Contents

Abstract

The paper contains a consistent presentation of the approach developed by the authors to analysis of nonlinear control systems, which exploits ideas and techniques of formal power series of independent noncommuting variables and the corresponding free algebras. The main part of the paper was conceived with a view of comparing our results with the results obtained by use of the differential-geometric approach. We consider control-linear systems with m controls. In a free associative algebra with m generators (which can be thought of as a free algebra of iterated integrals), a control system uniquely defines two special objects: the core Lie subalgebra and the graded left ideal. It turns out that each of these two objects completely defines a homogeneous approximation of the system. Our approach allows us to propose an algebraic (coordinate-independent) definition of the homogeneous approximation. This definition provides the uniqueness of the homogeneous approximation (up to a change of coordinates) and gives a way to find it directly, without preliminary finding privileged coordinates. The presented technique yields an effective description of all privileged coordinates and an explicit way of constructing an approximating system. In addition, we discuss the connection between the homogeneous approximation and an approximation in the sense of time optimality.

Acknowledgements. We would like to express our sincerest gratitude to Andrei Agrachev and Bronisław Jakubczyk for their sustained attention and encouragement of our research on algebraic methods in nonlinear control theory. We are deeply grateful to Piotr Mormul and Igor Zelenko for numerous substantial and helpful discussions on different questions concerning nonlinear control and homogeneous approximation.

The paper was partially written when the second author visited the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy; she is grateful to the Abdus Salam ICTP for the support and hospitality.

The authors sincerely thank the anonymous referee for his/her thorough analysis of the manuscript and valuable and friendly comments.

This work was partially supported by the Polish National Center of Science grant N N514 238438.

2010 *Mathematics Subject Classification*: Primary 93B25; Secondary 93B11, 93B15, 93B17.

Key words and phrases: nonlinear control system, formal power series, free algebra, core Lie subalgebra, left ideal generated by a system, R. Ree's theorem, homogeneous approximation, privileged coordinates, time optimality.

Received 28.9.2012; revised version 12.12.2013.

1. Introduction

In this paper we give an analysis of small-time approximation for control-linear systems by use of the approach based on formal power series of independent noncommuting variables and the corresponding free algebras. In particular, we propose an algebraic interpretation of concepts related to homogeneous approximation that are traditionally treated within differential-geometric methods. The free algebras approach to these problems described here was developed by the authors of the present paper during the last fifteen years.

More specifically, we consider the Cauchy problem for control-linear systems of the form

$$\dot{x} = \sum_{i=1}^{m} u_i X_i(x), \qquad x(0) = 0, \tag{1.1}$$

where $X_1(x), \dots, X_m(x)$ are real analytic vector fields. Our goal is to analyze, from an algebraic viewpoint, the concept of homogeneous approximation that was one of the points of interest in control theory during several decades [12, 23, 24, 6, 3, 8, 4]. The traditional approach is based on differential-geometric methods; a fundamental presentation can be found in [6].

Our interest in this field is connected with the study of time-optimal control problems. However, for time optimality it is more natural to consider control-affine systems instead of control-linear ones, and the end condition $x(\theta) = 0$ instead of the initial condition $x(0) = 0$. In essence, our main results concerning the application of formal power series and free algebras [49]–[55] are obtained just for such systems. In Subsection 1.1 we give a brief description of the main ideas. This subsection is independent of the rest of the paper; however, in the rest of the paper we mainly develop the approach described there. A sketch of the main results of the paper can be found in Subsection 1.2.

1.1. Series of nonlinear power moments and a nonlinear Markov moment problem. In [49] we proposed to apply the series method to a time-optimal control problem for nonlinear control-affine systems. As a first step, we suggested considering the time-optimal control problem as a nonlinear Markov moment problem.

This idea is well known in linear time optimality [37]–[39]. Consider a linear time-optimal control problem of the form

$$\dot{x} = Ax + bu, \qquad x(0) = x^0, \; x(\theta) = 0, \; |u(t)| \le 1, \qquad \theta \to \min, \tag{1.2}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}$, and A and b are a matrix and a vector of appropriate dimensions. Forget for a moment about the control constraints and optimal requirements, and consider the steering problem

$$\dot{x} = Ax + bu, \qquad x(0) = x^0,\ x(\theta) = 0, \tag{1.3}$$

i.e., the problem of finding a control $u(t)$, $t \in [0, \theta]$, that steers the given point x^0 to the origin in the given time θ. Due to the Cauchy formula, all such controls are described by the (vector) equality

$$x^0 = -\int_0^\theta e^{-tA} bu(t)\, dt.$$

Denoting $g(t) = -e^{-tA}b$, we see that *the steering problem* (1.3) *is equivalent to the Markov moment problem* [43, 5, 39]

$$x_i^0 = \int_0^\theta g_i(t)u(t)\, dt, \qquad i = 1, \ldots, n. \tag{1.4}$$

Since $g(t) = -e^{-tA}b = \sum_{k=0}^\infty \frac{(-1)^{k+1}}{k!} t^k A^k b$, equalities (1.4) can be rewritten as

$$x^0 = \sum_{k=0}^\infty v_k \int_0^\theta t^k u(t)\, dt. \tag{1.5}$$

Thus, the right hand side of (1.5) is a series of *power moments*

$$\xi_k(\theta, u) = \int_0^\theta t^k u(t)\, dt \tag{1.6}$$

of the function $u(t)$ with constant vector coefficients $v_k \in \mathbb{R}^n$. These coefficients can be found by the formula $v_k = \frac{(-1)^{k+1}}{k!} A^k b$, $k \geq 0$.

The steering problem (1.3) defines a (linear) operator $S(\theta, \cdot)$ that takes a control $u(t)$ to the corresponding initial point x^0, i.e., $S(\theta, u) = x^0$. Therefore, the right hand side of (1.5) gives a series expansion for this operator,

$$S(\theta, u) = \sum_{k=0}^\infty v_k \xi_k(\theta, u). \tag{1.7}$$

Let us apply a linear change of variables $y = Qx$ in the initial system. Obviously, it leads to the linear transformation of the corresponding series of power moments; namely, the series with coefficients v_k is mapped to the series with coefficients $\widehat{v}_k = Qv_k$. Hence, the new coefficients can be found directly from the old ones, *without finding the form of the system in the new variables.*

Now suppose θ is sufficiently small. The power moments have the following homogeneity property:

$$\xi_k(\theta, u) = \theta^{k+1} \xi_k(1, \tilde{u}), \quad \text{where} \quad \tilde{u}(t) = u(\theta t),\ t \in [0, 1].$$

This means that locally, for a small time θ, the order of smallness of the power moment ξ_k equals $k+1$. In particular, this order allows comparing terms of the series on the right hand side of (1.7).

This observation suggests the following idea: a small-time approximation of the control system can be described in terms of the series representation (1.7), as is common in calculus, when using Taylor series to approximate finite-dimensional mappings. Suppose the initial system is controllable. Then the vectors $v_0, \ldots, v_{n-1}$ are linearly independent.

Denote $Q = (v_0, \dots, v_{n-1})^{-1}$. Then the change of variables $y = Qx$ reduces the series in (1.7) to the form $\widehat{S}(\theta, u) = QS(\theta, u)$ whose componentwise representation is

$$(\widehat{S}(\theta, u))_i = \xi_{i-1}(\theta, u) + \rho_i(\theta, u), \qquad i = 1, \dots, n,$$

where $\rho_i(\theta, u) = \sum_{k=i}^{\infty}(Qv_k)_i \xi_k(\theta, u)$ contains terms of order greater than the order of ξ_{i-1}. (The order of terms of $\rho_i(\theta, u)$ can be made greater than the order of ξ_{n-1}, however this is not of importance further.) Thus, one may take the "series"

$$S_i^A = \xi_{i-1}(\theta, u), \qquad i = 1, \dots, n,$$

as a small-time approximation of the initial series S. Notice that S^A corresponds to the chained system

$$\dot{x}_1 = u, \quad \dot{x}_i = x_{i-1}, \ i = 2, \dots, n.$$

Hence, all controllable linear autonomous systems with a one-dimensional control are approximated by the chained system (up to a change of variables) in the sense mentioned above.

Now let us return to time optimality and, along with (1.2), consider the time-optimal control problem

$$\dot{x}_1 = u, \qquad \dot{x}_i = x_{i-1}, \ i = 2, \dots, n, \qquad x(0) = x^0, \ x(\theta) = 0, \ |u(t)| \le 1, \quad \theta \to \min. \quad (1.8)$$

It can be shown [38, 48] that solutions of these problems are equivalent in the following sense:

$$\theta_{x^0}/\theta^A_{Qx^0} \to 1, \qquad \frac{1}{\theta}\int_0^{\theta} |u_{x^0}(t) - u^A_{Qx^0}(t)|\, dt \to 0 \quad \text{ as } x^0 \to 0, \qquad (1.9)$$

where $(\theta_{x^0}, u_{x^0}(t))$ is a solution of (1.2), $(\theta^A_{x^0}, u^A_{x^0}(t))$ is a solution of (1.8), and $\theta = \min\{\theta_{x^0}, \theta^A_{Qx^0}\}$.

The class of nonautonomous linear control systems gives a variety of possible approximations. Namely, consider a steering problem of the form

$$\dot{x} = A(t)x + b(t)u, \qquad x(0) = x^0, \ x(\theta) = 0, \qquad (1.10)$$

where $A(t)$ and $b(t)$ are a matrix and a vector of appropriate dimensions with real analytic entries. This problem can also be rewritten in the form (1.7), where constant vector coefficients can be found from the formula $v_k = \frac{1}{k!}(-A(t) + d/dt)^k b(t)|_{t=0}$, $k \ge 0$.

Conversely, any set of vector coefficients v_k satisfying a natural convergence requirement $\|v_k\| \le k! C_1 C_2^k$, $C_1, C_2 > 0$, defines the series (1.7) corresponding to a system of the form (1.10). However, the system is not defined uniquely; for example, one can choose $A(t) = 0$ and $b(t) = \sum_{k=0}^{\infty} v_k t^k$.

Suppose the system is controllable. Then $\operatorname{rank}\{v_k\}_{k=0}^{\infty} = n$. Let $v_{m_1}, \dots, v_{m_n}$ be the first n linearly independent vectors from the sequence $\{v_k\}_{k=0}^{\infty}$, and $Q = (v_{m_1}, \dots, v_{m_n})^{-1}$. Then the change of variables $y = Qx$ reduces the series for (1.10) to the form $\widehat{S}(\theta, u) = QS(\theta, u)$,

$$(\widehat{S}(\theta, u))_i = \xi_{m_i}(\theta, u) + \rho_i(\theta, u), \qquad i = 1, \dots, n,$$

where $\rho_i(\theta, u) = \sum_{k=m_i+1}^{\infty}(Qv_k)_i \xi_k(\theta, u)$ contains terms of order (of smallness) greater than the order of ξ_{m_i}. Notice that the order of terms of $\rho_i(\theta, u)$, in general, may not be greater than the order of ξ_{m_n}. As an example, consider the system

$$\dot{x}_1 = u + tu, \qquad \dot{x}_2 = t^2 u.$$

Then $m_1 = 0$, $m_2 = 2$, and

$$(\widehat{S}(\theta, u))_1 = \xi_0(\theta, u) + \rho_1(\theta, u), \qquad (\widehat{S}(\theta, u))_2 = \xi_2(\theta, u),$$

where $\rho_1(\theta, u) = \xi_1(\theta, u)$. Here the order of $\rho_1(\theta, u)$ is greater than the order of $\xi_0(\theta, u)$, but less than the order of $\xi_2(\theta, u)$.

Hence, the series

$$(S^A(\theta, u))_i = \xi_{m_i}(\theta, u), \qquad i = 1, \dots, n,$$

can be considered as an approximation of the initial series S. A system corresponding to S^A is not defined uniquely; for example, it can be taken in the form

$$\dot{x}_i = -t^{m_i} u, \qquad i = 1, \dots, n. \tag{1.11}$$

This means that all controllable linear (real analytic) systems with a one-dimensional control are approximated by systems of the form (1.11). As above, this approximation implies the approximation in the sense of time-optimality [38, 48], i.e., (1.9) holds with $(\theta^A_{x^0}, u^A_{x^0}(t))$ a solution of the time-optimal control problem

$$\dot{x}_i = -t^{m_i} u, \quad i = 1, \dots, n, \qquad x(0) = x^0,\ x(\theta) = 0,\ |u(t)| \le 1, \qquad \theta \to \min.$$

Thus, the main idea of the previous analysis is as follows: replace a control system by a series of power moments, and approximate this series, taking into account the order of smallness of power moments.

Let us now go over to a nonlinear case. Consider the class of control-affine systems of the form

$$\dot{x} = a(t, x) + b(t, x)u, \tag{1.12}$$

where $a(t, x)$ and $b(t, x)$ are real analytic vector functions in a neighborhood of the origin. Suppose the origin is an equilibrium, which means $a(t, 0) \equiv 0$. As before, consider the steering problem to the origin, i.e.,

$$\dot{x} = a(t, x) + b(t, x)u, \qquad a(t, 0) \equiv 0, \qquad x(0) = x^0,\ x(\theta) = 0. \tag{1.13}$$

The first step is to find an appropriate series representation for this problem. As before, consider the operator $S(\theta, \cdot)$ that takes a control $u(t)$ to the corresponding initial point x^0, i.e., $S(\theta, u) = x^0$. More specifically, let us fix $\theta > 0$ and $u = u(t)$, $t \in [0, \theta]$. Substitute the control $u = u(t)$ into system (1.12) and invert the time $\tau = \theta - t$. Consider the Cauchy problem

$$\frac{d\widetilde{x}}{d\tau} = -a(\theta - \tau, \widetilde{x}) - b(\theta - \tau, \widetilde{x})u(\theta - \tau), \qquad \widetilde{x}(0) = 0,$$

and set $S(\theta, u) = \widetilde{x}(\theta)$. Obviously, if $x^0 = S(\theta, u) = \widetilde{x}(\theta)$ then $x(t) = x(\theta - \tau) = \widetilde{x}(\tau)$ satisfies (1.13) with $u = u(t)$. This means that x^0 is taken to the origin in time θ by the control $u(t)$ with respect to system (1.12).

The operator $S(\theta,\cdot)$ admits the series expansion [49, 51]

$$S(\theta,u)=\sum_{k=1}^{\infty}\sum_{m_1,\dots,m_k\ge 0} v_{m_1\dots m_k}\xi_{m_1\dots m_k}(\theta,u), \tag{1.14}$$

where $\xi_{m_1\dots m_k}(\theta,u)$ are *nonlinear power moments* of the form

$$\xi_{m_1\dots m_k}(\theta,u)=\int_0^{\theta}\int_0^{\tau_1}\cdots\int_0^{\tau_{k-1}}\tau_1^{m_1}\tau_2^{m_2}\cdots\tau_k^{m_k}u(\tau_1)u(\tau_2)\cdots u(\tau_k)\,d\tau_k\cdots d\tau_2\,d\tau_1, \tag{1.15}$$

and $v_{m_1\dots m_k}\in\mathbb{R}^n$ are constant vectors that can be found from $a(t,x)$ and $b(t,x)$ by certain formulas. (A similar expansion was used in [8] for the approximation along a trajectory.)

We are going to consider the series of nonlinear power moments (1.14) instead of the initial control system. Suppose a (real analytic) change of variables $y=Q(x)$ is applied in the system, where $Q(0)=0$. Let us find the series representation of the system in the new coordinates.

As for the linear case, we *do not use the form of the system in the new coordinates*; instead, we consider the transformation of the series itself. For brevity, let us write the Taylor series expansion for $Q(x)$ as $Q(x)=\sum_{q=1}^{\infty}\frac{1}{q!}Q^{(q)}(0)x^q$. Then

$$\widehat{S}(\theta,u)=Q(S(\theta,u))=\sum_{q=1}^{\infty}\frac{1}{q!}Q^{(q)}(0)(S(\theta,u))^q. \tag{1.16}$$

Therefore, we encounter the problem of finding powers of the series, i.e., products of nonlinear power moments.

Returning to the change of variables in the system, let us consider a product of two power moments (1.15). For example,

$$\begin{aligned}
\xi_{m_1}(\theta,u)\xi_{m_2}(\theta,u) &= \int_0^{\theta}\tau_1^{m_1}u(\tau_1)\,d\tau_1\int_0^{\theta}\tau_2^{m_2}u(\tau_2)\,d\tau_2\\
&=\int_0^{\theta}\int_0^{\tau_1}\tau_1^{m_1}\tau_2^{m_2}u(\tau_1)u(\tau_2)\,d\tau_2\,d\tau_1\\
&\quad+\int_0^{\theta}\int_0^{\tau_2}\tau_1^{m_1}\tau_2^{m_2}u(\tau_1)u(\tau_2)\,d\tau_1\,d\tau_2\\
&=\xi_{m_1m_2}(\theta,u)+\xi_{m_2m_1}(\theta,u);\\
\xi_{m_1}(\theta,u)\xi_{m_2m_3}(\theta,u) &= \int_0^{\theta}\tau_1^{m_1}u(\tau_1)\,d\tau_1\int_0^{\theta}\int_0^{\tau_2}\tau_2^{m_2}\tau_3^{m_3}u(\tau_2)u(\tau_3)\,d\tau_2\,d\tau_3\\
&=\xi_{m_1m_2m_3}(\theta,u)+\xi_{m_2m_1m_3}(\theta,u)+\xi_{m_2m_3m_1}(\theta,u),
\end{aligned}$$

and so on.

