mathcal{Z}_M(t,f,u)} e^{tf_0} p \right| \leq C_M \left(\|U\|_{L^1}^{M+1} + |U(t)|^{M+1} + \|U\|_{L^{M+1}}^{M+1} \right) \tag{7.16}$$

which gives the conclusion since $\|U\|_{L^1_t} \leq t^{\frac{M}{M+1}} \|U\|_{L^{M+1}_t}$. □

In Section 7.3.1, we define $\mathcal{Y}_M(t, f, U)$ and prove in Proposition 7.7 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 7.11 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 4.8 for the Magnus expansion in the interaction picture to obtain an approximation of the auxiliary state.

Proposition 7.7. *Let $\delta, \rho > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,\rho}_{3\delta}$. 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 $\|U\|_{L^\infty} \leq \eta_M$,*

$$\left| x(t; f, u, p) - e^{U(t)f_1} e^{\mathcal{Y}_M(t,f,U)} e^{tf_0} p \right| \leq C_M \|U\|_{L^1(0,t)}^{M+1} \tag{7.17}$$

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

$$G_t(s, y) := \sum_{\substack{k \in \mathbb{N}^* \\ \ell \in \mathbb{N}}} \frac{(s-t)^\ell}{\ell!} \frac{U(s)^k}{k!} \, \mathrm{ad}^\ell_{f_0} \, \mathrm{ad}^k_{f_1}(f_0)(y) \tag{7.18}$$

and this sum converges absolutely in $\mathcal{C}^{\omega,\rho'}_{3\delta}$ with $\rho' = \rho/e$. Moreover,

$$\begin{aligned} \mathcal{Y}_M(t, f, U) = \sum \frac{(-1)^{m-1}}{rm} \int_{\mathcal{T}_{\mathbf{r}}(t)} \frac{(\tau_r - t)^{\ell_r}}{\ell_r!} \frac{U(\tau_r)^{k_r}}{k_r!} \cdots \frac{(\tau_1 - t)^{\ell_1}}{\ell_1!} \frac{U(\tau_1)^{k_1}}{k_1!} \, \mathrm{d}\tau \\ \left[\cdots \left[\mathrm{ad}^{\ell_r}_{f_0}(\mathrm{ad}^{k_r}_{f_1}(f_0)), \mathrm{ad}^{\ell_{r-1}}_{f_0}(\mathrm{ad}^{k_{r-1}}_{f_1}(f_0)) \right], \ldots, \mathrm{ad}^{\ell_1}_{f_0}(\mathrm{ad}^{k_1}_{f_1}(f_0)) \right] \end{aligned} \tag{7.19}$$

where the sum is taken over $r \in [\![1, M]\!]$, $m \in [\![1, r]\!]$, $\mathbf{r} \in \mathbb{N}^m_r$, $\ell_1, \ldots, \ell_r \in \mathbb{N}$, $k_1, \ldots, k_r \in \mathbb{N}^*$ and the sum converges absolutely in $\mathcal{C}^{\omega,\rho'}_{3\delta}$.

Proof. Step 1: Convergence in (7.18) and (7.19). By (3.21), for every $s \in [0,t]$,

$$\left|\!\left|\!\left| \frac{(s-t)^\ell}{\ell!} \frac{U(s)^k}{k!} \operatorname{ad}_{f_0}^\ell \operatorname{ad}_{f_1}^k (f_0) \right|\!\right|\!\right|_{\rho'} \leq t^\ell \|U\|_{L^\infty}^k \frac{(k+\ell)!}{\ell! k!} \left(\frac{9}{\rho}\right)^{k+\ell} |||f|||_\rho^{k+\ell+1} \tag{7.20}$$

thus the sum in (7.18) converges absolutely in $\mathcal{C}^{\omega,\rho'}_{3\delta}$ when t and $\|U\|_{L^\infty}$ are $< \frac{\rho}{18|||f|||_\rho}$.

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 (3.2) and the non-decreasing of $q \in [\![1,\infty]\!] \mapsto \|\cdot\|_{L_t^q}$ for $t \in [0,1]$, we get

$$\begin{aligned} &\left| \int_{\mathcal{T}_{\mathbf{r}}(t)} \frac{(\tau_r - t)^{\ell_r}}{\ell_r!} \frac{U(\tau_r)^{k_r}}{k_r!} \cdots \frac{(\tau_1 - t)^{\ell_1}}{\ell_1!} \frac{U(\tau_1)^{k_1}}{k_1!} \, \mathrm{d}\tau \right| (r + |\ell| + |k| - 1)! \left(\frac{9\|f\|}{\rho}\right)^{r+|\ell|+|k|-1} \\ &\quad \leq (2^{r-1}t)^{|\ell|} (2^{r-1} \|U\|_{L_t^{|k|}})^{|k|} \left(\frac{36\|f\|}{\rho}\right)^{r+|\ell|+|k|-1} (r-1)! \end{aligned} \tag{7.21}$$

Thus, by (3.21), the sum in (7.19) converges absolutely in $\mathcal{C}^{\omega,\rho'}_{3\delta}$ when t and $\|U\|_{L^\infty}$ are $< \frac{\rho 2^{-M}}{18|||f|||_\rho}$.

Step 2: Proof of (7.17). Let T, η^* and F as in Proposition 7.2. We introduce the function $F_1 : [0,T] \times B_{2\delta} \to \mathbb{K}^d$ defined by

$$F_1(t,y) := F(y, U(t)) - f_0(y) = \sum_{j=1}^{+\infty} \frac{U(t)^j}{j!} \operatorname{ad}_{f_1}^j (f_0)(y) \tag{7.22}$$

where the sum converges in $\mathcal{C}^{\omega,\rho'}_{2\delta}$ 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 (7.22) satisfies

$$\|F_1\|_{L^1((0,t),\mathcal{C}^{M^2})} \leq C\|U\|_{L^1(0,t)} \leq CT\|U\|_{L^\infty(0,t)}. \tag{7.23}$$

Let Θ_M be as in Proposition 4.8 and

$$\eta_M = \min\left\{1, \eta^*, \frac{\rho 2^{-M}}{36\,|||f|||_\rho}, \Theta_M(T, \|f_0\|_{\mathcal{C}^{M^2+1}}) \frac{\min\{1,\delta\}}{CT}\right\}. \tag{7.24}$$

Let $p \in B_\delta$, $t \in [0, \eta_M]$, $u \in L^1((0,t);\mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta_M$. Then, the convergences of Step 1 hold and $\|F_1\|_{L^1((0,t),\mathcal{C}^{M^2})} \leq \Theta_M(T, \|f_0\|_{\mathcal{C}^{M^2+1}}) \min\{1;\delta\}$ thus we can apply Proposition 4.8 and Proposition 4.9 to the equation $\dot{x}_1 = f_0(x_1) + F_1(t, x_1)$

$$\left| x_1(t; F, U, p) - e^{\mathcal{Y}_M(t,f,U)} e^{tf_0} p \right| \leq C_M \|G_t\|^{M+1}_{L^1(0,t),\mathcal{C}^{M^2})}. \tag{7.25}$$

Moreover, there exists C' (depending only on η^*, f_0, f_1) such that

$$\|G_t\|_{L^1((0,t),\mathcal{C}^{M^2})} \leq C'\|U\|_{L^1(0,t)}. \tag{7.26}$$

Thus, we get (7.17) by applying the $e^{\eta^*\|f_1\|_{\mathcal{C}^1}}$-Lipschitz map $e^{U(t)f_1}$ to (7.25). □

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 7.8. *Let $\delta, \rho > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,\rho}_{3\delta}$. 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 (7.19) taken over $r \in [\![1, M]\!]$, $m \in [\![1, r]\!]$, $\mathbf{r} \in \mathbb{N}_r^m$, $\ell_1, \ldots, \ell_r \in \mathbb{N}$ *and* $k_1, \ldots, k_r \in \mathbb{N}^*$ *such that* $k_1 + \ldots + k_r = j$, *converges absolutely in* $\mathcal{C}_{3\delta}^{\omega,\rho'}$ *and its sum, denoted* $\mathcal{Y}_M^j(t, f, U)$, *satisfies*

$$\left|\!\left|\!\left|\mathcal{Y}_M^j(t, f, U)\right|\!\right|\!\right|_{\rho'} \leq C_M \left(\frac{\|U\|_{L_t^j}}{2\eta_M}\right)^j \tag{7.27}$$

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 in* $\mathcal{C}_{3\delta}^{\omega,\rho'}$.

Proof. Let $\eta_M > 0$ be as in Proposition 7.7, $t \in [0, \eta_M]$ 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 $\mathcal{C}_{3\delta}^{\omega,\rho'}$ because it is a subfamily of the one considered in Proposition 7.7. By (7.21), there exists $C_M > 0$ (independent of t and U) such that, for every $j \in \mathbb{N}^*$, (7.27) holds. The non-decreasing of $q \in [\![1, \infty]\!] \mapsto \|\cdot\|_{L_t^q}$ (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 7.9. *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*

$$\begin{cases} \dot{x}(t) = x(t)(X_0 + u(t)X_1), \\ x(0) = x^\star, \end{cases} \tag{7.28}$$

satisfies, for every $t \in \mathbb{R}_+$,

$$x(t) = x^\star \exp(tX_0) \exp(\mathcal{Y}_\infty(t, X, U)) \exp(U(t)X_1) \tag{7.29}$$

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*

$$\beta_t(s) = e^{-(t-s)X_0}\left(e^{U(s)X_1}X_0e^{-U(s)X_1} - X_0\right)e^{(t-s)X_0} = \sum_{\substack{k\in\mathbb{N}^*\\ \ell\in\mathbb{N}}} \frac{(s-t)^\ell}{\ell!}\frac{U(s)^k}{k!}\,\mathrm{ad}_{X_0}^\ell\,\mathrm{ad}_{X_1}^k(X_0) \tag{7.30}$$

i.e.

$$\begin{aligned}\mathcal{Y}_\infty(t, X, U) = \sum \frac{(-1)^{m-1}}{rm} \int_{\mathcal{T}_{\mathbf{r}}(t)} &\frac{(\tau_r - t)^{\ell_r}}{\ell_r!}\frac{U(\tau_r)^{k_r}}{k_r!} \cdots \frac{(\tau_1 - t)^{\ell_1}}{\ell_1!}\frac{U(\tau_1)^{k_1}}{k_1!}\,\mathrm{d}\tau \\ &\left[\cdots\left[\mathrm{ad}_{X_0}^{\ell_r}(\mathrm{ad}_{X_1}^{k_r}(X_0)), \mathrm{ad}_{X_0}^{\ell_{r-1}}(\mathrm{ad}_{X_1}^{k_{r-1}}(X_0))\right], \ldots, \mathrm{ad}_{X_0}^{\ell_1}(\mathrm{ad}_{X_1}^{k_1}(X_0))\right]\end{aligned} \tag{7.31}$$

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 2.26 has been generalized to an infinite alphabet in the proof of Theorem 2.39, it is possible to generalize Theorem 2.39 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

$$\dot{x}_1 = x_1(t)e^{U(t)X_1}X_0e^{-U(t)X_1} = x_1(t)\left(X_0 + \sum_{k\in\mathbb{N}^*}\frac{U(t)^k}{k!}\,\mathrm{ad}_{X_1}^k(X_0)\right). \tag{7.32}$$

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 2.39 (adapted to an infinite alphabet) and the algebra homomorphism sending Y_k to $\mathrm{ad}^k_{X_1}(X_0)$ prove that

$$x_1(t) = x^\star \exp(tX_0) \exp(\mathcal{Y}_\infty(t, X, U)). \tag{7.33}$$

which gives the conclusion. □

We now use the formal expansion (7.29) to obtain an alternative formula for $\mathcal{Z}_\infty(t, X, u)$ defined by Theorem 2.39, in terms of the primitive of the scalar input. For $r, \nu \in \mathbb{N}$, we introduce the finite dimensional subspace of $\mathcal{L}(X)$

$$\mathcal{L}_{r,\nu}(X) := \mathrm{span}\{\mathrm{eval}(b);\ \ b \in \mathrm{Br}(X), n_0(b) = \nu, n_1(b) = r\} \tag{7.34}$$

and $P_{r,\nu} : \widehat{\mathcal{L}}(X) \to \mathcal{L}_{r,\nu}(X)$ the associated canonical projection.