These relations can be described in the following terms. Instead of the linear space of linear power moments (1.6), in the nonlinear case we introduce the *algebra* of nonlinear power moments (1.15). Namely, consider the moments $\xi_{m_1\dots m_k}(\theta,u)$ as *words* generated by the *letters* $\xi_i(\theta,u)$, i.e., assume that the word $\xi_{m_1\dots m_k}(\theta,u)$ is a *concatenation* of the letters $\xi_{m_1}(\theta,u),\dots,\xi_{m_k}(\theta,u)$. Then the linear space of nonlinear power moments becomes an associative noncommutative algebra. It can be shown that nonlinear moments are linearly independent as functionals on u, therefore the above-mentioned algebra is free. Hence, this algebra is isomorphic to a (free) algebra of formal polynomials (with coefficients in $\mathbb{R}$) of noncommuting independent *abstract* variables $\{\xi_i,\, i \geq 0\}$. That is, monomials are of the form $\xi_{m_1\ldots m_k} = \xi_{m_1}\cdots\xi_{m_k}$. We denote this algebra by $\mathcal{A}$ and call it "the algebra of nonlinear power moments".

The series on the right hand side of (1.14) can therefore be described by the linear map $v : \mathcal{A} \to \mathbb{R}^n$ defined by $v(\xi_{m_1\ldots m_k}) = v_{m_1\ldots m_k}$. Moreover, this series has its formal analogue, namely the formal power series of ξ_i with coefficients in $\mathbb{R}^n$, i.e.,

$$S = \sum_{k=1}^{\infty} \sum_{m_1,\ldots,m_k \geq 0} v_{m_1\ldots m_k}\xi_{m_1\ldots m_k}.$$

Then the above-mentioned "usual" product of nonlinear power moments corresponds to the *shuffle product operation* in $\mathcal{A}$ [14, 46, 10, 2]; it is defined recurrently as

$$\begin{aligned}
\xi_{m_1} \sqcup\!\sqcup\, \xi_{q_1} &= \xi_{m_1 q_1} + \xi_{q_1 m_1}, \\
\xi_{m_1} \sqcup\!\sqcup\, \xi_{q_1\ldots q_r} &= \xi_{q_1\ldots q_r} \sqcup\!\sqcup\, \xi_{m_1} = \xi_{m_1 q_1\ldots q_r} + \xi_{q_1}(\xi_{m_1} \sqcup\!\sqcup\, \xi_{q_2\ldots q_r}), \qquad && r \geq 2, \\
\xi_{m_1\ldots m_k} \sqcup\!\sqcup\, \xi_{q_1\ldots q_r} &= \xi_{m_1}(\xi_{m_2\ldots m_k} \sqcup\!\sqcup\, \xi_{q_1\ldots q_r}) + \xi_{q_1}(\xi_{m_1\ldots m_k} \sqcup\!\sqcup\, \xi_{q_2\ldots q_r}), \qquad && k, r \geq 2.
\end{aligned}$$

As a result, the nonlinear power moments series for (1.16) can actually be found directly from the series (1.14). Recall that this allows us to find the series representation of the system after a change of variables directly via the initial series, without finding the form of the system in the new variables. Therefore, manipulations over the system can be reduced to purely algebraic procedures.

A number of questions concerning control-affine systems can be analyzed within the well developed "combinatorics on words" [42, 47, 32, 33]. As an example, let us mention a realizability problem. Namely, in contrast to the linear case, a set of vector coefficients $v_{m_1\ldots m_k}$ defining a series of a system of the form (1.13) cannot be arbitrary. Let us give an algebraic description of realizability conditions.

Consider the Lie algebra $\mathcal{L}$ freely generated by the same elements $\{\xi_i,\, i \geq 0\}$, with the Lie bracket

$$[\ell_1, \ell_2] = \ell_1\ell_2 - \ell_2\ell_1.$$

In these terms, the realizability theorem takes the following form [50]. Suppose a linear map $v : \mathcal{A} \to \mathbb{R}^n$ satisfies the condition

$$v(\mathcal{L}) = \mathbb{R}^n. \tag{1.17}$$

Recall that this is an accessibility condition, i.e., it guarantees that the set of those x^0 for which the steering problem (1.13) is solvable has a nonempty interior, and the origin belongs to the closure of this interior. Then the series (1.14) corresponds to a system of the form (1.13) if and only if

$$\|v_{m_1\ldots m_k}\| \leq k! C_1 C_2^{m_1+\cdots+m_k+k}, \qquad C_1, C_2 > 0,$$

and the following condition holds

$$\text{if} \quad v(\ell) = 0 \text{ for } \ell \in \mathcal{L} \quad \text{then} \quad v(\ell z) = 0 \text{ for any } z \in \mathcal{A}. \tag{1.18}$$

In other words, (1.18) means that the right ideal generated by $\mathrm{Ker}(v)\cap\mathcal{L}$ is contained in $\mathrm{Ker}(v)$.

Now let us pass to the approximation problem. Due to the homogeneity property

$$\xi_{m_1\dots m_k}(\theta,u)=\theta^{m_1+\cdots+m_k+k}\xi_{m_1\dots m_k}(1,\tilde u),\quad \text{where}\quad \tilde u(t)=u(\theta t),\ t\in[0,1],$$

it is natural to introduce the definition of the order of smallness as

$$\mathrm{ord}(\xi_{m_1\dots m_k})=m_1+\cdots+m_k+k.$$

This order generates the natural grading in $\mathcal{A}$, defined as

$$\mathcal{A}=\bigoplus_{m=1}^{\infty}\mathcal{A}^m,\quad \text{where}\quad \mathcal{A}^m=\mathrm{Lin}\{\xi_{m_1\dots m_k}:m_1+\cdots+m_k+k=m\}.$$

Consider a system of the form (1.13), and its series (1.14) (or, what is the same, the map $v:\mathcal{A}\to\mathbb{R}^n$); suppose (1.17) holds. Set $\mathcal{L}^m=\mathcal{L}\cap\mathcal{A}^m$, and denote

$$\mathcal{P}^m=\{\ell\in\mathcal{L}^m: v(\ell)\in v(\mathcal{L}^1\oplus\cdots\oplus\mathcal{L}^{m-1})\},\qquad m\geq 1.$$

For convenience, denote $\mathcal{A}^e=\mathcal{A}\oplus\mathbb{R}$, assuming $1\cdot a=a\cdot 1=a$ for any $a\in\mathcal{A}^e$. Introduce the right ideal generated by the sets $\mathcal{P}^m$, i.e.,

$$\mathcal{J}=\mathrm{Lin}\Big\{\ell z:\ell\in\bigoplus_{m=1}^{\infty}\mathcal{P}^m,\ z\in\mathcal{A}^e\Big\}.$$

Due to (1.17), the set $\bigoplus_{m=1}^{\infty}\mathcal{P}^m$ is of codimension n in $\mathcal{L}$. Choose any $\ell_1,\dots,\ell_n\in\mathcal{L}$ such that

$$\mathcal{L}=\mathrm{Lin}\{\ell_1,\dots,\ell_n\}+\bigoplus_{m=1}^{\infty}\mathcal{P}^m;$$

without loss of generality suppose $\ell_i\in\mathcal{L}^{w_i}$, $i=1,\dots,n$, and $w_1\leq\cdots\leq w_n$. Finally, introduce the inner product in $\mathcal{A}$ assuming that $\{\xi_{m_1\dots m_k}\}$ form an orthonormal basis.

The main "approximation theorem" can be formulated as follows [51]. *There exists a (real analytic) change of variables $y=Q(x)$ such that in the new variables the series of the system $\widehat S(\theta,u)=Q(S(\theta,u))$ is of the form*

$$(\widehat S(\theta,u))_i=\widetilde\ell_i(\theta,u)+\rho_i(\theta,u),\qquad i=1,\dots,n,$$

where $\widetilde\ell_i$ denotes the orthogonal projection of ℓ_i on the subspace $\mathcal{J}^\perp$, and ρ_i contains terms of order greater than the order of ℓ_i, i.e., $\rho_i\in\bigoplus_{m=w_i+1}^{\infty}\mathcal{A}^m$. Hence, the series S^A of the form

$$(S^A(\theta,u))_i=\widetilde\ell_i(\theta,u),\qquad i=1,\dots,n,$$

can be considered as an approximation of the initial series S. Moreover, the series S^A is realizable, i.e., it corresponds to some system of the form (1.13); *this system can be considered as an approximation of the initial control system.*

It can be shown that the linear subspace $\bigoplus_{m=1}^{\infty}\mathcal{P}^m$ is a Lie subalgebra of $\mathcal{L}$ and, moreover, can be an arbitrary Lie subalgebra of codimension n. Hence, the cited result gives a complete description of all possible approximations of systems (1.13).

Notice also that, under some additional conditions, this approximation implies the approximation in the sense of time optimality [51].

1.2. Sketch of the main results. With reference to our approach and results mentioned in the previous subsection, the question arose about a connection of our approximation and the concept of a homogeneous approximation [12, 23, 24, 6, 3, 8, 4]. This list of references is far from complete; during the last three decades several different approaches to the above-mentioned problem were proposed and developed. The present paper is conceived as an attempt to give an algebraic interpretation of the problem of homogeneous approximation, and to clarify the relationship between the algebraic and differential-geometric approaches. When comparing two approaches, it is natural to apply them to the same object. So, here we consider the Cauchy problem for control-linear systems of the form (1.1), as it was done in [6]. In this case we also deal with a free Lie algebra and a free associative algebra; however, unlike the algebras considered in Subsection 1.1, they are generated by a finite number of generating elements.

Namely, along with a control-linear system of the form (1.1), we consider its *endpoint map*, i.e., the operator $\mathcal{E}_{X_1,\ldots,X_m}(\theta,\cdot)$ taking a control $u=u(t)$ to the end point of the trajectory of (1.1), so that $\mathcal{E}_{X_1,\ldots,X_m}(\theta,u)=x(\theta)$ (Subsection 2.1). By a homogeneous approximation of system (1.1) we mean a system of the same form whose endpoint map is homogeneous and approximates the endpoint map of the initial system as $\theta\to 0$; for the precise definition see Subsection 3.1 (Definition 3.1). A substantial part of the results of the present paper is connected with the algebraic description of this concept.

In Subsection 2.2 we discuss the series representation of the endpoint map. Series of iterated integrals first proposed in [10] were adopted to the control theory context in [17, 19, 20]. We start with the following representation of the map $\mathcal{E}_{X_1,\ldots,X_m}(\theta,u)$, which can be considered as a partial case of the result of M. Fliess [19],

$$\mathcal{E}_{X_1,\ldots,X_m}(\theta,u)=\sum_{k=1}^{\infty}\sum_{1\le i_1,\ldots,i_k\le m}c_{i_1\ldots i_k}\eta_{i_1\ldots i_k}(\theta,u),$$

where

$$\eta_{i_1\ldots i_k}(\theta,u)=\int_0^{\theta}\int_0^{\tau_1}\cdots\int_0^{\tau_{k-1}}u_{i_1}(\tau_1)u_{i_2}(\tau_2)\cdots u_{i_k}(\tau_k)\,d\tau_k\cdots d\tau_2\,d\tau_1$$

are "iterated integrals" and $c_{i_1\ldots i_k}$ are constant vector coefficients. In Subsection 2.3 we study iterated integrals and, in particular, show that for any $\theta>0$ the linear span of iterated integrals forms a free algebra of functionals defined on the unit ball of $L_\infty([0,\theta];\mathbb{R}^m)$. This observation motivates introducing an abstract free associative graded algebra $\mathcal{F}$ (over $\mathbb{R}$) generated by m elements (letters) $\eta_1,\ldots,\eta_m$ as the algebra of words $\eta_{i_1\ldots i_k}=\eta_{i_1}\cdots\eta_{i_k}$ with the natural gradation $\mathcal{F}=\bigoplus_{k=1}^{\infty}\mathcal{F}^k$, where

$$\mathcal{F}^k=\mathrm{Lin}\{\eta_{i_1\ldots i_k}:1\le i_1,\ldots,i_k\le m\},\qquad k\ge 1.$$

By attaching the unity element 1 (the empty word), we get the algebra $\mathcal{F}^e=\mathcal{F}+\mathbb{R}$. Then the series for $\mathcal{E}_{X_1,\ldots,X_m}(\theta,u)$ has its formal analogue, namely, the *formal power series of independent noncommutating variables* $\eta_1,\ldots,\eta_m$ *with coefficients in* $\mathbb{R}^n$.

We also introduce the graded Lie algebra $\mathcal{L}=\bigoplus_{k=1}^{\infty}\mathcal{L}^k$ generated by the same m elements $\eta_1,\ldots,\eta_m$ with Lie bracket $[\ell_1,\ell_2]=\ell_1\ell_2-\ell_2\ell_1$. Since we are going to consider series instead of systems, we describe transformations over series that correspond to changes of variables. In particular, this justifies the consideration of the shuffle product operation $\sqcup\!\sqcup$ in $\mathcal{F}^e$ (Subsection 2.4), defined as $1 \sqcup\!\sqcup a = a \sqcup\!\sqcup 1 = a$ for any $a \in \mathcal{F}^e$, and recursively,

$$\eta_{i_1\ldots i_k} \sqcup\!\sqcup \eta_{j_1\ldots j_r} = (\eta_{i_1\ldots i_{k-1}} \sqcup\!\sqcup \eta_{j_1\ldots j_r})\eta_{i_k} + (\eta_{i_1\ldots i_k} \sqcup\!\sqcup \eta_{j_1\ldots j_{r-1}})\eta_{j_r}, \qquad k, r \geq 1.$$

This operation corresponds to the "usual product" of iterated integrals as functionals. Discussions on properties of iterated integrals, the shuffle product, and their usage for control systems can be found, for example, in [19], [2], [32], [21].

A concrete system of the form (1.1) is characterized by the linear map $c : \mathcal{F} \to \mathbb{R}^n$ defined as $c(\eta_{i_1\ldots i_k}) = c_{i_1\ldots i_k}$; it, in turn, defines the *core Lie subalgebra* $\mathcal{L}_{X_1,\ldots,X_m} \subset \mathcal{L}$ (Subsection 2.6) in the following way:

$$\mathcal{L}_{X_1,\ldots,X_m} = \bigoplus_{k=1}^{\infty} \mathcal{P}^k,$$

where $\mathcal{P}^k = \{\ell \in \mathcal{L}^k : c(\ell) \in c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k-1})\}$, $k \geq 1$.

The main achievements of the paper are based on the following observation: *The core Lie subalgebra* $\mathcal{L}_{X_1,\ldots,X_m}$ *(or, equivalently, the left ideal* $\mathcal{J}_{X_1,\ldots,X_m} = \mathrm{Lin}\{\mathcal{F}^e\mathcal{L}_{X_1,\ldots,X_m}\}$*, see Subsection* 4.1*) contains all the information about a homogeneous approximation of the system.* For some additional discussion on core Lie subalgebras, see [25], [26].

In Section 3 we turn to studying homogeneous approximation. We restrict ourselves to considering bracket generating systems, i.e., satisfying the Rashevsky–Chow condition $c(\mathcal{L}) = \mathbb{R}^n$. First, we discuss differential-geometric concepts of [6], such as nonholonomic derivatives, the order of a function, privileged coordinates, etc., interpreting them in terms of the properties of the linear map c. In particular, the definition of privileged coordinates requires using properties of the shuffle product. Here the central role is played by R. Ree's theorem [46] on a connection between the Lie algebra and the shuffle product. Namely, we introduce the inner product in $\mathcal{F}$ so that $\{\eta_{i_1\ldots i_k}\}$ forms an orthonormal basis. Then R. Ree's Theorem says that

$$\mathcal{L} = (\mathcal{F} \sqcup\!\sqcup \mathcal{F})^{\perp},$$

where $^{\perp}$ denotes orthogonal complement. This theorem allows us to clarify the algebraic sense of privileged coordinates and to propose a way for constructing them (Subsection 3.4).

In Section 4 we give an algebraic interpretation for the concepts mentioned in Section 3. First, generalizing R. Ree's theorem, we study the properties of the left ideal $\mathcal{J}_{X_1,\ldots,X_m} = \mathrm{Lin}\{\mathcal{F}^e\mathcal{L}_{X_1,\ldots,X_m}\}$ [51, 55, 54]. It is shown that this ideal gives a description of a principal part of the series representation. Namely, we obtain the "approximation theorem" (Theorems 4.21 and 4.22), which can be briefly formulated as follows: The endpoint map of a system of the form (1.1) can be reduced (by a polynomial change of coordinates) to the form

$$\left(\mathcal{E}_{\widehat{X}_1,\ldots,\widehat{X}_m}\right)_i = \widetilde{\ell}_i + \widehat{\rho}_i, \qquad i = 1, \ldots, n,$$

where elements $\ell_i \in \mathcal{L}^{w_i}$ are such that

$$\mathcal{L} = \mathrm{Lin}\{\ell_1, \ldots, \ell_n\} + \mathcal{L}_{X_1,\ldots,X_m},$$

$\widetilde{\ell}_i$ denotes the orthogonal projection of ℓ_i onto the subspace $\mathcal{J}^{\perp}_{X_1,\ldots,X_m}$, $\widehat{\rho}_i \in \bigoplus_{j=w_i+1}^{\infty} \mathcal{F}^j$. (Here $\widehat{X}_1,\ldots,\widehat{X}_m$ denote vector fields on the right hand side of the system in the new coordinates.) Another way of constructing the principal part is to use the basis which is dual to the Poincaré–Birkhoff–Witt basis of $\mathcal{F}$ (Subsection 4.4). We also give a description of all privileged coordinates [53], i.e., the coordinates in which such a representation holds (Subsection 4.7).

The "series" $\mathcal{E}$ with coordinates $\mathcal{E}_i = \widetilde{\ell}_i$ can be considered as a principal part of the endpoint map $\mathcal{E}_{\widehat{X}_1,\ldots,\widehat{X}_m}$. If this series is realized as a control-linear system then such a system can be considered as an approximation of the initial system. In Section 5 we consider the realizability problem [18, 28, 29, 30] and show that the above-mentioned series can be realized as a control-linear system. Therefore, the core Lie subalgebra $\mathcal{L}_{X_1,\ldots,X_m}$ (or, what is the same, the left ideal $\mathcal{J}_{X_1,\ldots,X_m}$) really defines a homogeneous approximation of the system. We propose an algebraic definition for a homogeneous approximation (Subsection 5.2, Definition 5.9, and Remark 5.10). Namely, the "approximation" property means that two systems have the same core Lie subalgebra whereas the "homogeneous" property means that $c(\mathcal{L}_{X_1,\ldots,X_m}) = 0$. In particular, this implies that the homogeneous approximation is unique, up to a (polynomial homogeneous) change of variables. Moreover, we show that a core Lie subalgebra can be an arbitrary graded Lie subalgebra of codimension n, which gives a complete algebraic classification of possible homogeneous approximations (Remark 5.11).

Section 6 is devoted to the important particular cases, namely, regular systems and homogeneous systems. Recall that $v = (v_1,\ldots,v_p)$ is called the *growth vector* of the system (at the origin) if $v_k = \dim c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^k)$, $k = 1,\ldots,p$, and $v_p = n$. We consider the growth vectors at all points of a certain neighborhood $U(0)$ of the origin. A system is called *regular* if its growth vector is constant in $U(0)$. We show that the core Lie subalgebra $\mathcal{L}_{X_1,\ldots,X_m}$ of a regular system is a Lie ideal (Lemma 6.4) or, what is the same, its left ideal $\mathcal{J}_{X_1,\ldots,X_m}$ is two-sided (Lemma 6.6). The converse is true for homogeneous systems: if a system is homogeneous and its core Lie subalgebra is a Lie ideal then this system is regular (Theorem 6.13). As is shown in Subsection 6.3, for a homogeneous system one can find its series representation at any point using only the information on its core Lie subalgebra $\mathcal{L}_{X_1,\ldots,X_m}$ (Lemma 6.11).

Finally, in Section 7 we study the connection between the homogeneous approximation, the sub-Riemannian metrics [6, 41, 7, 31], and the time optimality. Namely, we consider the time-optimal control problem for a control-linear system of the form

$$\dot{x} = \sum_{i=1}^{m} u_i(t) X_i(x), \qquad x(0) = 0,\ x(\theta) = s,\ \sum_{i=1}^{m} u_i^2(t) \le 1, \qquad \theta \to \min .$$

First, we prove that time-optimal controls $u^*(t)$ satisfy the equality $\sum_{i=1}^{m} u_i^{*\,2}(t) = 1$ a.e. (see Theorem 7.1). (This property is commonly accepted, but we could not find a complete and rigorous proof in the literature.) This theorem allows us to give a partial answer to the question analogous to the open problem proposed in [52] (see Remark 7.19). Since the time-optimal control also minimizes the length functional (Corollary 7.2), the optimal time coincides with the sub-Riemannian distance from the origin to the point s.

In Subsection 7.3 we introduce the concept of approximation in the sense of time optimality (Definition 7.16); one of the requirements of this definition, in essence, implies approximation in the sense of sub-Riemannian metrics. The main result of Section 7 is Theorem 7.17 describing conditions under which the homogeneous approximation of a control-linear system approximates it in the sense of time optimality.

Finally, we mention that the results of Sections 2, 4, and 5 that belong to the authors of the present paper are based on the original approach proposed in [49] and [51] for the case of control-affine systems; they can be found in [50], [53], [54], [25]. The results of Sections 6 and 7 (except Subsection 7.1) are mainly new.

2. Series method in a local analysis of control-linear systems

2.1. Endpoint map. In this paper we consider the class of control-linear systems of the form

$$\dot{x} = \sum_{i=1}^{m} u_i X_i(x), \qquad x \in U(0) \subset \mathbb{R}^n,\ u_1, \dots, u_m \in \mathbb{R}, \tag{2.1}$$

where $X_1(x), \dots, X_m(x)$ are real analytic vector fields in a neighborhood of the origin $U(0) \subset \mathbb{R}^n$. Below we are mainly interested in the behavior of trajectories of system (2.1) starting at the origin,

$$x(0) = 0. \tag{2.2}$$

For any $\theta > 0$, by $L_\infty([0, \theta]; \mathbb{R}^m)$ we denote the space of measurable and almost everywhere bounded vector functions $u(t) = (u_1(t), \dots, u_m(t))$, $t \in [0, \theta]$, with the norm

$$\|u\| = \operatorname*{ess\,sup}_{t \in [0,\theta]} \sqrt{\sum_{i=1}^{m} u_i^2(t)}.$$

By B^θ we denote the unit ball of the space $L_\infty([0, \theta]; \mathbb{R}^m)$,

$$B^\theta = \Big\{ u(t) = (u_1(t), \dots, u_m(t)) \in L_\infty([0, \theta]; \mathbb{R}^m) : \sum_{i=1}^{m} u_i^2(t) \le 1 \text{ a.e.},\ t \in [0, \theta] \Big\}.$$

Throughout the paper we consider systems of the form (2.1) with controls $u \in B^\theta$, $\theta > 0$. If the vector fields $X_1(x), \dots, X_m(x)$ are fixed then there exists $T_0 > 0$ such that for any $\theta \in (0, T_0)$ trajectories of (2.1)–(2.2) corresponding to such controls are well defined.

Now we introduce one of the central concepts of this section.

Definition 2.1. For any $\theta \in (0, T_0)$ and $u \in B^\theta$, denote by $x(t; u)$ the solution of the Cauchy problem (2.1)–(2.2). Suppose the mapping $\mathcal{E}_{X_1,\dots,X_m}$ takes a pair (θ, u) to the end point of the trajectory, i.e.,

$$\mathcal{E}_{X_1,\dots,X_m}(\theta, u) = x(\theta; u).$$

We call $\mathcal{E}_{X_1,\dots,X_m}$ the *endpoint map* (at the origin) of system (2.1).

In the present paper we study local (for small θ) properties of this map.

2.2. Series representation. We begin with a representation of $\mathcal{E}_{X_1,\dots,X_m}(\theta,u)$ depending on θ and u, and not including a trajectory $x(t;u)$. Such representations, which generalize the well-known Cauchy formula for linear differential equations, were proposed by V. Volterra and developed by N. Wiener who used series of multidimensional integrals to describe the response of nonlinear systems. Discussion of different approaches can be found in [9, 22, 40, 1, 56, 27, 13, 57, 35, 21]. For control-affine systems, M. Fliess [17, 19, 20] proposed to apply the Chen series [10]. This leads to the following theorem which is a partial case of the result of M. Fliess [19].

Theorem 2.2 (M. Fliess [19]). *Consider a system of the form* (2.1) *and suppose that the vector fields $X_1,\dots,X_m$ are real analytic in a neighborhood of the origin. Then there exists $T\in(0,T_0]$ such that the endpoint map is represented in the form of a series*

$$\mathcal{E}_{X_1,\dots,X_m}(\theta,u)=\sum_{k=1}^{\infty}\sum_{1\le i_1,\dots,i_k\le m} c_{i_1\dots i_k}\eta_{i_1\dots i_k}(\theta,u), \tag{2.3}$$

which is absolutely convergent for any $\theta\in(0,T)$ and any $u\in B^\theta$, where

$$\eta_{i_1\dots i_k}(\theta,u)=\int_0^\theta\int_0^{\tau_1}\cdots\int_0^{\tau_{k-1}} u_{i_1}(\tau_1)u_{i_2}(\tau_2)\cdots u_{i_k}(\tau_k)\,d\tau_k\cdots d\tau_2\,d\tau_1 \tag{2.4}$$

are "iterated integrals" and $c_{i_1\dots i_k}$ are constant vector coefficients that can be found by

$$c_{i_1\dots i_k}=X_{i_k}\cdots X_{i_1}E(0), \tag{2.5}$$

where $E(x)=x$ is the identity map.

Remark 2.3. On the right hand side of (2.5), we regard the vector fields X_i as the differential operators of the first order defined as $X_i\psi=\psi'_xX_i$. Then a composition of k such operators $X_{i_k}\cdots X_{i_1}$ is the differential operator of order k. Throughout this paper we consider such operators as acting on vector functions, assuming that this action is componentwise. We defer the detailed discussion to Subsection 2.5.

Remark 2.4. Equality (2.4) says that iterated integrals depend on θ and u. To be more precise, below we consider them as functionals of u for any fixed θ. We discuss the exact sense of the iterated integrals in Subsection 2.3.

Remark 2.5. Let us clarify the convergence of the series more specifically. Since the vector fields $X_1,\dots,X_m$ are real analytic, there exist positive constants C_1 and C_2 such that the estimates $\|c_{i_1\dots i_k}\|\le C_1C_2^kk!$ hold. Since $|\eta_{i_1\dots i_k}(\theta,u)|\le\frac{1}{k!}\theta^k$ for any $u\in B^\theta$, for any $k\ge1$ we get

$$\Big\|\sum_{1\le i_1,\dots,i_k\le m} c_{i_1\dots i_k}\eta_{i_1\dots i_k}(\theta,u)\Big\|\le C_1(mC_2\theta)^k. \tag{2.6}$$

Hence, the series in (2.3) is absolutely convergent if $mC_2\theta<1$. This gives the condition for T, namely, $T<\frac{1}{mC_2}$. Below, without loss of generality, we assume $T=T_0$.

For the sake of completeness, we give a sketch of the proof of Theorem 2.2. The main goal here is to show that the proof does not require any special methods and additional concepts. For brevity, we write $X_{i_k}\cdots X_{i_1}(x)$ instead of $X_{i_k}\cdots X_{i_1}E(x)$.

Suppose $\theta > 0$ is sufficiently small, a control $u(t)$ is fixed, and $x(t) = x(t;u)$ is the solution of the Cauchy problem (2.1)–(2.2). Integrating (2.1) with respect to t from 0 to θ and taking into account (2.2), we get

$$x(\theta) = \sum_{i=1}^{m} \int_0^\theta X_i(x(t))u_i(t)\,dt. \tag{2.7}$$

Note that

$$\frac{d}{dt}X_i(x(t)) = (X_i(x(t)))'_x\dot{x}(t) = \sum_{j=1}^{m}(X_i(x(t)))'_x X_j(x(t))u_j(t) = \sum_{j=1}^{m} X_jX_i(x(t))u_j(t)$$

and

$$u_i(t) = -\frac{d}{dt}\int_t^\theta u_i(\tau)\,d\tau.$$

Then, integrating by parts the right hand side of (2.7), we get

$$\begin{aligned} x(\theta) &= \sum_{i=1}^{m}\Big(-X_i(x(t))\int_t^\theta u_i(\tau)\,d\tau\Big|_0^\theta + \int_0^\theta \sum_{j=1}^{m} X_jX_i(x(\tau_1))u_j(\tau_1)\int_{\tau_1}^\theta u_i(\tau_2)\,d\tau_2\,d\tau_1\Big) \\ &= \sum_{i=1}^{m} c_i\eta_i(\theta,u) + \sum_{1\le i_1,i_2\le m}\int_0^\theta X_{i_2}X_{i_1}(x(\tau_1))u_{i_2}(\tau_1)\int_{\tau_1}^\theta u_{i_1}(\tau_2)\,d\tau_2\,d\tau_1. \end{aligned}$$

We can repeat the described procedure, integrating by parts the second term on the right hand side of the last equality, and so on. After q such steps we obtain

$$x(\theta) = \sum_{k=1}^{q}\sum_{1\le i_1,\dots,i_k\le m} c_{i_1\dots i_k}\eta_{i_1\dots i_k}(\theta,u) + R_q(\theta,u),$$

where

$$\begin{aligned} &R_q(\theta,u) \\ &= \sum_{1\le i_1,\dots,i_{q+1}\le m}\int_0^\theta\int_{\tau_1}^\theta\cdots\int_{\tau_q}^\theta X_{i_{q+1}}\cdots X_{i_1}(x(\tau_1))u_{i_{q+1}}(\tau_1)\cdots u_{i_1}(\tau_{q+1})\,d\tau_{q+1}\cdots d\tau_2\,d\tau_1. \end{aligned}$$

By use of the analyticity of the vector fields $X_1,\dots,X_m$, it is not hard to prove that $R_q(\theta,u)\to 0$ as $q\to\infty$ for any sufficiently small $\theta>0$ and any $u\in B^\theta$. This completes the proof of Theorem 2.2.

Let us briefly discuss representation (2.3). The right hand side of (2.3) includes "objects" of two kinds. The objects of the first kind are the constant coefficients—vectors in $\mathbb{R}^n$—of the form (2.5). They are determined by the vector fields $X_1,\dots,X_m$ (more precisely, by the values of these vector fields and their derivatives at the origin) and, moreover, they depend on local coordinates. The objects of the second kind are the iterated integrals (2.4). They are "completely independent" in the sense that they are the same for all systems of the form (2.1). It turns out that the set of iterated integrals can be regarded as a *free associative algebra*; we introduce it in the next subsection.

2.3. Iterated integrals and free associative algebras. Let us now introduce the exact definition of iterated integrals.

Definition 2.6. For $\theta > 0$, $k \geq 1$, and $1 \leq i_1, \ldots, i_k \leq m$, consider the functional $\eta_{i_1\ldots i_k}(\theta, \cdot) : B^\theta \to \mathbb{R}$ that takes each control $u \in B^\theta$ to the number $\eta_{i_1\ldots i_k}(\theta, u)$ defined by (2.4). This functional is called an *iterated integral* [19].

Note that the linear span (over $\mathbb{R}$) of all iterated integrals equipped with the *concatenation product operation*

$$\eta_{i_1\ldots i_k}(\theta, \cdot) \vee \eta_{j_1\ldots j_s}(\theta, \cdot) = \eta_{i_1\ldots i_k j_1\ldots j_s}(\theta, \cdot)$$

forms an associative algebra. Moreover, one-dimensional integrals $\eta_i(\theta, \cdot)$, $i = 1, \ldots, m$, can be considered as the generators of this algebra, so one can write

$$\eta_{i_1\ldots i_k}(\theta, \cdot) = \eta_{i_1}(\theta, \cdot) \vee \cdots \vee \eta_{i_k}(\theta, \cdot).$$

Here we use $\vee$ to avoid confusing concatenation with multiplication of integrals as real numbers (when $u \in B^\theta$ is substituted).

In this subsection we give the exact definition of this algebra and discuss some of its properties.

Below we often deal with controls defined on different intervals. For the sake of convenience, let us adopt the following notation.

Notation 2.7. By definition, for any $\alpha > 0$ and any $u(t)$, $t \in [0, \beta]$, set $u^\alpha(t) = u(\alpha t)$, $t \in [0, \beta/\alpha]$.

In particular, for any $\theta > 0$ one has $u(t) = u^\theta(t/\theta)$. Taking this into account, let us rewrite an iterated integral of the form (2.4) in the following way:

$$\begin{aligned}\eta_{i_1\ldots i_k}(\theta, u) &= \int_0^\theta \int_0^{\tau_1} \cdots \int_0^{\tau_{k-1}} u_{i_1}(\tau_1) u_{i_2}(\tau_2) \cdots u_{i_k}(\tau_k)\, d\tau_k \cdots d\tau_2\, d\tau_1 \\ &= \int_0^\theta \int_0^{\tau_1} \cdots \int_0^{\tau_{k-1}} u^\theta_{i_1}\left(\frac{\tau_1}{\theta}\right) u^\theta_{i_2}\left(\frac{\tau_2}{\theta}\right) \cdots u^\theta_{i_k}\left(\frac{\tau_k}{\theta}\right) d\tau_k \cdots d\tau_2\, d\tau_1 \\ &= \theta^k \int_0^1 \int_0^{\tau_1} \cdots \int_0^{\tau_{k-1}} u^\theta_{i_1}(\tau_1) u^\theta_{i_2}(\tau_2) \cdots u^\theta_{i_k}(\tau_k)\, d\tau_k \cdots d\tau_2\, d\tau_1 \\ &= \theta^k \eta_{i_1\ldots i_k}(1, u^\theta).\end{aligned}$$

This equality holds for any $u \in B^\theta$ or, what is the same, for any $u^\theta \in B^1$. In other words, for any $\theta > 0$ and any $u \in B^1$ we have

$$\eta_{i_1\ldots i_k}(\theta, u^{1/\theta}) = \theta^k \eta_{i_1\ldots i_k}(1, u).$$

Hence, k equals the asymptotic order of the iterated integral $\eta_{i_1\ldots i_k}(\theta, u^{1/\theta})$ with respect to θ as $\theta \to 0$ for any fixed control $u \in B^1$ such that $\eta_{i_1\ldots i_k}(1, u) \neq 0$. This justifies the following

Definition 2.8. We say that k is the *order* of the iterated integral $\eta_{i_1\ldots i_k}(\theta, \cdot)$.