Proposition 7.10. *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 7.9 and $\mathcal{Z}_\infty(t, X, u)$ defined by Theorem 2.39. Then, in $\widehat{\mathcal{L}}(X)$,*

$$\mathcal{Z}_\infty(t, X, u) = \mathrm{CBHD}_\infty\left(\mathcal{Y}_\infty(t, X, U), U(t)X_1\right). \tag{7.35}$$

In particular, for every $M \in \mathbb{N}^*$, $r \in [\![1, M]\!]$ and $\nu \in \mathbb{N}$,

$$P_{r,\nu}\mathcal{Z}_M(t, X, u) = P_{r,\nu}\,\mathrm{CBHD}_M\left(\mathcal{Y}_M(t, X, U), U(t)X_1\right). \tag{7.36}$$

In this statement, CBHD_∞ is defined in Corollary 2.32, CBHD_M is its truncation used in Corollary 4.4 and $\mathcal{Z}_M(t, X, u)$ is defined in Theorem 2.39 and used in Proposition 4.8.

Proof. We deduce from Proposition 7.9 and Theorem 2.39 that

$$\exp(\mathcal{Z}_\infty(t, X, u)) = \exp(\mathcal{Y}_\infty(t, X, U)) \exp(U(t)X_1). \tag{7.37}$$

Thus Corollary 2.32 proves (7.35). Let $M \in \mathbb{N}^*$, $r \in [\![1, M]\!]$, $\nu \in \mathbb{N}$. We deduce from (7.35) that

$$P_{r,\nu}\mathcal{Z}_\infty(t, X, u) = P_{r,\nu}\,\mathrm{CBHD}_\infty\left(\mathcal{Y}_\infty(t, X, U), U(t)X_1\right). \tag{7.38}$$

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

$$P_{r,\nu}\,\mathrm{CBHD}_\infty\left(\mathcal{Y}_\infty(t, X, U), U(t)X_1\right) = P_{r,\nu}\,\mathrm{CBHD}_M\left(\mathcal{Y}_\infty(t, X, U), U(t)X_1\right). \tag{7.39}$$

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

$$P_{r,\nu}\,\mathrm{CBHD}_M\left(\mathcal{Y}_\infty(t, X, U), U(t)X_1\right) = P_{r,\nu}\,\mathrm{CBHD}_M\left(\mathcal{Y}_M(t, X, U), U(t)X_1\right), \tag{7.40}$$

which ends the proof of (7.36). □

7.3.3 Error formula for analytic vector fields

We prove in Proposition 7.11 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 7.11. *Let $\delta, \rho > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,\rho}_{3\delta}$. 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 $\|U\|_{L^\infty} \le \eta_M$,*

$$\left|e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}e^{tf_0}p - e^{\mathcal{Z}_M(t,f,u)}e^{tf_0}p\right| \le C_M\left(|U(t)|^{M+1} + \int_0^t |U|^{M+1}\right). \tag{7.41}$$

Proof. We split the difference as

$$\begin{aligned} e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}e^{tf_0}p - e^{\mathrm{CBHD}_M(\mathcal{Y}_M(t,f,U),U(t)f_1)}e^{tf_0}p \\ + e^{\mathrm{CBHD}_M(\mathcal{Y}_M(t,f,U),U(t)f_1)}e^{tf_0}p - e^{\mathcal{Z}_M(t,f,u)}e^{tf_0}p. \end{aligned} \tag{7.42}$$

Taking into account that $\|\mathcal{Y}_M(t,f,U)\|_{\mathcal{C}^{M^2}} \leq C\|U\|_{L^1(0,t)}$, the first line is bounded by Corollary 4.4. Using Grönwall's lemma and Proposition 7.12 bounds the second line. □

Proposition 7.12. *Let $\delta, \rho > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,\rho}_{3\delta}$ 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 $\|U\|_{L^\infty} \leq \eta_M$,*

$$|||\mathcal{Z}_M(t,f,u) - \mathrm{CBHD}_M\left(\mathcal{Y}_M(t,f,U),U(t)f_1\right)|||_{\rho''} \leq C_M\left(|U(t)|^{M+1} + \int_0^t |U|^{M+1}\right). \tag{7.43}$$

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 7.8, one can write

$$\mathcal{Y}_M(t,f,U) = \sum_{j=1}^{M}\mathcal{Y}^j_M(t,f,U) + \sum_{j>M}\mathcal{Y}^j_M(t,f,U) =: \overline{\mathcal{Y}}_M(t,f,U) + \mathcal{R}_M(t,f,U), \tag{7.44}$$

where the remainder satisfies $|||\mathcal{R}_M(t,f,U)|||_{\rho'} \leq C\|U\|^{M+1}_{L^{M+1}_t}$. By the triangular and Young inequalities, it is therefore sufficient to prove (7.43) with $\mathcal{Y}_M$ replaced by the finite truncation $\overline{\mathcal{Y}}_M(t,f,U)$.

Step 2: Identification at the free level. Let $\Lambda : \mathcal{L}(X) \to \mathcal{C}^\omega_{3\delta}$ be the homomorphism of Lie algebra such that $\Lambda(X_i) = f_i$. The relation (7.36) 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 $r \in [\![1,M]\!]$, $\nu \in \mathbb{N}$

$$P_{r,\nu}\mathcal{Z}_M(t,f,u) = P_{r,\nu}\,\mathrm{CBHD}_M\left(\overline{\mathcal{Y}}_M(t,f,U),U(t)f_1\right). \tag{7.45}$$

By definition

$$\mathcal{Z}_M(t,f,u) = \sum_{\nu\in\mathbb{N}}\sum_{r=1}^{M} P_{r,\nu}\mathcal{Z}_M(t,f,u) \tag{7.46}$$

where the sum converges in $\mathcal{C}^{\omega,\rho'}_{3\delta}$ for appropriate $\rho' \in (0,\rho)$, by Proposition 4.9. Thus, with the notations of (2.40),

$$\mathcal{Z}_M(t,f,u) - \mathrm{CBHD}_M\left(\overline{\mathcal{Y}}_M(t,f,U),U(t)f_1\right) = \sum_{jh_1+h_2>M} F_{2,h}(\mathcal{Y}^j_M(t,f,U),U(t)f_1), \tag{7.47}$$

where the sum is taken over $j,h_1,h_2 \in [\![1,M]\!]$.

Step 3: Proof of (7.43). From now on, $\eta_M > 0$ is given by Proposition 7.7 and Lemma 7.8, $t \in [0,\eta_M]$, $u \in L^1((0,t);\mathbb{K})$ is such that $\|U\|_{L^\infty} < \eta_M$ and $\rho'' \in (0,\rho')$. For each term in the finite sum (7.47), one has, thanks to Lemma 7.8,

$$\begin{aligned} \left|\left|\left|F_{2,h}(\mathcal{Y}^j_M(t,f,U),U(t)f_1)\right|\right|\right|_{\rho''} &\leq C\left|\left|\left|\mathcal{Y}^j_M(t,f,U)\right|\right|\right|^{h_1}_{\rho'}|||U(t)f_1|||^{h_2}_{\rho'} \\ &\leq C'\|U\|^{jh_1}_{L^j_t}|U(t)|^{h_2} \leq C'\|U\|^{jh_1}_{L^{M+1}_t}t^{h_1-\frac{jh_1}{M+1}}|U(t)|^{h_2} \end{aligned} \tag{7.48}$$

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 6.6 can be enhanced to involve only the primitive of the input, at least for generalized Hall bases where X_1 is minimal, which in turn improves the estimate of Proposition 6.7 (see Proposition 7.14 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 7.13. *Let $X = \{X_0, X_1\}$, $\mathcal{B}$ be a generalized 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 $t \in [0,T]$,*

$$|\xi_b(t;1,u)| \leq c_k \|U\|_{L^k_t}^k \frac{(c_k t)^{n_0(b)-1}}{(n_0(b)-1)!} \tag{7.49}$$

and

$$|\dot{\xi}_b(t;1,u)| \leq \begin{cases} c_k |U(t)|^k & \text{when } n_0(b) = 1, \\ c_k |U(t)|^k \frac{(c_k t)^{n_0(b)-1}}{(n_0(b)-1)!} + c_k^2 \|U\|_{L^k_t}^k \frac{(c_k t)^{n_0(b)-2}}{(n_0(b)-2)!} & \text{when } n_0(b) \geq 2. \end{cases} \tag{7.50}$$

Proof. As for Lemma 6.6, estimate (7.49) is obtained, for each b, by time integration of (7.50). Moreover, still as in Lemma 6.6, 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 (7.49) and (7.50) 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 6.5, when $n(b) = 1$, there exist $m, \overline{m} \in \mathbb{N}$ such that $b = \mathrm{ad}_{X_0}^m \overline{\mathrm{ad}}_{X_0}^{\overline{m}}(X_1)$. 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 6.5 and the previous invariant properties, we may assume that $b = \mathrm{ad}_{b_1}^m(b_2)$ 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
$$|\dot{\xi}_b(t)| = \frac{|U(t)|^m}{m!} \tag{7.51}$$
so (7.50) with $c_k = 1$ holds since $n_0(b) = 1$ and $k = m$.

- If $b_1 \neq X_1$, then b_1 satisfies (7.50) 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 (7.50) 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,
$$|\dot{\xi}_b(t)| \leq \frac{c_{k_1}^m \|U\|_{L^{k_1}_t}^{mk_1}}{m!} \frac{(c_{k_1} t)^{m\nu_1 - m}}{(\nu_1 - 1)!^m} \left(c_{k_2} |U(t)|^{k_2} \frac{(c_{k_2} t)^{\nu_2 - 1}}{(\nu_2 - 1)!} + c_{k_2}^2 \|U\|_{L^{k_2}_t}^{k_2} \frac{(c_{k_2} t)^{\nu_2 - 2}}{(\nu_2 - 2)!} \mathbf{1}_{\nu_2 \geq 2} \right). \tag{7.52}$$
Thanks to Hölders' inequality,
$$\|U\|_{L^{k_1}_t}^{mk_1} \|U\|_{L^{k_2}_t}^{k_2} \leq \|U\|_{L^k_t}^k t^m. \tag{7.53}$$
Thanks to Hölder's inequality and Young's inequality,
$$\|U\|_{L^{k_1}_t}^{mk_1} |U(t)|^{k_2} \leq t^m \left(\|U\|_{L^k_t}^k t^{-1} + |U(t)|^k \right). \tag{7.54}$$

Moreover, thanks to (3.2), for $i \in [\![1,2]\!]$,

$$\frac{1}{m!}\frac{1}{(\nu_1-1)!^m}\frac{1}{(\nu_2-i)!} \leq 2^{(m+1)(\nu-i)}\frac{1}{(\nu-i)!}. \tag{7.55}$$

Combining these inequalities proves (7.50) with $c_k := 2^{k+2}\max\{c_j; j \in [\![1,k-1]\!]\}$.

□

These enhanced estimates yield the following result.

Proposition 7.14. *Let $X = \{X_0, X_1\}$, $\mathcal{B}$ a generalized 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 $r, \delta > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,r}_{4\delta}$. 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 $\mathcal{B}\cap S_M$ converges uniformly on B_δ and, for each $t \in [0,T]$ and $p \in B_\delta$,*

$$\left| x(t;f,u,p) - \overrightarrow{\prod_{b\in\mathcal{B}\cap S_M}} e^{\xi_b(t;1,u)f_b}p \right| \leq C_M\|U\|^{M+1}_{L^{M+1}(0,t)}. \tag{7.56}$$

Proof. The proof is the same as the proof of Proposition 6.7. The only difference is that we use estimates of Lemma 7.13 instead of those of Lemma 6.6. 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 (6.34). □

7.5 Failure of the primitive estimate for multiple inputs

Proposition 7.6 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 $f_0, f_1 \in \mathcal{C}^\infty_\delta$, in the degenerate case $M = 0$ and the particular case $f_0(0) = 0$, $p = 0$, estimate (7.14) 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 $\|U\|_{L^\infty} \leq 1$,

$$|x(t;u,0)| \leq C_T\|U\|_{L^\infty}. \tag{7.57}$$

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 7.15. *Let $T > 0$ and consider the following system on $\mathbb{R}^2$:*

$$\begin{cases} \dot{x}_1 = u, \\ \dot{x}_2 = vx_1, \end{cases} \tag{7.58}$$

where u and v are two scalar inputs. There exists $u_n, v_n \in L^1(0,T)$ such that

$$\|U_n\|_{L^\infty} + \|V_n\|_{L^\infty} \to 0 \quad \text{and} \quad |x(t;(u_n,v_n),0)| \not\to 0, \tag{7.59}$$

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

$$\|U_n\|_{L^\infty} + \|V_n\|_{L^\infty} \leq \frac{2}{n}. \tag{7.60}$$

Moreover, $x_1(t) = U_n(t) = (\sin n^2 t)/n$ and

$$x_2(T) = \int_0^T v_n(t)U_n(t)\,\mathrm{d}t = \int_0^T \sin^2(n^2 t)\,\mathrm{d}t \to \frac{T}{2}, \tag{7.61}$$

as $n \to +\infty$. This proves (7.59).