Notice that this notion of order corresponds to the order in which the terms of the series (2.3) are added.

Definition 2.9. Suppose $\theta > 0$ is fixed. Consider the associative algebra $\mathcal{F}_\theta$ of functionals (over $\mathbb{R}$)

$$\mathcal{F}_\theta = \mathrm{Lin}\{\eta_{i_1\ldots i_k}(\theta, \cdot) : k \geq 1,\, 1 \leq i_1, \ldots, i_k \leq m\},$$

with the product operation

$$\eta_{i_1\dots i_k}(\theta,\cdot)\vee\eta_{j_1\dots j_s}(\theta,\cdot)=\eta_{i_1\dots i_k j_1\dots j_s}(\theta,\cdot).$$

We call $\mathcal{F}_\theta$ the *Fliess algebra* or the *algebra of iterated integrals.* One-dimensional integrals $\eta_i(\theta,\cdot)$, $i=1,\dots,m$, are the generators of $\mathcal{F}_\theta$. The natural filtration is given by the sequence of subspaces $\sum_{k=1}^q\mathcal{F}_\theta^k$, $q\geq 1$, where

$$\mathcal{F}_\theta^k=\mathrm{Lin}\{\eta_{i_1\dots i_k}(\theta,\cdot):1\leq i_1,\dots,i_k\leq m\},\qquad k\geq 1.$$

The main observation here is that this associative algebra is *free* [19]. Before proving this claim, let us give some preliminary remarks. Suppose the control $u^1(t)$, $t\in[0,\theta^1]$, steers the origin to the point z, and the control $u^2(t)$, $t\in[0,\theta^2]$, steers the point z to the point x. More precisely, the solution $x^1(t)$ of the Cauchy problem

$$\dot{x}=\sum_{i=1}^m u_i^1(t)X_i(x),\qquad x(0)=0,$$

satisfies the condition $x^1(\theta^1)=z$, and the solution $x^2(t)$ of the Cauchy problem

$$\dot{x}=\sum_{i=1}^m u_i^2(t)X_i(x),\qquad x(0)=z,$$

satisfies $x^2(\theta^2)=x$. Let us denote by $u^1\circ u^2$ the *concatenation* of controls $u^1(t)$ and $u^2(t)$ defined by

$$(u^1\circ u^2)(t)=\begin{cases}u^1(t) & \text{for } t\in[0,\theta^1],\\ u^2(t-\theta^1) & \text{for } t\in(\theta^1,\theta^1+\theta^2].\end{cases}\tag{2.8}$$

Then, obviously, the control $u^1\circ u^2$ steers the origin to the point x, i.e., the solution $x^3(t)$ of the Cauchy problem

$$\dot{x}=\sum_{i=1}^m(u^1\circ u^2)_i(t)X_i(x),\qquad x(0)=0,$$

satisfies the condition $x^3(\theta^1+\theta^2)=x$.

LEMMA 2.10. *For any controls $u^1\in B^{\theta^1}$ and $u^2\in B^{\theta^2}$, and any iterated integral, the following identity holds:*

$$\eta_{i_1\dots i_k}(\theta^1+\theta^2,u^1\circ u^2)=\sum_{j=0}^k\eta_{i_1\dots i_j}(\theta^2,u^2)\,\eta_{i_{j+1}\dots i_k}(\theta^1,u^1),$$

where for any θ and u it is assumed that $\eta_{i_s\dots i_q}(\theta,u)=1$ if $s>q$.

Proof. Denote $u=u^1\circ u^2$. Consider the integration domain for $\eta_{i_1\dots i_k}(\theta^1+\theta^2,u)$; it is a simplex in $\mathbb{R}^k$. Note that it can be represented as the union of $k+1$ polyhedrons

$$\begin{aligned}&\{(\tau_1,\dots,\tau_k):0\leq\tau_k\leq\dots\leq\tau_1\leq\theta^1+\theta^2\}\\&\qquad=\bigcup_{j=0}^k\{(\tau_1,\dots,\tau_k):0\leq\tau_k\leq\dots\leq\tau_{j+1}\leq\theta^1\leq\tau_j\leq\dots\leq\tau_1\leq\theta^1+\theta^2\}\end{aligned}$$

with pairwise nonintersecting interiors. Moreover, each polyhedron equals the Cartesian product of two simplices. Hence, $\eta_{i_1\ldots i_k}(\theta^1+\theta^2,u)$ equals the sum (over $j=0,\ldots,k$) of the integrals

$$\int_{\theta^1}^{\theta^1+\theta^2}\cdots\int_{\theta^1}^{\tau_{j-1}}\int_0^{\theta^1}\cdots\int_0^{\tau_{k-1}} u_{i_1}(\tau_1)\cdots u_{i_k}(\tau_k)\,d\tau_k\cdots d\tau_1$$
$$=\Big(\int_{\theta^1}^{\theta^1+\theta^2}\cdots\int_{\theta^1}^{\tau_{j-1}} u_{i_1}(\tau_1)\cdots u_{i_j}(\tau_j)\,d\tau_j\cdots d\tau_1\Big)$$
$$\times\Big(\int_0^{\theta^1}\cdots\int_0^{\tau_{k-1}} u_{i_{j+1}}(\tau_{j+1})\cdots u_{i_k}(\tau_k)\,d\tau_k\cdots d\tau_{j+1}\Big).$$

Taking into account (2.8), we rewrite this expression as $\eta_{i_1\ldots i_j}(\theta^2,u^2)\eta_{i_{j+1}\ldots i_k}(\theta^1,u^1)$. ■

Now we are ready to prove the following result.

Lemma 2.11 ([19]). *Let $\theta>0$ be fixed. Suppose*

$$\sum_{k\ge1,\,1\le i_1,\ldots,i_k\le m}\alpha_{i_1\ldots i_k}\eta_{i_1\ldots i_k}(\theta,u)=0 \tag{2.9}$$

for all $u\in B^\theta$, where $\alpha_{i_1\ldots i_k}\in\mathbb{R}$ and only a finite number of terms on the left hand side are nonzero. Then all coefficients $\alpha_{i_1\ldots i_k}$ on the left hand side vanish.

As a consequence, for any $\theta>0$ the algebra $\mathcal{F}_\theta$ is free, and the representation $\mathcal{F}_\theta=\sum_{k=1}^\infty\mathcal{F}_\theta^k$ defines a graded structure.

Proof. Below we use the equality $\eta_{i_1\ldots i_k}(T,u)=T^k\eta_{i_1\ldots i_k}(1,u^T)$, which holds for any $T>0$. Notice that here u ranges over the set B^T iff u^T ranges over B^1.

First, for any $\tau\in[0,\theta]$ consider an arbitrary control $u\in B^\theta$ such that $u(t)=0$ for $t\in[\tau,\theta]$. Then $\eta_{i_1\ldots i_k}(\theta,u)=\eta_{i_1\ldots i_k}(\tau,u)=\tau^k\eta_{i_1\ldots i_k}(1,u^\tau)$ for arbitrary $u^\tau\in B^1$. Hence, (2.9) implies

$$\sum_{k\ge1}\tau^k\sum_{1\le i_1,\ldots,i_k\le m}\alpha_{i_1\ldots i_k}\eta_{i_1\ldots i_k}(1,u)=0,\qquad u\in B^1.$$

For any fixed $u\in B^1$, the left hand side is a polynomial in $\tau\in[0,\theta]$, hence for any $k\ge1$,

$$\sum_{1\le i_1,\ldots,i_k\le m}\alpha_{i_1\ldots i_k}\eta_{i_1\ldots i_k}(1,u)=0,\qquad u\in B^1. \tag{2.10}$$

Thus, the statement of the lemma is reduced to the following claim: if (2.10) holds for all $u\in B^1$ then $\alpha_{i_1\ldots i_k}=0$ for all $1\le i_1,\ldots,i_k\le m$.

We prove this claim by induction on k. For $k=1$, the proof is clear. For any $k\ge2$, suppose that the equality

$$\sum_{1\le i_1,\ldots,i_{k-1}\le m}\widetilde{\alpha}_{i_1\ldots i_{k-1}}\eta_{i_1\ldots i_{k-1}}(1,u)=0,\qquad u\in B^1,$$

yields $\widetilde{\alpha}_{i_1\ldots i_{k-1}}=0$ for all $1\le i_1,\ldots,i_{k-1}\le m$. Take an arbitrary $t>0$ and two controls $u^1\in B^1$ and $u^2\in B^t$. It follows from (2.10) that

$$\sum_{1\le i_1,\ldots,i_k\le m}\alpha_{i_1\ldots i_k}\eta_{i_1\ldots i_k}(T,u)=0,\qquad u\in B^T,$$

for any $T > 0$. Hence, setting $T = 1 + t$ and $u = u^1 \circ u^2$, and applying Lemma 2.10, we get

$$\sum_{1\le i_1,\dots,i_k\le m} \alpha_{i_1\dots i_k}\eta_{i_1\dots i_k}(1+t, u^1\circ u^2)$$
$$= \sum_{1\le i_1,\dots,i_k\le m} \alpha_{i_1\dots i_k}\sum_{j=0}^{k}\eta_{i_1\dots i_j}(t,u^2)\eta_{i_{j+1}\dots i_k}(1,u^1)$$
$$= \sum_{1\le i_1,\dots,i_k\le m} \alpha_{i_1\dots i_k}\sum_{j=0}^{k} t^j\eta_{i_1\dots i_j}(1,(u^2)^t)\eta_{i_{j+1}\dots i_k}(1,u^1) = 0.$$

Denote $u^3 = (u^2)^t \in B^1$. Then the last equality can be rewritten as

$$\sum_{j=0}^{k} t^j \sum_{1\le i_1,\dots,i_k\le m} \alpha_{i_1\dots i_k}\eta_{i_1\dots i_j}(1,u^3)\,\eta_{i_{j+1}\dots i_k}(1,u^1) = 0, \qquad u^1, u^3 \in B^1.$$

For any fixed $u^1, u^3 \in B^1$ the left hand side is a polynomial in t, hence, in particular,

$$\sum_{1\le i_1,\dots,i_k\le m} \alpha_{i_1\dots i_k}\eta_{i_1\dots i_{k-1}}(1,u^3)\,\eta_{i_k}(1,u^1) = 0, \qquad u^1, u^3 \in B^1.$$

For any fixed $u^1 \in B^1$ we can rewrite this equality as

$$\sum_{1\le i_1,\dots,i_{k-1}\le m}\Big(\sum_{1\le i_k\le m} \alpha_{i_1\dots i_k}\eta_{i_k}(1,u^1)\Big)\eta_{i_1\dots i_{k-1}}(1,u^3)$$
$$= \sum_{1\le i_1,\dots,i_{k-1}\le m} \widetilde{\alpha}_{i_1\dots i_{k-1}}\eta_{i_1\dots i_{k-1}}(1,u^3) = 0.$$

Hence, by the induction assumption,

$$\widetilde{\alpha}_{i_1\dots i_{k-1}} = \sum_{1\le i_k\le m} \alpha_{i_1\dots i_k}\eta_{i_k}(1,u^1) = 0, \qquad u^1 \in B^1,$$

and therefore $\alpha_{i_1\dots i_k} = 0$. ∎

COROLLARY 2.12. *Let $\theta > 0$ be fixed. Suppose*

$$\sum_{k=1}^{\infty}\sum_{1\le i_1,\dots,i_k\le m} \alpha_{i_1\dots i_k}\eta_{i_1\dots i_k}(\theta,u) = 0 \tag{2.11}$$

for all $u \in B^\theta$, where $\alpha_{i_1\dots i_k} \in \mathbb{R}$ satisfy the estimate $|\alpha_{i_1\dots i_k}| \le C_1C_2^k k!$, $C_1, C_2 > 0$, $mC_2\theta < 1$. Then all coefficients $\alpha_{i_1\dots i_k}$ on the left hand side vanish.

As a consequence, the representation of the endpoint map $\mathcal{E}_{X_1,\dots,X_m}(\theta,u)$ in the form of a series of iterated integrals is unique.

Proof. As in the proof of the previous lemma, for any $\tau \in [0,\theta]$ consider arbitrary controls $u \in B^\theta$ such that $u(t) = 0$ for $t \in [\tau,\theta]$. Then (2.11) implies that for any fixed $u \in B^1$,

$$\sum_{k=1}^{\infty}\tau^k\sum_{1\le i_1,\dots,i_k\le m} \alpha_{i_1\dots i_k}\eta_{i_1\dots i_k}(1,u) = 0,$$

i.e., the convergent power series in τ vanishes. Hence, for any $k \ge 1$,

$$\sum_{1\le i_1,\dots,i_k\le m} \alpha_{i_1\dots i_k}\eta_{i_1\dots i_k}(1,u) = 0, \qquad u \in B^1,$$

where the sum on the left hand side is finite. Now the statement follows from Lemma 2.11. ∎

Thus, due to Lemma 2.11, the algebra of functionals $\mathcal{F}_\theta$ is free (for any $\theta > 0$). This motivates introducing an abstract free associative graded algebra generated by m elements. Namely, let us consider the set of m abstract free elements called *letters*; we denote them by $\eta_1, \ldots, \eta_m$. Strings of letters are called *words*; we denote them by $\eta_{i_1\ldots i_k} = \eta_{i_1} \cdots \eta_{i_k}$. In the set of words, the natural concatenation operation is introduced:

$$\eta_{i_1\ldots i_k} \cdot \eta_{j_1\ldots j_s} = \eta_{i_1\ldots i_k j_1\ldots j_s}.$$

Below we usually omit the sign of this operation.

All finite linear combinations of words (over $\mathbb{R}$) form a free associative algebra with the natural gradation $\mathcal{F} = \bigoplus_{k=1}^{\infty} \mathcal{F}^k$, where the homogeneous subspace $\mathcal{F}^k$ is defined as the linear span of products of k generators,

$$\mathcal{F}^k = \mathrm{Lin}\{\eta_{i_1\ldots i_k} = \eta_{i_1} \cdots \eta_{i_k} : 1 \le i_1, \ldots, i_k \le m\}, \qquad k \ge 1. \tag{2.12}$$

Then $\mathcal{F}$ is naturally isomorphic to $\mathcal{F}_\theta$ for any $\theta > 0$.

NOTATION 2.13. By $\mathcal{F}$ we denote a free associative algebra (over $\mathbb{R}$) with m (abstract) generators $\eta_1, \ldots, \eta_m$ and the natural gradation $\mathcal{F} = \bigoplus_{k=1}^{\infty} \mathcal{F}^k$, where the homogeneous subspaces $\mathcal{F}^k$ are given by (2.12).

In other words, $\mathcal{F}$ is the associative $\mathbb{R}$-algebra of formal noncommuting polynomials of m independent variables. Lemma 2.11 implies that the algebras $\mathcal{F}_\theta$ and $\mathcal{F}$ are isomorphic.

Sometimes it is convenient to supplement the algebra $\mathcal{F}$ with the unity element 1 (which can be thought of as the empty word) and consider the algebra

$$\mathcal{F}^e = \mathcal{F} + \mathbb{R}$$

assuming $1 \cdot a = a \cdot 1 = a$ for any $a \in \mathcal{F}^e$. Throughout the paper we assume $\eta_{i_p\ldots i_q} = 1$ if $p > q$.

Taking into account the graded structure, we introduce the following convenient definition.

DEFINITION 2.14. We say that an element $a \in \mathcal{F}$ *is of order* k and write $\mathrm{ord}(a) = k$ iff $a \in \mathcal{F}^k$. If an element is of some order, we say that it is *homogeneous*.

We also introduce the free Lie algebra $\mathcal{L}$ which is generated by the same set of generators $\eta_1, \ldots, \eta_m$ with bracket $[\ell_1, \ell_2] = \ell_1\ell_2 - \ell_2\ell_1$. (Notice that $\mathcal{F}$ is the universal enveloping for $\mathcal{L}$.) It inherits the gradation $\mathcal{L} = \bigoplus_{k=1}^{\infty} \mathcal{L}^k$, where $\mathcal{L}^k = \mathcal{L} \cap \mathcal{F}^k$, $k \ge 1$. The Lie algebra $\mathcal{L}$ will play an important role in our further constructions.