Remark 7.16. *Although Proposition 7.6 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 [36].*

8 On direct intrinsic representations of the state

The expansions studied above in this paper 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 2.16 and Remark 8.7). 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 8.1. *Let $\delta > 0$ and $z \in \mathcal{C}^1_\delta$ such that $\|z\|_{\mathcal{C}^0} \leq \delta$. Then*

$$|e^z(0) - z(0)| \leq |z(0)| \|Dz\|_{\mathcal{C}^0} e^{\|Dz\|_{\mathcal{C}^0}}. \tag{8.1}$$

Proof. Let $x(t) := e^{tz}(0)$ for $t \in [0,1]$. Then, for every $t \in [0,1]$,

$$|x(t) - tz(0)| \leq \int_0^t |z(x(\tau)) - z(0)| \, \mathrm{d}\tau \leq \frac{t^2}{2} \|Dz\|_{\mathcal{C}^0} |z(0)| + \int_0^t \|Dz\|_{\mathcal{C}^0} |x(\tau) - \tau z(0)| \, \mathrm{d}\tau \tag{8.2}$$

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 8.2. *Let $M \in \mathbb{N}^*$, $\delta > 0$ and $q \in \mathbb{N}^*$.*

1. *Let $I = [\![0, q]\!]$ or $I = [\![1, q]\!]$. Let $f_i \in \mathcal{C}^{M^2}_\delta$ 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 (4.51) with $p = 0$ and $Z_M(t, f, u)$ denotes the vector field defined in Proposition 4.3 (called $Z_M(t, \sum_{i \in I} u_i f_i)$ in this statement), then, as $T \to 0$,*

$$x(t; f, u, 0) = Z_M(t, f, u)(0) + O\left(t^{M+1} + t|x(t; f, u, 0)|\right). \tag{8.3}$$

   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 $t \in [0,T]$,

$$|x(t; f, u, 0) - Z_M(t, f, u)(0)| \leq C\left(t^{M+1} + t|x(t; f, u, 0)|\right). \tag{8.4}$$

2. *Let $T > 0$, $f_0, \ldots, f_q \in \mathcal{C}^{M^2+1}_{2\delta}$ 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 (6.1) with $p = 0$ and $\mathcal{Z}(t; f, u)$ denotes the vector field defined in Proposition 6.4, then, as $\|u\|_{L^1} \to 0$,*

$$x(t; f, u, 0) = \mathcal{Z}_M(t, f, u)(0) + O\left(\|u\|^{M+1}_{L^1_t} + \|u\|_{L^1_t} |x(t; f, u, 0)|\right). \tag{8.5}$$

3. *Let $f_0, f_1 \in \mathcal{C}^\omega_{3\delta}$ with $f_0(0) = 0$. Let $T > 0$ as in Proposition 7.2. For $u \in L^1((0,T);\mathbb{K})$, if $x(t;f,u,0)$ denotes the solution to (7.1) with $p = 0$ and $\mathcal{Z}(t,f,u)$ denotes the vector field defined in Proposition 6.4 (with $q = 1$), then, as $(T, \|U\|_{L^\infty}) \to 0$,*

$$x(t;f,u,0) = \mathcal{Z}_M(t,f,u)(0) + O\left(\|U\|^{M+1}_{L^{M+1}_t} + \|U\|_{L^1_t}|x(t;f,u,0)| + |x(t;f,u,0)|^2\right) \tag{8.6}$$

where $U(s) := \int_0^s u$.

Proof. Proof of the first statement. By Proposition 4.3, 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 $t \in [0,T^*]$,

$$\left|x(t;f,u,0) - e^{Z_M(t,f,u)}(0)\right| \le C_1 t^{M+1}. \tag{8.7}$$

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 $t \in [0,T^*]$,

$$\|Z_M(t,f,u)\|_{\mathcal{C}^1} \le C_2 t. \tag{8.8}$$

Thus, by Lemma 8.1, 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 $t \in [0,T^*]$,

$$\left|e^{Z_M(t,f,u)}(0) - Z_M(t,f,u)(0)\right| \le C_3 t\,|Z_M(t,f,u)(0)|\,. \tag{8.9}$$

Then, by triangular inequality, for every $u \in L^\infty((0,T^*);\mathbb{K}^q)$ with $\|u\|_{L^\infty} \le 1$ and $t \in [0,T^*]$

$$|x(t;f,u,0) - Z_M(t,f,u)(0)| \le C_1 t^{M+1} + C_3 t |Z_M(t,f,u)(0)| \tag{8.10}$$

and in particular, for $t \le T \le 1/(2C_3)$

$$|Z_M(t,f,u)(0)| \le 2\,|x(t;f,u,0)| + 2C_1 t^{M+1}. \tag{8.11}$$

This gives (8.4) 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 6.4, then applies Lemma 8.1 to $\mathcal{Z}_M(t,f,u)$ and concludes thanks to the following estimate, implied by the explicit expressions of the vector field

$$\|\mathcal{Z}_M(t,f,u)\|_{\mathcal{C}^1} = \underset{\|u\|_{L^1}\to 0}{O}\left(\|u\|_{L^1(0,t)}\right). \tag{8.12}$$

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 7.12 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 $(t, \|U\|_{L^\infty_t}) \to 0$

$$\|\mathcal{Z}_M(t,f,u)\|_{\mathcal{C}^1} = O\left(|U(t)| + \|U\|_{L^1(0,t)}\right). \tag{8.13}$$

Thus, using $f_0(0) = 0$, Proposition 7.6 and the same strategy as above, we obtain in the asymptotics $(t, \|U\|_{L^\infty_t}) \to 0$

$$x(t;f,u,0) = \mathcal{Z}_M(t,f,u)(0) + O\left(|U(t)|^{M+1} + \int_0^t |U|^{M+1} + \left(|U(t)| + \|U\|_{L^1}\right)|x(t;f,u,0)|\right). \tag{8.14}$$