REMARK 2.15. Below we systematically consider formal power series of elements of $\mathcal{F}$ over $\mathbb{R}$ or $\mathbb{R}^n$. Namely, if the sum in $a = \sum \alpha_{i_1\ldots i_k} \eta_{i_1\ldots i_k}$ (where the coefficients $\alpha_{i_1\ldots i_k}$ are from $\mathbb{R}$ or $\mathbb{R}^n$) is taken over an infinite set of indices, we mean that a is a formal power series.

Thus, along with the endpoint map and its series representation (2.3), we can consider its "abstract analog", the formal power series (with coefficients in $\mathbb{R}^n$) of elements of $\mathcal{F}$

of the form

$$\mathcal{E}_{X_1,\dots,X_m} = \sum_{k=1}^{\infty} \sum_{1\le i_1,\dots,i_k\le m} c_{i_1\dots i_k}\eta_{i_1\dots i_k}. \tag{2.13}$$

REMARK 2.16. Corollary 2.12 implies that there exists a unique formal power series (2.13) corresponding to the endpoint map $\mathcal{E}_{X_1,\dots,X_m}(\theta,u)$, i.e., to the Cauchy problem (2.1)–(2.2). A description of all such formal power series is given in Section 5.

2.4. Changes of variables and shuffles. Notice that a change of variables in system (2.1) leads to some transformation of the series representation of the endpoint map. Namely, suppose we know the series representation

$$\mathcal{E}_{X_1,\dots,X_m}(\theta,u) = \sum_{k=1}^{\infty} \sum_{1\le i_1,\dots,i_k\le m} c_{i_1\dots i_k}\eta_{i_1\dots i_k}(\theta,u),$$

where $c_{i_1\dots i_k}$ are constant vector coefficients. Clearly, this representation (due to Corollary 2.12) coincides with (2.3), however, here we "forget" that the coefficients $c_{i_1\dots i_k}$ can be found via the vector fields $X_1,\dots,X_m$ by formula (2.5).

Suppose $y=Q(x)$ is a real analytic change of variables defined in a neighborhood of the origin and such that $Q(0)=0$. Then in the new coordinates the initial system takes the form

$$\dot y = \sum_{i=1}^{m} u_i\widehat{X}_i(y), \qquad y\in \widehat{U}(0)\subset\mathbb{R}^n, \tag{2.14}$$

where $\widehat{X}_i(y) = Q'(x)X_i(x)|_{x=Q^{-1}(y)}$, $i=1,\dots,m$. For any sufficiently small $\theta>0$ and any $u\in B^\theta$, we get

$$\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m}(\theta,u) = Q(\mathcal{E}_{X_1,\dots,X_m}(\theta,u)).$$

Let us find the series representation for the endpoint map $\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m}(\theta,u)$ of the system in the new variables (2.14). We are going to do this without using the explicit form of the vector fields $\widehat{X}_i(y)$. Instead, let us expand Q into a Taylor series, $Q(x)=\sum_{q=1}^{\infty}\frac{1}{q!}Q^{(q)}(0)x^q$, where, for brevity, we use the notation

$$Q^{(q)}(0)x^q = \sum_{j_1+\dots+j_n=q} \frac{q!}{j_1!\cdots j_n!}\,\frac{\partial^{j_1+\dots+j_n}Q(0)}{\partial x_1^{j_1}\cdots\partial x_n^{j_n}}\,x_1^{j_1}\cdots x_n^{j_n}.$$

Then we get the representation

$$\begin{aligned}\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m}(\theta,u) &= Q(\mathcal{E}_{X_1,\dots,X_m}(\theta,u)) = \sum_{q=1}^{\infty}\frac{1}{q!}Q^{(q)}(0)(\mathcal{E}_{X_1,\dots,X_m}(\theta,u))^q\\ &= \sum \alpha^{j_1\dots j_n}_{i_1^1\dots i_{k_1}^1\dots i_1^n\dots i_{k_n}^n}(\eta_{i_1^1\dots i_{k_1}^1}(\theta,u))^{j_1}\cdots(\eta_{i_1^n\dots i_{k_n}^n}(\theta,u))^{j_n},\end{aligned} \tag{2.15}$$

where

$$\alpha^{j_1\dots j_n}_{i_1^1\dots i_{k_1}^1\dots i_1^n\dots i_{k_n}^n} = \frac{1}{j_1!\cdots j_n!}\,\frac{\partial^{j_1+\dots+j_n}Q(0)}{\partial x_1^{j_1}\cdots\partial x_n^{j_n}}\,(c_{i_1^1\dots i_{k_1}^1})_1^{j_1}\cdots(c_{i_1^n\dots i_{k_n}^n})_n^{j_n},$$

$(v)_i$ denotes the ith component of the vector $v\in\mathbb{R}^n$, and the last sum in (2.15) is taken over all $j_1,\dots,j_n\ge 0$, all $k_1,\dots,k_n\ge 1$, and all $1\le i_1^1,\dots,i_{k_n}^n\le m$. (Here we do not care about convergence, because we are only interested in formal transformations; the convergence of the resulting series is guaranteed by the analyticity of the vector fields $X_1, \dots, X_m$ and the map Q.)

Now we are going to represent $\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m}(\theta, u)$ as a series of iterated integrals with constant vector coefficients. To this end, we need to express products of iterated integrals as linear combinations of such integrals.

Let us calculate the product of two iterated integrals. Notice that

$$\eta_{p_1 \dots p_q}(\theta, u) = \int_{0 \le \tau_q \le \dots \le \tau_1 \le \theta} \prod_{j=1}^{q} u_{p_j}(\tau_j)\, d\tau_1 \cdots d\tau_q.$$

So, we have

$$\begin{aligned} &\eta_{i_1 \dots i_k}(\theta, u)\eta_{i_{k+1} \dots i_{k+r}}(\theta, u) \\ &= \int_{0 \le \tau_k \le \dots \le \tau_1 \le \theta} \prod_{j=1}^{k} u_{i_j}(\tau_j)\, d\tau_1 \cdots d\tau_k \int_{0 \le \tau_{k+r} \le \dots \le \tau_{k+1} \le \theta} \prod_{j=k+1}^{r} u_{i_j}(\tau_j)\, d\tau_{k+1} \cdots d\tau_{k+r}. \end{aligned} \tag{2.16}$$

In order to multiply two integrals over the domains $0 \le \tau_k \le \dots \le \tau_1 \le \theta$ and $0 \le \tau_{k+r} \le \dots \le \tau_{k+1} \le \theta$, we should "shuffle" two sets of variables $\{\tau_1, \dots, \tau_k\}$ and $\{\tau_{k+1}, \dots, \tau_{k+r}\}$ in all possible ways, preserving the "interior order" in each set. The following definition is useful.

Definition 2.17. The sequence $(j_1, \dots, j_{k+r})$ is called a *shuffle permutation* of the sequences $(1, \dots, k)$ and $(k+1, \dots, k+r)$ if it is a permutation of the sequence $(1, \dots, k+r)$ and possesses the following property:

$$\text{if} \quad 1 \le j_p < j_q \le k \quad \text{or} \quad k+1 \le j_p < j_q \le k+r, \quad \text{then} \quad p < q.$$

We denote by $S_{k,r}$ the set of all such shuffle permutations.

Taking into account this definition, we obtain

$$\begin{aligned} &\int_{0 \le \tau_k \le \dots \le \tau_1 \le \theta} \prod_{j=1}^{k} u_{i_j}(\tau_j)\, d\tau_1 \cdots d\tau_k \int_{0 \le \tau_{k+r} \le \dots \le \tau_{k+1} \le \theta} \prod_{j=k+1}^{r} u_{i_j}(\tau_j)\, d\tau_{k+1} \cdots d\tau_{k+r} \\ &\qquad = \sum_{(j_1,\dots,j_{k+r}) \in S_{k,r}} \int_{0 \le \tau_{j_{k+r}} \le \dots \le \tau_{j_1} \le \theta} \prod_{q=1}^{k+r} u_{i_{j_q}}(\tau_{j_q})\, d\tau_{j_{k+r}} \cdots d\tau_{j_1}. \end{aligned}$$

Hence, (2.16) gives

$$\eta_{i_1 \dots i_k}(\theta, u)\eta_{i_{k+1} \dots i_{k+r}}(\theta, u) = \sum_{(j_1,\dots,j_{k+r}) \in S_{k,r}} \eta_{i_{j_1} \dots i_{j_{k+r}}}(\theta, u). \tag{2.17}$$

In an associative algebra, the corresponding operation is called the shuffle product [14, 46, 10, 2].

Definition 2.18. The *shuffle product* $\sqcup\!\sqcup$ in $\mathcal{F}$ is defined by the rule

$$\eta_{i_1 \dots i_k} \sqcup\!\sqcup\, \eta_{i_{k+1} \dots i_{k+r}} = \sum_{(j_1,\dots,j_{k+r}) \in S_{k,r}} \eta_{i_{j_1} \dots i_{j_{k+r}}}.$$

This operation is commutative and associative.

Note that commutativity and associativity follow immediately from (2.17).

Thus, the "usual product" of iterated integrals as functionals corresponds to the shuffle product in the abstract algebra. One can express this statement as follows:

$$\eta_{i_1\dots i_k}(\theta,u)\eta_{s_1\dots s_r}(\theta,u)=(\eta_{i_1\dots i_k}\sqcup\!\sqcup\,\eta_{s_1\dots s_r})(\theta,u),$$

where on the right hand side we mean that first, one finds the shuffle product of abstract elements $\eta_{i_1\dots i_k}$ and $\eta_{s_1\dots s_r}$ in $\mathcal{F}$, and then replaces the resulting element of $\mathcal{F}$ by the corresponding element of $\mathcal{F}_\theta$.

From the practical point of view, it is more convenient to use another way of finding the shuffle product. It is convenient to extend the shuffle product to the algebra $\mathcal{F}^e$ assuming $1\sqcup\!\sqcup\, a=a\sqcup\!\sqcup\, 1=a$ for any $a\in\mathcal{F}^e$. Then it can be easily proved that Definition 2.18 is equivalent to the following

Definition 2.19. The *shuffle product* in $\mathcal{F}$ is defined by the recurrent formula

$$\eta_{i_1\dots i_k}\sqcup\!\sqcup\,\eta_{j_1\dots j_r}=(\eta_{i_1\dots i_{k-1}}\sqcup\!\sqcup\,\eta_{j_1\dots j_r})\eta_{i_k}+(\eta_{i_1\dots i_k}\sqcup\!\sqcup\,\eta_{j_1\dots j_{r-1}})\eta_{j_r},\qquad k,r\ge 1,\tag{2.18}$$

or, which gives the same,

$$\eta_{i_1\dots i_k}\sqcup\!\sqcup\,\eta_{j_1\dots j_r}=\eta_{i_1}(\eta_{i_2\dots i_k}\sqcup\!\sqcup\,\eta_{j_1\dots j_r})+\eta_{j_1}(\eta_{i_1\dots i_k}\sqcup\!\sqcup\,\eta_{j_2\dots j_r}),\qquad k,r\ge 1.\tag{2.19}$$

These formulas admit the following generalization, which can also be easily obtained from Definition 2.18.

Lemma 2.20. *For any* $0\le s\le k+r$,

$$\eta_{i_1\dots i_k}\sqcup\!\sqcup\,\eta_{j_1\dots j_r}=\sum_{\substack{0\le q\le k,\,0\le t\le r\\ q+t=s}}(\eta_{i_1\dots i_q}\sqcup\!\sqcup\,\eta_{j_1\dots j_t})(\eta_{i_{q+1}\dots i_k}\sqcup\!\sqcup\,\eta_{j_{t+1}\dots j_r}).\tag{2.20}$$

With this concept in hand, let us return to transformations of the endpoint map. Recall that the representation of the endpoint map in the form of a series of iterated integrals is unique due to Corollary 2.12. Hence, Remark 2.16 and (2.15) give the following description of the formal power series $\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m}$:

$$\begin{aligned}\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m}&=Q(\mathcal{E}_{X_1,\dots,X_m})=\sum_{q=1}^{\infty}\frac{1}{q!}Q^{(q)}(0)(\mathcal{E}_{X_1,\dots,X_m})^{\sqcup\!\sqcup\, q}\\&=\sum_{q=1}^{\infty}\sum_{j_1+\dots+j_n=q}\frac{1}{j_1!\cdots j_n!}\frac{\partial^{j_1+\dots+j_n}Q(0)}{\partial x_1^{j_1}\cdots\partial x_n^{j_n}}(\mathcal{E}_{X_1,\dots,X_m})_1^{\sqcup\!\sqcup\, j_1}\sqcup\!\sqcup\,\cdots\sqcup\!\sqcup\,(\mathcal{E}_{X_1,\dots,X_m})_n^{\sqcup\!\sqcup\, j_n},\end{aligned}\tag{2.21}$$

where the shuffle product of series is calculated termwise and $a^{\sqcup\!\sqcup\, q}$ denotes the shuffle q-power of a, that is, $a^{\sqcup\!\sqcup\, q}=a\sqcup\!\sqcup\,\cdots\sqcup\!\sqcup\, a$ (q times) for $q\ge 1$, $a^{\sqcup\!\sqcup\, 0}=1$. We will return to this representation later.

Here and further, when applying a real analytic transformation to a series of elements of $\mathcal{F}$, we mean that all polynomials are regarded as shuffle polynomials.

Example 2.21. Consider the system with two controls

$$\begin{aligned}\dot x_1&=u_1,\\ \dot x_2&=x_1u_2,\\ \dot x_3&=\tfrac{1}{6}x_1^3u_2.\end{aligned}\tag{2.22}$$

First, let us find the series representation (2.3) for the endpoint map $\mathcal{E}_{X_1,X_2}$. Since the system is feedforward (i.e., the kth component of X_i depends only on $x_1,\dots,x_{k-1}$), we can find this representation immediately, by integrating all these equations one by one. Taking into account that $x(0)=0$, we get

$$\begin{aligned} x_1(t) &= \int_0^t u_1(\tau)\,d\tau, \\ x_2(t) &= \int_0^t x_1(\tau_1)u_2(\tau_1)\,d\tau_1 = \int_0^t\int_0^{\tau_1} u_1(\tau_2)u_2(\tau_1)\,d\tau_2\,d\tau_1, \\ x_3(t) &= \frac{1}{6}\int_0^t x_1^3(\tau_1)u_2(\tau_1)\,d\tau_1 = \frac{1}{6}\int_0^t\Big(\int_0^{\tau_1} u_1(\tau_2)\,d\tau_2\Big)^3 u_2(\tau_1)\,d\tau_1 \\ &= \int_0^t\int_0^{\tau_1}\int_0^{\tau_2}\int_0^{\tau_3} u_1(\tau_2)u_1(\tau_3)u_1(\tau_4)u_2(\tau_1)\,d\tau_4\,d\tau_3\,d\tau_2\,d\tau_1. \end{aligned}$$

Taking into account the definition (2.4) of iterated integrals, we get

$$\mathcal{E}_{X_1,X_2} = \begin{pmatrix} \eta_1 \\ \eta_{21} \\ \eta_{2111} \end{pmatrix}.$$

Equivalently, it can be easily checked that all vectors (2.5) vanish except $c_1=e_1$, $c_{21}=e_2$, and $c_{2111}=e_3$.

Now, let us demonstrate how the series representation transforms under a change of variables. For example, consider

$$y = Q(x) = \begin{pmatrix} x_1 \\ x_2 - x_2^2 \\ x_3 \end{pmatrix}.$$

Then the series representation of the system in the new variables can be found directly, without finding the vector fields $\widehat{X}_1$ and $\widehat{X}_2$,

$$\mathcal{E}_{\widehat{X}_1,\widehat{X}_2} = Q(\mathcal{E}_{X_1,X_2}) = \begin{pmatrix} \eta_1 \\ \eta_{21} - \eta_{21}\sqcup\!\sqcup\,\eta_{21} \\ \eta_{2111} \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_{21} - 2\eta_{2121} - 4\eta_{2211} \\ \eta_{2111} \end{pmatrix}.$$

Let us write the system in the new variables. Obviously, $x_1=y_1$ and $x_3=y_3$. Let us find x_2 from the equation $y_2=x_2-x_2^2$. Since the change of variables maps a neighborhood of the origin to a neighborhood of the origin, we get $x_2=\frac{1}{2}(1-\sqrt{1-4y_2})$. Hence,

$$\begin{aligned} \dot y_1 &= u_1, \\ \dot y_2 &= y_1u_2 - 2y_1u_2\big(\tfrac{1}{2}(1-\sqrt{1-4y_2})\big) = y_1\sqrt{1-4y_2}\,u_2, \\ \dot y_3 &= \tfrac{1}{6}y_1^3u_2, \end{aligned}$$

that is,

$$\widehat{X}_1(y) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \widehat{X}_2(y) = \begin{pmatrix} 0 \\ y_1\sqrt{1-4y_2} \\ \frac{1}{6}y_1^3 \end{pmatrix}.$$

Then the form of $\mathcal{E}_{\widehat{X}_1,\widehat{X}_2}$ can be found by use of the vector fields $\widehat{X}_1(y)$ and $\widehat{X}_2(y)$; however, this way is much more complicated even for such a simple example.