The following proposition and Hölder inequality give the conclusion. □

Proposition 8.3. *Let $\delta > 0$, $f_0, f_1 \in \mathcal{C}^\omega_\delta$ 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 $t \in [0,T]$,*

$$|U(t)| \le C\left(|x(t;f,u,0)| + \|U\|_{L^1_t}\right). \tag{8.15}$$

Proof. With the notations of Proposition 7.2, $x(t;f,u,0) = e^{U(t)f_1}x_1(t;F,U,0)$ 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 $\|U\|_{L^\infty} \leq \eta^*$,

$$|x(t;f,u,0) - x_1(t;F,U,0) - U(t)f_1(0)| \leq C_1|U(t)|^2 + C_1|U(t)||x_1(t;F,U,0)|. \tag{8.16}$$

Moreover, by Grönwall's lemma, there exists $C_2 > 0$ such that

$$|x_1(t;F,U,0)| \leq C_2\|U\|_{L^1(0,t)}. \tag{8.17}$$

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 (8.16) and using (8.17), we get the conclusion, when $\|U\|_{L^\infty}$ is small enough. □

Remark 8.4. *Estimate* (8.6) *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 (4.51), we define $y(t) := \theta(x(t))$. Then, one checks that y is the solution to

$$\dot{y}(t) = \sum_{i\in I} u_i(t)g_i(y(t)) \quad \text{and} \quad y(0) = q, \tag{8.18}$$

where $g_i := \theta_* f_i$ and $q := \theta(p)$. By iterating Lemma 3.23, Lie brackets of the vector fields defining the dynamics for y can be computed explicitly from those of x. More precisely, for every $b \in \mathrm{Br}(X)$,

$$g_b = \theta_* f_b \tag{8.19}$$

with the notation of Definition 3.13. 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 $b \in \mathrm{Br}(X)$,

$$g_b(q) = Lf_b(p). \tag{8.20}$$

Conversely, if the f_i and g_i for $i \in I$ are analytic vector fields, the existence of points p and q and a linear invertible map L_p such that (8.20) holds is a sufficient condition for the existence of a local smooth diffeomorphism θ with $\theta(p) = q$ and such that, for every controls u_i, there holds $y(t) = \theta(x(t))$ where x and y denote the solutions to (4.51) and (8.18) for the same set of controls. This nice property is proved in [49, Theorem 1] and was then extended with a more general geometric viewpoint in [66] (see also [4, Theorem 5.5] for a modern presentation).

When (8.20) only holds for brackets up to some length $M \in \mathbb{N}$ and the controls are uniformly bounded in L^∞, one can prove (see [50]) the existence of a local smooth diffeomorphism θ and a constant C such that

$$|y(t) - \theta(x(t))| \leq Ct^{M+1}. \tag{8.21}$$

Up to our knowledge, the converse, which is conjectured to be true in [50], is a nice open problem.

Open problem 8.5. *Let* $I = [\![1,q]\!]$ *and* $X = \{X_1, \dots X_q\}$*. Let* $p, q \in \mathbb{K}^d$*. Assume that there exists a smooth diffeomorphism* θ *from a neighborhood of* p *to a neighborhood of* q *and* $M \in \mathbb{N}$ *such that, for every controls* $u_1, \dots u_q \in L^\infty(0,T)$ *with* $\|u_i\| \leq 1$*, estimate* (8.21) *holds for every 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* $|b| \leq M$*,* (8.20) *holds?*

Open problem 8.6. *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* (8.21) *as it is valid for less choices of controls since* u_0 *is heavily constrained.*

Remark 8.7. *Property* (8.20) *is specific to Lie brackets and does not hold for products of differential operators. For example, consider on* $\mathbb{R}^2$ *the vector fields* $f_0(x) = (0, x_1)$ *and* $f_1(x) = (1, 0)$ *and let* $\theta(x) := (x_1, x_2 + x_1^2)$. *Then* $g_0(y) = (0, y_1)$ *and* $g_1(y) = (1, 2y_1)$. *In particular, one has* $(f_1 \cdot \nabla)f_1 = 0$ *but* $(g_1 \cdot \nabla)g_1 = (0, 2)$. *So one cannot hope for a relation such as* (8.20) *to hold for products of differential operators. 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 (4.51) with $p = 0$. Let $Z_M(t, u)$ be the vector field defined in Proposition 4.3 (and called $Z_M(t, \sum_{i\in I} u_i f_i)$ in this statement). By Proposition 4.3, 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 a 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. [5, 6]).

This also corresponds to replacing the terms $x(t; u)+o(|x(t; u)|)$ in Proposition 8.2 by $\theta(x(t; u))$, where θ is a smooth local diffeomorphism of $\mathbb{K}^d$.

We start with a definition.

Definition 8.8. *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* $t \in [0, T]$,

$$\beta(\lambda t, u^\lambda) = \lambda^n \beta(t, u) \tag{8.22}$$

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 8.9. *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* $t \in [0, T]$,

$$|\beta(t, u)| \leq Ct^{n+1}. \tag{8.23}$$

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 8.10. *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* $f_i \in \mathcal{C}^\infty_\delta$ *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 $t \in [0, T_M]$,

$$|\theta(x(t;u)) - y_M(t;u)| \leq C_M t^{M+1}, \tag{8.24}$$

and

$$y_M(t;u) = \theta(0) + \sum_{|b| \leq M} \beta_b(t,u) g_b(\theta(0)), \tag{8.25}$$

where $g_b = \theta_* f_b$ and $x(t;u)$ is the solution to (4.51) 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 (8.25) 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 8.11. *Let $X = \{X_0, X_1\}$. Let $T > 0$ and consider, in $\mathbb{R}^3$, $f_0(x) := (0, x_1 + x_1^2, x_1 x_2)$ and $f_1(x) := (1, 0, 0)$, i.e. the following affine system with drift*

$$\begin{cases} \dot{x}_1 = u, \\ \dot{x}_2 = x_1 + x_1^2, \\ \dot{x}_3 = x_1 x_2, \end{cases} \tag{8.26}$$

together with the initial data $x(0) = 0$. There exists a monomial basis $\mathcal{B}$ of $\mathcal{L}(X)$, such that, for every functionals $\beta_b : [0,T] \times L^\infty((0,T);\mathbb{R}) \to \mathbb{R}$ for $b \in \mathcal{B}$, homogenous 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 (8.24) 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

$$b_1 = X_1, \qquad b_2 = [X_0, X_1], \qquad c_1 = [X_0, [X_0, X_1]], \tag{8.27}$$

$$f_{b_1}(x) = e_1, \qquad f_{b_2}(x) = -(1 + 2x_1)e_2 - x_2 e_3, \qquad f_{c_1}(x) = x_1^2 e_3. \tag{8.28}$$

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:

$$b_3 = [X_1, [X_0, X_1]], \qquad c_2 = [X_1, \mathrm{ad}^2_{X_0}(X_1)], \tag{8.29}$$

$$f_{b_3}(x) = -2e_2, \qquad f_{c_2}(x) = 2x_1 e_3. \tag{8.30}$$

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:

$$b_4 = \mathrm{ad}^2_{X_1} \mathrm{ad}^2_{X_0}(X_1), \qquad b_5 = [[X_0, X_1], [X_1, [X_0, X_1]]], \tag{8.31}$$

$$f_{b_4}(x) = 2e_3, \qquad f_{b_5}(x) = -2e_3. \tag{8.32}$$

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

$$y_6(t) = \theta(0) + D\theta(0)\Big(\beta_1(t,u)e_1 - \beta_2(t,u)e_2 - 2\beta_3(t,u)e_2 + 2(\beta_4(t,u) - \beta_5(t,u))e_3\Big) \tag{8.33}$$

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 (8.26) yields

$$x(t;u) = U(t)e_1 + V(t)e_2 + \int_0^t U^2(s)\,\mathrm{d}s e_2 + \frac{1}{2}V^2(t)e_3 + \int_0^t U(s)\int_0^s U^2(s')\,\mathrm{d}s'\,\mathrm{d}s e_3, \tag{8.34}$$

where the five terms are respectively functionals homogeneous of degree 1 through 5 with respect to time in the sense of Definition 8.8. 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 $M \in [\![1,6]\!]$

$$\theta(x(t)) = \theta(0) + \sum_{k=1}^{M} \gamma_k(t,u) + \mathcal{O}(t^{M+1}). \tag{8.35}$$

In particular

$$\begin{aligned}\gamma_4(t,u) = {} & \frac{1}{2}V^2(t)\partial_3\theta(0) + U(t)\int_0^t U^2\partial_{12}\theta(0) + \frac{1}{2}V^2(t)\partial_{22}\theta(0) \\ & + \frac{1}{2}U^2(t)V(t)\partial_{112}\theta(0) + \frac{1}{4!}U^4(t)\partial_1^4\theta(0)\end{aligned} \tag{8.36}$$

and

$$\begin{aligned}\gamma_6(t,u) = {} & U(t)\int_0^t U(s)\int_0^s U(s')^2 ds' ds\partial_{13}\theta(0) + \frac{1}{2}V^3(t)\partial_{23}\theta(0) + \frac{1}{2}\left(\int_0^t U^2\right)^2\partial_{22}\theta(0) \\ & + \frac{1}{4}U^2(t)V^2(t)\partial_{113}\theta(0) + \frac{1}{2}U(t)V(t)\int_0^t U^2\partial_{122}\theta(0) + \frac{1}{6}V^3(t)\partial_{222}\theta(0) \\ & + \frac{1}{6}U^3(t)\int_0^t U^2\partial_{1112}\theta(0) + \frac{1}{4}U^2(t)V^2(t)\partial_{1122}\theta(0) \\ & + \frac{1}{4!}V(t)U^4(t)\partial_1^4\partial_2\theta(0) + \frac{1}{6!}U^6(t)\partial_1^6\theta(0).\end{aligned} \tag{8.37}$$

Step 3: Denying (8.24). 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 (8.24) 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 (8.24), Lemma 8.9 and (8.33) 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 8.12. *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 Joackim Bernier for discussions about the numerical literature and the importance of the convergence issues discussed in this paper for numerical splitting schemes.

Karine Beauchard and Frédéric Marbach benefit from the support of ANR project TRECOS, grant ANR-20-CE40-0009-01.

References

[1] 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.

[2] Andrei Agrachev and Revaz Gamkrelidze. Exponential representation of flows and a chronological enumeration. *Mat. Sb. (N.S.)*, 107(149)(4):467–532, 639, 1978.

[3] 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.

[4] Andrei Agrachev and Yuri Sachkov. *Control theory from the geometric viewpoint.* Berlin: Springer, 2004.

[5] Jaime Arango and Adriana Gómez. Flows and diffeomorphisms. *Rev. Colomb. Mat.*, 32(1):13–27, 1998.

[6] Jaime Arango and Adriana Gómez. Diffeomorphisms as time one maps. *Aequationes Math.*, 64(3):304–314, 2002.

[7] George Arfken, Hans Weber, and Frank Harris. *Mathematical methods for physicists. A comprehensive guide. 7th ed.* Amsterdam: Elsevier/Academic Press, 7th ed. edition, 2013.

[8] Henry Baker. Alternants and Continuous Groups. *Proc. London Math. Soc. (2)*, 3:24–47, 1905.

[9] Fabrice Baudoin. *An introduction to the geometry of stochastic flows.* Imperial College Press, London, 2004.

[10] Karine Beauchard and Frédéric Marbach. Quadratic obstructions to small-time local controllability for scalar-input systems. *J. Differential Equations*, 264(5):3704–3774, 2018.

[11] Gérard Ben Arous. Flots et séries de Taylor stochastiques. *Probab. Theory Related Fields*, 81(1):29–77, 1989.

[12] 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.

[13] Sergio Blanes and Fernando Casas. On the convergence and optimization of the Baker-Campbell-Hausdorff formula. *Linear Algebra Appl.*, 378:135–158, 2004.

[14] Sergio Blanes and Fernando Casas. *A concise introduction to geometric numerical integration.* Monographs and Research Notes in Mathematics. CRC Press, Boca Raton, FL, 2016.

[15] Sergio Blanes, Fernando Casas, and Ander Murua. Splitting and composition methods in the numerical integration of differential equations. *Bol. Soc. Esp. Mat. Apl., SēMA*, 45:89–145, 2008.