Let us consider another change of variables:

$$y = Q(x) = \begin{pmatrix} 3x_1^5 - 25x_1^3 + 60x_1 \\ x_1 + x_2 \\ x_1x_2 - x_3 \end{pmatrix}.$$

Since the equation $y_1 = 3x_1^5 - 25x_1^3 + 60x_1$ is not solvable by radicals, $\widehat{X}_1(y)$ and $\widehat{X}_2(y)$ cannot be expressed explicitly (by radicals). Hence, we encounter some difficulties finding the series representation of $\mathcal{E}_{\widehat{X}_1,\widehat{X}_2}$ via $\widehat{X}_1$ and $\widehat{X}_2$. However, using the direct formula $\mathcal{E}_{\widehat{X}_1,\widehat{X}_2} = Q(\mathcal{E}_{X_1,X_2})$ we easily find that

$$\mathcal{E}_{\widehat{X}_1,\widehat{X}_2} = Q(\mathcal{E}_{X_1,X_2}) = \begin{pmatrix} 3\eta_1^{ш5} - 25\eta_1^{ш3} + 60\eta_1 \\ \eta_1 + \eta_{21} \\ \eta_1 ш \eta_{21} - \eta_{2111} \end{pmatrix} = \begin{pmatrix} 360\eta_{11111} - 150\eta_{111} + 60\eta_1 \\ \eta_1 + \eta_{21} \\ \eta_{121} + 2\eta_{211} - \eta_{2111} \end{pmatrix}.$$

2.5. An associative algebra of differential operators and a Lie algebra of vector fields. It is well known that any fixed set of m vector fields $X_1, \dots, X_m$ generates a (filtered) associative algebra of differential operators $F = \sum_{k=1}^{\infty} F^k$, where F^k is the linear span (over $\mathbb{R}$) of differential operators of order k of the form $X_{i_k} \cdots X_{i_1}$, $1 \le i_1, \dots, i_k \le m$, with composition being the algebraic product operation. Usually, such differential operators are supposed to act on (smooth or, in our case, real analytic) functions, that is, mappings from $U(0) \subset \mathbb{R}^n$ to $\mathbb{R}$. However, we prefer to define them as acting componentwise on vector functions, that is, mappings from $U(0) \subset \mathbb{R}^n$ to $\mathbb{R}^n$. In particular, the series coefficient (2.5) equals the value (at the origin) of the image of the identity map $E(x) = x$ under the corresponding differential operator from F.

Let us also consider the (filtered) Lie algebra of vector fields generated by the set $X_1, \dots, X_m$. It can be introduced as $L = \sum_{k=1}^{\infty} L^k$, where

$$L^1 = \mathrm{Lin}\{X_1, \dots, X_m\}$$

(the linear span is taken over $\mathbb{R}$) and L^k are defined recurrently by

$$L^{k+1} = [L^1, L^k], \qquad k \ge 1,$$

where $[\cdot,\cdot]$ denotes the Lie bracket of vector fields, $[X_i, X_j] = X_iX_j - X_jX_i$.

Let us now discuss the connections between the algebras F and L and the free algebras $\mathcal{F}$ and $\mathcal{L}$. Denote by φ the natural anti-homomorphism $\varphi : \mathcal{F} \to F$ defined by the rule

$$\varphi(\eta_{i_1 \dots i_k}) = X_{i_k} \cdots X_{i_1}, \qquad k \ge 1,\ 1 \le i_1, \dots, i_k \le m.$$

Then

$$\varphi(a_1a_2) = \varphi(a_2)\varphi(a_1) \quad \text{for any } a_1, a_2 \in \mathcal{F}.$$

Obviously, φ maps the free Lie algebra $\mathcal{L}$ to the Lie algebra L, and satisfies

$$\varphi([\ell_1, \ell_2]) = [\varphi(\ell_2), \varphi(\ell_1)] \quad \text{for any } \ell_1, \ell_2 \in \mathcal{L}.$$

Hence, the restriction of φ to $\mathcal{L}$ is an anti-homomorphism $\varphi : \mathcal{L} \to L$.

Let us also consider the linear map $c:\mathcal{F}\to\mathbb{R}^n$ defined as

$$c(a)=\varphi(a)E(0),\qquad a\in\mathcal{F}.$$

In other words, c is defined on basis elements by the formula

$$c(\eta_{i_1\dots i_k})=X_{i_k}\cdots X_{i_1}E(0)=c_{i_1\dots i_k},$$

where $c_{i_1\dots i_k}$ are the vector coefficients of $\eta_{i_1\dots i_k}$ in (2.13), and is extended to the whole algebra $\mathcal{F}$ by linearity. Then (2.13) can be rewritten in the form

$$\mathcal{E}_{X_1,\dots,X_m}=\sum_{k=1}^{\infty}\sum_{1\le i_1,\dots,i_k\le m}c(\eta_{i_1\dots i_k})\eta_{i_1\dots i_k}. \tag{2.23}$$

The subspace $\sum_{k=1}^{\infty}c(\mathcal{L}^k)\subset\mathbb{R}^n$ determines the dimension of the orbit of the system through the origin. In particular, the orbit is of full dimension iff the *Rashevsky–Chow condition* [45, 11]

$$\sum_{k=1}^{\infty}c(\mathcal{L}^k)=\mathbb{R}^n \tag{2.24}$$

holds. For control-linear systems like (2.1) this condition also implies local controllability; this means that any point from a certain neighborhood of the origin can be reached from any other point of this neighborhood.

Definition 2.22. A system of the form (2.1) that satisfies the Rashevsky–Chow condition (2.24) is called *bracket generating* (or *completely nonholonomic*).

Throughout the paper, we consider only bracket generating systems.

Definition 2.23. The minimal number p that guarantees the equality $\sum_{k=1}^{p}c(\mathcal{L}^k)=\mathbb{R}^n$ is called the *degree of nonholonomy*. Set

$$v_k=\dim c(\mathcal{L}^1\oplus\cdots\oplus\mathcal{L}^k),\qquad k=1,\dots,p\ (v_p=n). \tag{2.25}$$

The sequence $v=(v_1,\dots,v_p)$ is called the (*small*) *growth vector* of the system.

Both concepts, the degree of nonholonomy and the growth vector, are invariant under changes of variables and nonsingular feedbacks, and, in some way, describe the behavior of the system in a neighborhood of the origin. However, the precise description of the local behavior of the system is a more delicate question. Below we develop a technique which allows us to carry out such local analysis.

The anti-homomorphism φ (more specifically, the linear map c) induces special structures in the free Lie algebra $\mathcal{L}$. The simplest property is given by the following lemma.

Lemma 2.24. *$\mathrm{Ker}(c)\cap\mathcal{L}$ is a Lie subalgebra in $\mathcal{L}$.*

Proof. The proof is clear: Consider $\ell_1,\ell_2\in\mathrm{Ker}(c)\cap\mathcal{L}$, and denote $Y_i=\varphi(\ell_i)$, $i=1,2$. Then $c(\ell_i)=Y_i(0)=0$, $i=1,2$. This implies that

$$\begin{aligned}c([\ell_1,\ell_2])&=[\varphi(\ell_2),\varphi(\ell_1)]E(0)=Y_2Y_1E(0)-Y_1Y_2E(0)\\&=Y_1'(x)Y_2(x)|_{x=0}-Y_2'(x)Y_1(x)|_{x=0}=0.\ \blacksquare\end{aligned}$$

LEMMA 2.25. *If* $\ell \in \mathrm{Ker}(c) \cap \mathcal{L}$ *then* $(a\ell) \in \mathrm{Ker}(c)$ *for any* $a \in \mathcal{F}$.

Proof. It is sufficient to prove the statement for any element a of the form $a = \eta_{i_1\ldots i_k}$, where $k \geq 1$, $1 \leq i_1, \ldots, i_k \leq m$. Denote $Y = \varphi(\ell)$. Then $Y(0) = 0$, and therefore

$$c(\eta_{i_1\ldots i_k}\ell) = YX_{i_k}\cdots X_{i_1}E(0) = (X_{i_k}\cdots X_{i_1}E(x))'_x Y(x)|_{x=0} = 0. \blacksquare$$

Lemma 2.25 means that Ker(c) contains the left ideal generated by Ker$(c) \cap \mathcal{L}$, i.e.,

$$\mathrm{Lin}\big(\mathcal{F}^e(\mathrm{Ker}(c) \cap \mathcal{L})\big) \subset \mathrm{Ker}(c).$$

Below we obtain more precise properties using the filtered structures in $\mathcal{L}$ and L. Our main concept is introduced in the next subsection.

2.6. Core Lie subalgebra. Consider subspaces of $\mathcal{L}$ of the form

$$\mathcal{P}^k = \{\ell \in \mathcal{L}^k : c(\ell) \in c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k-1})\}, \qquad k \geq 1, \tag{2.26}$$

where for $k = 1$, $\mathcal{P}^1 = \{\ell \in \mathcal{L}^1 : c(\ell) = 0\}$, and set

$$\mathcal{L}_{X_1,\ldots,X_m} = \bigoplus_{k=1}^{\infty} \mathcal{P}^k. \tag{2.27}$$

LEMMA 2.26. $\mathcal{L}_{X_1,\ldots,X_m}$ *is a (graded) Lie subalgebra of* $\mathcal{L}$.

Proof. Let us show that $\mathcal{L}_{X_1,\ldots,X_m}$ is a Lie subalgebra. Obviously, it is sufficient to show that the Lie bracket of two homogeneous elements from $\mathcal{L}_{X_1,\ldots,X_m}$ belongs to $\mathcal{L}_{X_1,\ldots,X_m}$.

Suppose $\ell_i \in \mathcal{P}^{k_i}$, $i = 1, 2$. Then $c(\ell_i) \in c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k_i - 1})$. This means that there exist two elements $\ell_i' \in \mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k_i-1}$, $i = 1, 2$, such that $c(\ell_i) = c(\ell_i')$, i.e., $c(\ell_i - \ell_i') = 0$. Due to Lemma 2.24, $c([\ell_1 - \ell_1', \ell_2 - \ell_2']) = 0$. Hence,

$$\begin{aligned} c([\ell_1, \ell_2]) &= c([\ell_1 - \ell_1', \ell_2 - \ell_2']) + c\big([\ell_1', \ell_2] + [\ell_1, \ell_2'] - [\ell_1', \ell_2']\big) \\ &= c\big([\ell_1', \ell_2] + [\ell_1, \ell_2'] - [\ell_1', \ell_2']\big) \in c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k_1+k_2-1}), \end{aligned}$$

i.e., $[\ell_1, \ell_2] \in \mathcal{P}^{k_1+k_2}$. This implies that $\mathcal{L}_{X_1,\ldots,X_m}$ is a Lie subalgebra. It remains to note that $\mathcal{L}_{X_1,\ldots,X_m}$ is graded by definition. $\blacksquare$

So, to each control-linear system of the form (2.1) we assign the Lie subalgebra $\mathcal{L}_{X_1,\ldots,X_m}$.

LEMMA 2.27. *The Lie subalgebra* $\mathcal{L}_{X_1,\ldots,X_m}$ *is invariant with respect to nonsingular changes of variables in system* (2.1).

Proof. Suppose that a change of variables $y = Q(x)$ is applied so that $Q(0) = 0$ and $\det Q'(0) \neq 0$. Then the vector fields $X_1, \ldots, X_m$ in the new variables take the form $\widehat{X}_i(y) = Q'(x)X_i(x)|_{x=Q^{-1}(y)}$, $i = 1, \ldots, m$. Let us denote by $\widehat{c} : \mathcal{L} \to \mathbb{R}^n$ the linear map defined by $\widehat{c}(\eta_{i_1\ldots i_k}) = \widehat{X}_{i_k}\cdots\widehat{X}_{i_1}E(0)$. Then, as is well known, for any $\ell \in \mathcal{L}$ one has $\widehat{c}(\ell) = Q'(0)c(\ell)$.

Due to the definition of $\mathcal{P}^k = \mathcal{L}_{X_1,\ldots,X_m} \cap \mathcal{L}^k$,

$$\ell \in \mathcal{L}_{X_1,\ldots,X_m} \cap \mathcal{L}^k \text{ iff there exists } \ell' \in \mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k-1} \text{ such that } c(\ell - \ell') = 0.$$

Since $\widehat{c}(\ell-\ell') = Q'(0)c(\ell-\ell')$ and $\det Q'(0) \neq 0$, we get $\widehat{c}(\ell-\ell') = 0$ iff $c(\ell-\ell') = 0$. Hence, $\ell \in \mathcal{L}_{X_1,\dots,X_m} \cap \mathcal{L}^k$ iff $\ell \in \mathcal{L}_{\widehat{X}_1,\dots,\widehat{X}_m} \cap \mathcal{L}^k$, $k \geq 1$. This implies $\mathcal{L}_{X_1,\dots,X_m} = \mathcal{L}_{\widehat{X}_1,\dots,\widehat{X}_m}$. ■

Now we introduce one of the main concepts of the present paper.

Definition 2.28. We call the Lie subalgebra $\mathcal{L}_{X_1,\dots,X_m}$ defined by (2.26)–(2.27) the *core Lie subalgebra* corresponding to system (2.1).

The core Lie subalgebra $\mathcal{L}_{X_1,\dots,X_m}$ is intrinsic coordinate-independent object. Below we show that just this subalgebra is responsible for the homogeneous approximation of the system.

Let us explain the term "core Lie subalgebra". First, notice that the map $c : \mathcal{L} \to \mathbb{R}^n$ induces the filtration in $\mathbb{R}^n$ defined by $\mathbb{R}^n = \bigcup_{i=1}^p c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^i)$. Let us introduce the associated graded linear space. Namely, consider the factor subspaces $[c(\mathcal{L}^1)] = c(\mathcal{L}^1)$ and $[c(\mathcal{L}^i)] = c(\mathcal{L}^i)/c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{i-1})$, $i = 2,\dots,p$. Then the direct sum $V^n = [c(\mathcal{L}^1)] \oplus \cdots \oplus [c(\mathcal{L}^p)]$ is a graded linear space isomorphic to the initial filtered space $\mathbb{R}^n$. Now consider the induced graded linear map $g : \mathcal{L} \to V^n$ defined for $\ell \in \mathcal{L}^i$ by $g(\ell) = [c(\ell)]$ if $i = 1,\dots,p$, and by $g(\ell) = 0$ if $i \geq p+1$. Then $\mathcal{L}_{X_1,\dots,X_m}$ equals the core of g, i.e., $\mathcal{L}_{X_1,\dots,X_m} = \mathrm{Ker}(g)$. This implies that $\mathrm{Im}(g) = V^n$ is isomorphic to $\mathcal{L}/\mathrm{Ker}(g)$. In particular, this yields the following lemma.

Lemma 2.29. *The subspace $\mathcal{L}_{X_1,\dots,X_m}$ is of codimension n in the space $\mathcal{L}$.*

Proof. We give a proof that is independent of the discussion above.

For any $k \geq 1$, let us decompose $\mathcal{L}^k$ into a direct sum as $\mathcal{L}^k = \mathcal{P}^k \oplus \mathcal{M}^k$, where $\mathcal{M}^k$ is a complement subspace for $\mathcal{P}^k$. Notice that $\mathcal{M}^k = \{0\}$ for all $k \geq p+1$, where p is the degree of nonholonomy of the system. Hence,

$$\mathcal{L} = \mathcal{L}_{X_1,\dots,X_m} \oplus (\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^p).$$

Recall that by definition $c(\mathcal{P}^k) \subset c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k-1})$. It is easy to prove by induction that

$$c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^k) = c(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^k), \qquad k \geq 1.$$

Hence, $c(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^p) = c(\mathcal{L}) = \mathbb{R}^n$. It follows from the definition of $\mathcal{M}^k$ that

$$c(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^k) = c(\mathcal{M}^1) \oplus \cdots \oplus c(\mathcal{M}^k), \qquad k \geq 2,$$

and

$$\dim c(\mathcal{M}^k) = \dim \mathcal{M}^k, \qquad k \geq 1.$$

Hence,

$$\dim c(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^k) = \dim(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^k), \qquad k \geq 1.$$

Therefore,

$$\operatorname{codim} \mathcal{L}_{X_1,\dots,X_m} = \dim(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^p) = \dim c(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^p) = n. \ ■$$

Corollary 2.30. *If homogeneous elements $\ell_1,\dots,\ell_n \in \mathcal{L}$ are such that*

$$\mathcal{L} = \mathrm{Lin}\{\ell_1,\dots,\ell_n\} + \mathcal{L}_{X_1,\dots,X_m}$$

then the vectors $c(\ell_1),\dots,c(\ell_n)$ are linearly independent.