[16] Sergio Blanes, Fernando Casas, José Oteo, and Javier Ros. A pedagogical approach to the Magnus expansion. *Eur. J. Phys.*, 31(4):907–918, 2010.

[17] Sergio Blanes and Per Christian Moan. Splitting methods for non-autonomous Hamiltonian equations. *J. Comput. Phys.*, 170(1):205–230, 2001.

[18] 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.

[19] Nicolas Bourbaki. *Éléments d'histoire des mathématiques*. Masson, Paris, 1984. Reprints of the historical notes from the ıt Eléments.

[20] 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.

[21] Boris Buffoni and John Toland. *Analytic theory of global bifurcation. An introduction*. Princeton, NJ: Princeton University Press, 2003.

[22] 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).

[23] John Edward Campbell. On a Law of Combination of Operators (Second Paper). *Proc. Lond. Math. Soc.*, 29:14–32, 1897/98.

[24] 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.

[25] Fabienne Castell. Asymptotic expansion of stochastic flows. *Probab. Theory Related Fields*, 96(2):225–239, 1993.

[26] Kuo-Tsai Chen. Iterated integrals and exponential homomorphisms. *Proc. London Math. Soc. (3)*, 4:502–512, 1954.

[27] 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.

[28] Jean-Michel Coron. *Control and nonlinearity*, volume 136. Providence, RI: American Mathematical Society (AMS), 2007.

[29] Bruce Driver. On truncated logarithms of flows on a riemannian manifold. *arXiv preprint arXiv:1810.02414*, 2018.

[30] Evgenii Dynkin. Calculation of the coefficients in the Campbell-Hausdorff formula. *Doklady Akad. Nauk SSSR (N.S.)*, 57:323–326, 1947.

[31] 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.

[32] David Finkelstein. On relations between commutators. *Comm. Pure Appl. Math.*, 8:245–250, 1955.

[33] Michel Fliess. Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bull. Soc. Math. France*, 109(1):3–40, 1981.

[34] 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.

[35] 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.

[36] 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.

[37] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2010.

[38] Marshall Hall. A basis for free Lie rings and higher commutators in free groups. *Proc. Am. Math. Soc.*, 1:575–581, 1950.

[39] Philip Hall. A contribution to the theory of groups of prime-power order. *Proc. Lond. Math. Soc. (2)*, 36:29–95, 1933.

[40] Felix Hausdorff. Die symbolische exponentialformel in der gruppentheorie. *Ber. Verh. Kgl. Schs. Ges. Wiss. Leipzig., Math.-phys. Kl.*, 58:19–48, 1906.

[41] 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.

[42] Nathan Jacobson. *Lie algebras*. Dover Publications, Inc., New York, 1979. Republication of the 1962 original.

[43] Frédéric Jean. *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*. Cham: Springer; Bilbao: BCAM – Basque Center for Applied Mathematics, 2014.

[44] Velimir Jurdjevic and Héctor Sussmann. Control systems on Lie groups. *J. Differential Equations*, 12:313–329, 1972.

[45] Matthias Kawski. Controllability via chronological calculus. In *Proceedings of the IEEE Conference on Decision and Control*, volume 3, pages 2920–2925. IEEE, 1999.

[46] Matthias Kawski. Calculating the logarithm of the chen fliess series. In *Proc. MTNS, Perpignan*, 2000.

[47] 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.

[48] Matthias Kawski and Héctor Sussmann. Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory. In *Operators, systems, and linear algebra (Kaiserslautern, 1997)*, European Consort. Math. Indust., pages 111–128. Teubner, Stuttgart, 1997.

[49] Arthur Krener. On the equivalence of control systems and the linearization of nonlinear systems. *SIAM J. Control*, 11:670–676, 1973.

[50] Arthur Krener. Local approximation of control systems. *J. Differ. Equations*, 19:125–133, 1975.

[51] John Lee. *Introduction to smooth manifolds*, volume 218. New York, NY: Springer, 2nd revised edition, 2013.

[52] Roger Lyndon. A theorem of Friedrichs. *Michigan Math. J.*, 3:27–29, 1955.

[53] Wilhelm Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure Appl. Math.*, 7:649–673, 1954.

[54] Robert McLachlan. Composition methods in the presence of small parameters. *BIT*, 35(2):258–268, 1995.

[55] Per Christian Moan and Jitse Niesen. Convergence of the Magnus series. *Found. Comput. Math.*, 8(3):291–301, 2008.

[56] Clément Mouhot and Cédric Villani. On Landau damping. *Acta Math.*, 207(1):29–201, 2011.

[57] Ander Murua. The Hopf algebra of rooted trees, free Lie algebras, and Lie series. *Found. Comput. Math.*, 6(4):387–426, 2006.

[58] 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.

[59] Rimhak Ree. Lie elements and an algebra associated with shuffles. *Ann. of Math. (2)*, 68:210–220, 1958.

[60] Christophe Reutenauer. *Free Lie algebras*, volume 7 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1993.

[61] 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.

[62] Leonardo Saenz and Rodolfo Suarez. A combinatorial approach to the generalized Baker-Campbell-Hausdorff-Dynkin formula. *Systems Control Lett.*, 45(5):357–370, 2002.

[63] Wilhelm Specht. Die linearen Beziehungen zwischen höheren Kommutatoren. *Math. Z.*, 51:367–376, 1948.

[64] 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.

[65] Robert Strichartz. The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations. *J. Funct. Anal.*, 72(2):320–345, 1987.

[66] Héctor Sussmann. An extension of a theorem of Nagano on transitive Lie algebras. *Proc. Am. Math. Soc.*, 45:349–356, 1974.

[67] Héctor Sussmann. Lie brackets and local controllability: a sufficient condition for scalar-input systems. *SIAM J. Control Optim.*, 21(5):686–713, 1983.

[68] 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.

[69] Hale Trotter. On the product of semi-groups of operators. *Proc. Am. Math. Soc.*, 10:545–551, 1959.

[70] 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.

[71] James Wei. Note on the global validity of the Baker-Hausdorff and Magnus theorems. *J. Mathematical Phys.*, 4:1337–1341, 1963.

[72] Franz Wever. Operatoren in Lieschen Ringen. *J. Reine Angew. Math.*, 187:44–55, 1949.