Proof. Due to Lemma 2.29, $\operatorname{codim} \mathcal{L}_{X_1,\dots,X_m} = n$. Hence, the assumption of the lemma implies $\mathcal{L} = \mathrm{Lin}\{\ell_1,\dots,\ell_n\} \oplus \mathcal{L}_{X_1,\dots,X_m}$ and $\dim \mathrm{Lin}\{\ell_1,\dots,\ell_n\} = n$. For any $k \geq 1$, set $\mathcal{M}^k = \mathrm{Lin}\{\ell_1, \ldots, \ell_n\} \cap \mathcal{L}^k$. Since $\ell_1, \ldots, \ell_n$ are homogeneous, $\mathrm{Lin}\{\ell_1, \ldots, \ell_n\} = \bigoplus_{k=1}^{\infty} \mathcal{M}^k$, therefore $\mathcal{L}^k = \mathcal{P}^k \oplus \mathcal{M}^k$ for any $k \geq 1$. Moreover, there exists p such that $\mathcal{M}^k = \{0\}$ for all $k \geq p+1$. Hence, $\mathrm{Lin}\{\ell_1, \ldots, \ell_n\} = \mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^p$. Similarly to the proof of Lemma 2.29, we have

$$\begin{aligned}\dim \mathrm{Lin}\{c(\ell_1), \ldots, c(\ell_n)\} &= \dim c(\mathrm{Lin}\{\ell_1, \ldots, \ell_n\}) = \dim c(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^p) \\ &= \dim(\mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^p) = \dim \mathrm{Lin}\{\ell_1, \ldots, \ell_n\} = n. \ \blacksquare\end{aligned}$$

EXAMPLE 2.31. Let us return to system (2.22) from Example 2.21. We have

$$c(\eta_1) = c_1 = e_1 \neq 0, \quad c(\eta_2) = c_2 = 0, \quad c([\eta_2, \eta_1]) = c_{21} - c_{12} = e_2 \notin \mathrm{Lin}\{e_1\},$$
$$c([[\eta_2, \eta_1], \eta_1]) = c_{211} - 2c_{121} + c_{112} = 0, \quad c([[\eta_2, \eta_1], \eta_2]) = 2c_{212} - c_{122} - c_{221} = 0,$$
$$c([[[\eta_2, \eta_1], \eta_1], \eta_1]) = c_{2111} - 3\eta_{1211} + 3c_{1121} - c_{1112} = e_3 \notin \mathrm{Lin}\{e_1, e_2\},$$

and all other brackets vanish. Hence, the degree of nonholonomy equals $p = 4$, and the growth vector equals $v = (1, 2, 2, 3)$.

Now, let us find the core Lie subalgebra $\mathcal{L}_{X_1,X_2}$. Since

$$\begin{gathered}\mathcal{P}^1 = \mathrm{Lin}\{\eta_2\}, \quad \mathcal{P}^2 = \{0\}, \quad \mathcal{P}^3 = \mathrm{Lin}\big\{[[\eta_2, \eta_1], \eta_1], [[\eta_2, \eta_1], \eta_2]\big\} = \mathcal{L}^3, \\ \mathcal{P}^4 = \mathrm{Lin}\big\{[[[\eta_2, \eta_1], \eta_1], \eta_2], [[[\eta_2, \eta_1], \eta_2], \eta_2]\big\},\end{gathered} \tag{2.28}$$

and $\mathcal{P}^k = \mathcal{L}^k$ for $k \geq 5$, we have $\mathcal{L}_{X_1,X_2} = \sum_{k=1}^{\infty} \mathcal{P}^k$. Obviously, $\mathcal{L}_{X_1,X_2}$ is a subalgebra and $\mathrm{codim}\, \mathcal{L}_{X_1,X_2} = 3$. Let us find three homogeneous elements that define a complement of $\mathcal{L}_{X_1,X_2}$. For example, we may choose

$$\ell_1 = \eta_1, \quad \ell_2 = -2[\eta_2, \eta_1], \quad \ell_3 = 3[[[\eta_2, \eta_1], \eta_1], \eta_1] - [[[\eta_2, \eta_1], \eta_2], \eta_2]. \tag{2.29}$$

Then $\mathcal{L} = \mathrm{Lin}\{\ell_1, \ell_2, \ell_3\} + \mathcal{L}_{X_1,X_2}$. Notice that the vectors $c(\ell_1) = e_1$, $c(\ell_2) = -2e_2$, and $c(\ell_3) = 3e_3$ are linearly independent.

3. Homogeneous approximation, nonholonomic derivatives, weights, and privileged coordinates from the algebraic viewpoint

3.1. Definition of a homogeneous approximation. The concept of homogeneous approximation plays an important role in nonlinear control theory [12, 23, 24, 6, 3, 8, 4]. Though it can be introduced in a coordinate-free manner, the most clear definitions include some special "privileged" coordinates, in which the two systems—the initial and approximating ones—can be effectively compared.

Let us introduce homogeneous approximations in terms of the endpoint map.

DEFINITION 3.1. Consider a bracket generating control-linear system of the form (2.1). A bracket generating control-linear system

$$\dot{z} = \sum_{i=1}^{m} u_i Z_i(z), \qquad z \in U(0) \subset \mathbb{R}^n,\ u_1, \ldots, u_m \in \mathbb{R}, \tag{3.1}$$

with real analytic vector fields $Z_1(z), \ldots, Z_m(z)$ is called a *homogeneous approximation* for the initial system if

(i) its endpoint map $\mathcal{E}_{Z_1,\ldots,Z_m}$ is homogeneous,

$$\mathcal{E}_{Z_1,\ldots,Z_m}(\theta,u^{1/\theta}) = H_\theta(\mathcal{E}_{Z_1,\ldots,Z_m}(1,u)) \quad \text{for any } \theta>0,\ u\in B^1,$$

where H_θ is a dilation defined by $H_\theta(z) = (\theta^{w_1}z_1,\ldots,\theta^{w_n}z_n)$, and $1\le w_1\le\cdots\le w_n$ are some integers;

(ii) there is a real analytic change of variables $y=Q(x)$ in the initial system ($Q(0)=0$, $\det Q'(0)\neq 0$) such that $\mathcal{E}_{Z_1,\ldots,Z_m}$ approximates the endpoint map of the initial system in the new coordinates; namely, for any $u\in B^1$,

$$H_\theta^{-1}\big(Q(\mathcal{E}_{X_1,\ldots,X_m}(\theta,u^{1/\theta}))-\mathcal{E}_{Z_1,\ldots,Z_m}(\theta,u^{1/\theta})\big)\to 0 \quad \text{as } \theta\to 0.$$

In this section we examine some concepts that are encountered when studying homogeneous approximations [6]. Our analysis is based on the series approach and the free algebras introduced above.

3.2. Nonholonomic derivatives and the order of functions. Suppose a bracket generating control-linear system of the form (2.1) is fixed. Following [6], let us say that the differential operators of the first order $X_1,\ldots,X_m$ are *nonholonomic derivatives of the first order*. Then any operator $X_{i_k}\cdots X_{i_1}$ is naturally considered as a nonholonomic derivative of the kth order, $1\le i_1,\ldots,i_k\le m$, $k\ge 1$. Nonholonomic derivatives are used in the following definition.

Suppose a real analytic function $f=f(x):U(0)\to\mathbb{R}$ is given. The number s is called the *order* of the function $f=f(x)$ at the point $x=0$ if

(i) $X_{i_k}\cdots X_{i_1}f(0)=0$ for all $k\le s-1$ and all $1\le i_1,\ldots,i_k\le m$;
(ii) $X_{j_s}\cdots X_{j_1}f(0)\neq 0$ for a certain set $1\le j_1,\ldots,j_s\le m$.

In the case of the coordinate functions $f_i(x)=x_i$, $i=1,\ldots,n$, this definition can be reformulated by use of the set of the vectors (2.5). Namely, the order of the function $f_i(x)=x_i$ coincides with the minimal k such that $(c_{j_1\ldots j_k})_i\neq 0$ for a certain set $1\le j_1,\ldots,j_k\le m$.

This can also be expressed in terms of the series representation of $\mathcal{E}_{X_1,\ldots,X_m}$. Namely, the order of the coordinate function $f_i(x)=x_i$ coincides with the minimal order of an iterated integral (2.4) entering the ith component of the right hand side of (2.3) with a nonzero coefficient.

3.3. Weight of coordinates. Recall that we consider the bracket generating system. Let v be its growth vector (2.25); for convenience, set $v_0=0$.

Suppose the coordinates are chosen so that $c(\mathcal{L}^1\oplus\cdots\oplus\mathcal{L}^i)=\mathrm{Lin}\{e_1,\ldots,e_{v_i}\}$, $i=1,\ldots,p$. This can be achieved by a certain linear nonsingular change of variables in the initial system; such coordinates are called *linearly adapted* [6].

Let us recall the following definition, which is suitable for linearly adapted coordinates. The minimal number w_i such that $e_i\in c(\mathcal{L}^1\oplus\cdots\oplus\mathcal{L}^{w_i})$ is called the *weight of the coordinate* x_i, $i=1,\ldots,n$. In other words, the weight of x_i coincides with the minimal order w_i of a homogeneous Lie element $\ell\in\mathcal{L}^{w_i}$ such that $(c(\ell))_i\neq 0$, $i=1,\ldots,n$.

It is worth noting that the sequence of the weights $\{w_1, \ldots, w_n\}$ is the same for all linearly adapted coordinates.

3.4. Privileged coordinates and R. Ree's theorem. Following [6], we say that linearly adapted coordinates $x_1, \ldots, x_n$ are *privileged* if the order of any coordinate function $f_i(x) = x_i$, $i = 1, \ldots, n$, coincides with the weight of this coordinate. It is proved in [6] that one can construct privileged coordinates by a certain polynomial change of variables. Moreover, in such coordinates a homogeneous approximation of the initial system can be easily constructed.

Our next goal is to express the definition of privileged coordinates in terms of the map c. To this end, we use the result of the remarkable paper of R. Ree [46], namely the theorem on a connection of the Lie algebra and the shuffle product.

Definition 3.2. Define the *inner product operation* $\langle \cdot, \cdot \rangle$ in $\mathcal{F}$, assuming the basis

$$\{\eta_{i_1 \ldots i_k} : k \geq 1,\ 1 \leq i_1, \ldots, i_k \leq m\}$$

is orthonormal, i.e.,

$$\langle \eta_{i_1 \ldots i_k}, \eta_{j_1 \ldots j_s} \rangle = \begin{cases} 1 & \text{if } k = s,\ i_q = j_q,\ q = 1, \ldots, k, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the subspaces $\mathcal{F}^k$ are orthogonal to each other, hence the sums of subspaces like (2.27) are orthogonal. Below we also use the symbol $\oplus^\perp$, which denotes the orthogonal sum. However, to avoid cumbersome notation, for direct sums of homogeneous subspaces we keep the symbol $\oplus$.

Theorem 3.3 (R. Ree [46]). *An element of $\mathcal{F}$ belongs to the Lie algebra $\mathcal{L}$ if and only if it is orthogonal to the shuffle product of any two elements of $\mathcal{F}$,*

$$\ell \in \mathcal{L} \text{ iff } \langle \ell, a_1 \sqcup\!\sqcup\, a_2 \rangle = 0 \text{ for any } a_1, a_2 \in \mathcal{F}.$$

In other words, Ree's theorem says that

$$\mathcal{L} = (\mathcal{F} \sqcup\!\sqcup\, \mathcal{F})^\perp,$$

where the symbol $^\perp$ denotes the orthogonal complement. Hence,

$$\mathcal{F} = \mathcal{L} \oplus^\perp \mathrm{Lin}\{\mathcal{F} \sqcup\!\sqcup\, \mathcal{F}\},$$

where the symbol $\oplus^\perp$ denotes the orthogonal sum. Since the subspaces $\mathcal{F}^k$ are orthogonal to each other, for any homogeneous subspace we get the decomposition

$$\mathcal{F}^k = \mathcal{L}^k \oplus^\perp \mathrm{Lin}\{\mathcal{F}^i \sqcup\!\sqcup\, \mathcal{F}^{k-i} : i = 1, \ldots, k-1\}, \qquad k \geq 1.$$

It is easy to prove by induction that

$$\mathcal{F}^k = \mathcal{L}^k \oplus^\perp \mathrm{Lin}\{\mathcal{L}^{i_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup\, \mathcal{L}^{i_q} : q \geq 2,\ i_1 + \cdots + i_q = k,\ i_1, \ldots, i_q \geq 1\}, \qquad k \geq 1.$$

It is convenient to write this decomposition in the form

$$\mathcal{F}^k = \mathcal{L}^k \oplus^\perp (\mathcal{L}^{\mathrm{sh}} \cap \mathcal{F}^k), \quad k \geq 1, \tag{3.2}$$

where

$$\mathcal{L}^{\mathrm{sh}} = \mathrm{Lin}\{z_1 \sqcup\!\sqcup \cdots \sqcup\!\sqcup\, z_q : q \geq 2,\ z_1, \ldots, z_q \in \mathcal{L}\},$$

or briefly,

$$\mathcal{F} = \mathcal{L} \oplus^{\perp} \mathcal{L}^{\mathrm{sh}}. \tag{3.3}$$

Now, for any $k \geq 1$ consider an orthonormal basis B_k of the subspace $\mathcal{L}^k$,

$$B_k = \{b_{k,j} : j = 1, \dots, d_k\}, \qquad d_k = \dim \mathcal{L}^k,$$

and an orthonormal basis $\widehat{B}_k$ of $\mathcal{L}^{\mathrm{sh}} \cap \mathcal{F}^k$,

$$\widehat{B}_k = \{\widehat{b}_{k,j} : j = 1, \dots, \widehat{d}_k\}, \qquad \widehat{d}_k = \dim(\mathcal{L}^{\mathrm{sh}} \cap \mathcal{F}^k) = \dim(\mathcal{F}^k) - \dim(\mathcal{L}^k) = m^k - d_k.$$

Then the set $\bigcup_{k\geq 1}(B_k \cup \widehat{B}_k)$ is an orthonormal basis of $\mathcal{F}$. Hence, the series on the right hand side of (2.23) can be re-expanded in this orthonormal basis, which gives

$$\mathcal{E}_{X_1,\dots,X_m} = \sum_{k=1}^{\infty}\Big(\sum_{j=1}^{d_k} c(b_{k,j})b_{k,j} + \sum_{j=1}^{\widehat{d}_k} c(\widehat{b}_{k,j})\widehat{b}_{k,j}\Big). \tag{3.4}$$

This representation leads to the following reformulation of the definitions of the order and weight of coordinates.

The *order* of the coordinate function $f_i(x) = x_i$ equals the minimal order k_i of a basis element $b_{k_i,j}$ or $\widehat{b}_{k_i,j}$ entering the ith component of the right hand side of (3.4) with a nonzero coefficient, i.e., such that $(c(b_{k_i,j}))_i \neq 0$ or $(c(\widehat{b}_{k_i,j}))_i \neq 0$.

Suppose the coordinates are linearly adapted. The *weight* of the coordinate x_i equals the minimal order w_i of a basis element $b_{w_i,j}$ entering the ith component of the right hand side of (3.4) with a nonzero coefficient, i.e., such that $(c(b_{w_i,j}))_i \neq 0$.

In particular, it is clear that the order of a coordinate function is less than or equal to the weight of this coordinate. Moreover, the order of $f_i(x) = x_i$ is strictly less than the weight of x_i if and only if there exists a basis element $\widehat{b}_{k_i,j}$ such that $(c(\widehat{b}_{k_i,j}))_i \neq 0$ and $k_i < w_i$. Thus, linearly adapted coordinates are *privileged* if for any $i = 1, \dots, n$,

$$(c(\widehat{b}_{k,j}))_i = 0 \qquad \text{for all } k < w_i \text{ and } j = 1, \dots, \widehat{d}_k.$$

Therefore, one can try to construct privileged coordinates excluding all $\widehat{b}_{k,j}$ such that $1 \leq k < w_i$, $j = 1, \dots, \widehat{d}_k$, from the ith component of (3.4). However, the bases B_k and $\widehat{B}_k$ are inconvenient for the practical implementation of this idea, since they do not involve any information about the concrete system of the form (2.1). In the next section we give another basis which is suitable in this situation.

4. The left ideal and dual basis in the associative algebra

4.1. The left ideal generated by a system. In this subsection we introduce the concept that, along with the core Lie subalgebra, plays the central role in our constructions.

Definition 4.1. We call the subspace

$$\mathcal{J}_{X_1,\dots,X_m} = \mathrm{Lin}\{\mathcal{F}^e\mathcal{L}_{X_1,\dots,X_m}\} = \mathrm{Lin}\{a\ell : a \in \mathcal{F}^e,\ \ell \in \mathcal{L}_{X_1,\dots,X_m}\}$$

the *left ideal* corresponding to system (2.1).

Notice that, due to its definition, the left ideal $\mathcal{J}_{X_1,\dots,X_m}$ is graded, i.e.,

$$\mathcal{J}_{X_1,\dots,X_m} = \bigoplus_{k=1}^{\infty}(\mathcal{J}_{X_1,\dots,X_m} \cap \mathcal{F}^k). \tag{4.1}$$

Moreover, it is invariant with respect to nonsingular changes of variables in the system, which follows directly from Lemma 2.27.

Lemma 4.2. *If $a \in \mathcal{J}_{X_1,\dots,X_m} \cap \mathcal{F}^k$ then $c(a) \in c(\mathcal{F}^1 \oplus \dots \oplus \mathcal{F}^{k-1})$.*

Proof. Without loss of generality assume $a = \eta_{j_1\dots j_q}\ell$, where $\ell \in \mathcal{P}^s$, $q + s = k$, $q \geq 0$, and $s \geq 1$ (if $q = 0$ then $a = \ell$). Since $\ell \in \mathcal{P}^s$, there exists $\ell' \in \mathcal{L}^1 \oplus \dots \oplus \mathcal{L}^{s-1}$ such that $c(\ell - \ell') = 0$. Hence, due to Lemma 2.25, $c(\eta_{j_1\dots j_q}(\ell - \ell')) = 0$, which implies $c(a) = c(\eta_{j_1\dots j_q}\ell) = c(\eta_{j_1\dots j_q}\ell') \in c(\mathcal{F}^1 \oplus \dots \oplus \mathcal{F}^{k-1})$.

Notice that for $a \in \mathcal{J}_{X_1,\dots,X_m} \cap \mathcal{F}^1$ the lemma means $c(a) = 0$. ∎

Due to Lemma 2.29, the core Lie subalgebra $\mathcal{L}_{X_1,\dots,X_m}$ is of codimension n in $\mathcal{L}$. Let us fix an arbitrary set $\{\ell_1, \dots, \ell_n\}$ of homogeneous elements of $\mathcal{L}$ such that

$$\mathcal{L} = \mathrm{Lin}\{\ell_1, \dots, \ell_n\} + \mathcal{L}_{X_1,\dots,X_m}. \tag{4.2}$$

Due to Lemma 2.29, this sum is direct. Without loss of generality assume

$$\mathrm{ord}(\ell_i) \leq \mathrm{ord}(\ell_j) \quad \text{if } 1 \leq i < j \leq n. \tag{4.3}$$

It is worth noting that the orders of elements $\ell_1, \dots, \ell_n$ satisfying (4.3) are defined uniquely, since the number of such elements of order $k \geq 1$ equals $\dim(\mathcal{L}^k) - \dim(\mathcal{P}^k)$.

Denote by $\{\ell_j\}_{j=n+1}^{\infty}$ a homogeneous basis of $\mathcal{L}_{X_1,\dots,X_m}$. Then $\{\ell_j\}_{j=1}^{\infty}$ is a (homogeneous) basis of $\mathcal{L}$.

Now we are going to use the well-known Poincaré–Birkhoff–Witt theorem [47], which says that the set

$$\{\ell_{j_1} \cdots \ell_{j_r} : 1 \leq j_1 \leq \dots \leq j_r,\ r \geq 1\} \tag{4.4}$$

forms a basis of $\mathcal{F}$.

Lemma 4.3. *The set*

$$\{\ell_{j_1} \cdots \ell_{j_r} : n+1 \leq j_1 \leq \dots \leq j_r,\ r \geq 1\} \tag{4.5}$$

forms a basis of the subalgebra

$$M = \mathrm{Lin}\{\ell_{i_1} \cdots \ell_{i_k} : i_1, \dots, i_k \geq n+1,\ k \geq 1\}. \tag{4.6}$$

Proof. Let us prove that any element of the form

$$x = \ell_{q_1} \cdots \ell_{q_s}, \qquad q_1, \dots, q_s \geq n+1,$$

equals a linear combination of elements (4.5).

For $s = 1$, there is nothing to prove. Suppose $s \geq 2$ and introduce the following definition. We say that $\ell_{q_{p_1}}$ and $\ell_{q_{p_2}}$ form an *inversion* in x if $p_1 < p_2$ and $q_{p_1} > q_{p_2}$. Let d be the number of inversions in the element x. Then $d \leq s(s-1)/2$. We say that the pair (s, d) is the *disorder* of the element x.

If the disorder of x equals $(s,0)$ then x belongs to the set (4.5). Suppose the disorder of x equals (s,d) with $d>0$. Then for a certain $1\le i\le s-1$ one has $q_i>q_{i+1}$. Thus,

$$\ell_{q_i}\ell_{q_{i+1}}=[\ell_{q_i},\ell_{q_{i+1}}]+\ell_{q_{i+1}}\ell_{q_i}.$$

By definition, $q_i\ge n+1$ and $q_{i+1}\ge n+1$. Hence, the elements ℓ_{q_i} and $\ell_{q_{i+1}}$ belong to the Lie subalgebra $\mathcal{L}_{X_1,\ldots,X_m}$, which gives $[\ell_{q_i},\ell_{q_{i+1}}]\in\mathcal{L}_{X_1,\ldots,X_m}$. Therefore, $[\ell_{q_i},\ell_{q_{i+1}}]$ can be represented as a linear combination of elements ℓ_j with $j\ge n+1$, i.e.,

$$\ell_{q_i}\ell_{q_{i+1}}=\sum_{j\ge n+1}\beta_j\ell_j+\ell_{q_{i+1}}\ell_{q_i},\qquad \beta_j\in\mathbb{R}.$$

Denoting

$$y_j=\ell_{q_1}\cdots\ell_{q_{i-1}}\ell_j\,\ell_{q_{i+2}}\cdots\ell_{q_s},\quad z=\ell_{q_1}\cdots\ell_{q_{i-1}}\ell_{q_{i+1}}\ell_{q_i}\ell_{q_{i+2}}\cdots\ell_{q_s},$$

we get

$$x=\sum_{j\ge n+1}\beta_jy_j+z,$$

where y_j and z belong to (4.6), and moreover y_j are of disorder $(s-1,d_j)$ and z is of disorder (s,d') with $d'<d$. Thus, x is represented as a sum of terms from (4.6) each of which has the disorder smaller (in the lexicographic sense) than the disorder of x. Obviously, after a finite number of such steps the element x is reduced to a sum of elements from (4.6) whose disorders equal $(k_i,0)$, $k_i\ge1$. It was noticed above that all such terms are of the form (4.5). Hence, x equals a linear combination of elements (4.5).

Thus, any element from (4.6) is a linear combination of elements (4.5). On the other hand, elements (4.5) are linearly independent, since they belong to the Poincaré–Birkhoff–Witt basis (4.4). Therefore, they form a basis of M. ■

Corollary 4.4. *The set*

$$\{\ell_{j_1}\cdots\ell_{j_r}:1\le j_1\le\cdots\le j_r,\ r\ge1,\ j_r\ge n+1\}\tag{4.7}$$

forms a basis of the left ideal $\mathcal{J}_{X_1,\ldots,X_m}$.

Proof. Obviously, it is sufficient to prove that any element of the form $a\ell_i$, where $a\in\mathcal{F}$ and $i\ge n+1$, can be represented uniquely as a linear combination of elements (4.7). Since a can be expressed via the Poincaré–Birkhoff–Witt basis (4.4), it is sufficient to prove this fact for any element of the form

$$z=(\ell_{j_1}\cdots\ell_{j_h})(\ell_{j_{h+1}}\cdots\ell_{j_s})\ell_i,$$

where $j_1\le\cdots\le j_h\le n<j_{h+1}\le\cdots\le j_s$, $i\ge n+1$.

For $s=h$, there is nothing to prove. Consider the case $s\ge h+1$. Due to Lemma 4.3, the element $(\ell_{j_{h+1}}\cdots\ell_{j_s})\ell_i$ is a linear combination of elements (4.5). Hence, z is a linear combination of elements of the form

$$(\ell_{j_1}\cdots\ell_{j_h})(\ell_{i_1}\cdots\ell_{i_k}),$$

where $j_1\le\cdots\le j_h\le n$, $n+1\le i_1\le\cdots\le i_k$, and $k\ge1$.

Thus, any element of $\mathcal{J}_{X_1,\ldots,X_m}$ is a linear combination of elements (4.7). On the other hand, elements (4.7) are linearly independent, since they belong to the Poincaré–Birkhoff–Witt basis (4.4). Therefore, they form a basis of the left ideal $\mathcal{J}_{X_1,\ldots,X_m}$. ■

Corollary 4.5. *For any* $k \geq 1$,

$$\mathcal{J}_{X_1,\ldots,X_m} \cap \mathcal{L}^k = \mathcal{P}^k,$$

and therefore

$$\mathcal{J}_{X_1,\ldots,X_m} \cap \mathcal{L} = \mathcal{L}_{X_1,\ldots,X_m}.$$

Proof. The inclusion $\mathcal{P}^k \subset \mathcal{J}_{X_1,\ldots,X_m} \cap \mathcal{L}^k$ follows from the definition. Let us show that $\mathcal{J}_{X_1,\ldots,X_m} \cap \mathcal{L}^k \subset \mathcal{P}^k$.

Due to Corollary 4.4, any element $a \in \mathcal{J}_{X_1,\ldots,X_m}$ can be expressed as a linear combination of elements of (4.7). On the other hand, a basis of $\mathcal{L}$ is given by the elements $\{\ell_j\}_{j=1}^{\infty}$ that belong to the Poincaré–Birkhoff–Witt basis (4.4).

If $a \in \mathcal{J}_{X_1,\ldots,X_m} \cap \mathcal{L}$ then it is a linear combination of elements from the intersection of the sets (4.7) and $\{\ell_j\}_{j=1}^{\infty}$, which equals $\{\ell_j\}_{j=n+1}^{\infty}$. Obviously, $a \in \mathcal{J}_{X_1,\ldots,X_m} \cap \mathcal{L}^k$ is a linear combination of elements $\{\ell_j\}_{j=n+1}^{\infty} \cap \mathcal{L}^k \subset \mathcal{P}^k$. ∎

As a consequence, two structures induced by the control system, namely $\mathcal{L}_{X_1,\ldots,X_m}$ and $\mathcal{J}_{X_1,\ldots,X_m}$, define each other uniquely.

4.2. Orthogonal complement to the left ideal and a generalization of R. Ree's theorem. It turns out that, in the homogeneous approximation problem, an important role is played by the orthogonal complement of the left ideal $\mathcal{J}_{X_1,\ldots,X_m}$, i.e.,

$$\mathcal{J}^{\perp}_{X_1,\ldots,X_m} = \{x \in \mathcal{F} : \langle x, a\rangle = 0 \text{ for any } a \in \mathcal{J}_{X_1,\ldots,X_m}\}.$$

Note that (4.1) implies

$$\mathcal{J}^{\perp}_{X_1,\ldots,X_m} = \bigoplus_{k=1}^{\infty} (\mathcal{J}^{\perp}_{X_1,\ldots,X_m} \cap \mathcal{F}^k). \tag{4.8}$$

In this subsection we study properties of $\mathcal{J}^{\perp}_{X_1,\ldots,X_m}$.

Lemma 4.6. *Suppose* $x = \sum_{i_1,\ldots,i_k} \gamma_{i_1\ldots i_k}\eta_{i_1\ldots i_k}$, *where* $\gamma_{i_1\ldots i_k} \in \mathbb{R}$. *Then* $x \in \mathcal{J}^{\perp}_{X_1,\ldots,X_m}$ *iff* $\sum_{i_{s+1}\ldots i_k} \gamma_{i_1\ldots i_k}\eta_{i_{s+1}\ldots i_k} \perp \mathcal{P}^{k-s}$ *for any* $s = 0, \ldots, k-1$ *and any fixed set of indices* $i_1, \ldots, i_s$.

Proof. The proof follows immediately from the definitions. In fact, $x \in \mathcal{J}^{\perp}_{X_1,\ldots,X_m}$ iff x is orthogonal to any element of the form $\eta_{i_1^0\ldots i_s^0}\,\ell$, where $0 \leq s \leq k-1$ and $\ell \in \mathcal{P}^{k-s}$, i.e.,

$$\begin{aligned}\langle x, \eta_{i_1^0\ldots i_s^0}\ell\rangle &= \Big\langle \sum_{i_1,\ldots,i_k} \gamma_{i_1\ldots i_k}\eta_{i_1\ldots i_s}\eta_{i_{s+1}\ldots i_k}, \eta_{i_1^0\ldots i_s^0}\ell\Big\rangle \\ &= \Big\langle \sum_{i_{s+1},\ldots,i_k} \gamma_{i_1^0\ldots i_s^0 i_{s+1}\ldots i_k}\eta_{i_{s+1}\ldots i_k}, \ell\Big\rangle = 0,\end{aligned}$$

which proves the lemma. ∎

Lemma 4.7. *Suppose* $a, b \in \mathcal{J}^{\perp}_{X_1,\ldots,X_m}$. *Then* $a \sqcup\!\sqcup b \in \mathcal{J}^{\perp}_{X_1,\ldots,X_m}$.

Proof. Due to (4.8), it is sufficient to prove the lemma for $a \in \mathcal{F}^k$ and $b \in \mathcal{F}^r$ for arbitrary $k, r \geq 1$. Let $a = \sum_{i_1,\ldots,i_k} \alpha_{i_1\ldots i_k}\eta_{i_1\ldots i_k}$ and $b = \sum_{j_1,\ldots,j_r} \beta_{j_1\ldots j_r}\eta_{j_1\ldots j_r}$.

It is sufficient to prove that $a\sqcup b$ is orthogonal to any element of the form $x\ell$, where $x\in\mathcal{F}^s$ and $\ell\in\mathcal{P}^{k+r-s}$, $0\le s\le k+r-1$. Using Lemma 2.20, we get

$$\begin{aligned}
a\sqcup b &= \sum_{\substack{i_1,\dots,i_k\\ j_1,\dots,j_r}} \alpha_{i_1\dots i_k}\beta_{j_1\dots j_r}\eta_{i_1\dots i_k}\sqcup\eta_{j_1\dots j_r}\\
&= \sum_{\substack{i_1,\dots,i_k\\ j_1,\dots,j_r}}\ \sum_{\substack{0\le q\le k,\,0\le t\le r\\ q+t=s}} \alpha_{i_1\dots i_k}\beta_{j_1\dots j_r}(\eta_{i_1\dots i_q}\sqcup\eta_{j_1\dots j_t})(\eta_{i_{q+1}\dots i_k}\sqcup\eta_{j_{t+1}\dots j_r})\\
&= \sum_{\substack{0\le q\le k,\,0\le t\le r\\ q+t=s}}\ \sum_{\substack{i_1,\dots,i_q\\ j_1,\dots,j_t}} (\eta_{i_1\dots i_q}\sqcup\eta_{j_1\dots j_t}) \sum_{\substack{i_{q+1},\dots,i_k\\ j_{t+1},\dots,j_r}} \alpha_{i_1\dots i_k}\beta_{j_1\dots j_r}(\eta_{i_{q+1}\dots i_k}\sqcup\eta_{j_{t+1}\dots j_r}).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\langle x\ell, a\sqcup b\rangle\\
&\quad= \sum_{\substack{0\le q\le k,\,0\le t\le r\\ q+t=s}}\ \sum_{\substack{i_1,\dots,i_q\\ j_1,\dots,j_t}} \Big\langle x\,\ell, (\eta_{i_1\dots i_q}\sqcup\eta_{j_1\dots j_t}) \sum_{\substack{i_{q+1},\dots,i_k\\ j_{t+1},\dots,j_r}} \alpha_{i_1\dots i_k}\beta_{j_1\dots j_r}(\eta_{i_{q+1}\dots i_k}\sqcup\eta_{j_{t+1}\dots j_r})\Big\rangle\\
&\quad= \sum_{\substack{0\le q\le k,\,0\le t\le r\\ q+t=s}}\ \sum_{\substack{i_1,\dots,i_q\\ j_1,\dots,j_t}} \langle x, \eta_{i_{q+1}\dots i_k}\sqcup\eta_{j_{t+1}\dots j_r}\rangle \Big\langle \ell, \sum_{\substack{i_{q+1},\dots,i_k\\ j_{t+1},\dots,j_r}} \alpha_{i_1\dots i_k}\beta_{j_1\dots j_r}(\eta_{i_{q+1}\dots i_k}\sqcup\eta_{j_{t+1}\dots j_r})\Big\rangle.
\end{aligned} \tag{4.9}$$

Consider each term of (4.9).

If $q<k$ and $t<r$, then $\eta_{i_{q+1}\dots i_k}\in\mathcal{F}$ and $\eta_{j_{t+1}\dots j_r}\in\mathcal{F}$. Hence, due to R. Ree's theorem, $\langle \ell, \eta_{i_{q+1}\dots i_k}\sqcup\eta_{j_{t+1}\dots j_r}\rangle=0$.

If $t=r$ then $\eta_{j_{t+1}\dots j_r}=1$. Since $q+t=s$, we get $q=s-r\le k-1$. Hence in this case,

$$\Big\langle \ell, \sum_{\substack{i_{q+1},\dots,i_k\\ j_{t+1},\dots,j_r}} \alpha_{i_1\dots i_k}\beta_{j_1\dots j_r}(\eta_{i_{q+1}\dots i_k}\sqcup\eta_{j_{t+1}\dots j_r})\Big\rangle = \beta_{j_1\dots j_r}\Big\langle \ell, \sum_{i_{s-r+1},\dots,i_k} \alpha_{i_1\dots i_k}\eta_{i_{s-r+1}\dots i_k}\Big\rangle = 0$$

due to Lemma 4.6, since $a\in\mathcal{J}^{\perp}_{X_1,\dots,X_m}$.

Analogously, if $q=k$ then $\eta_{i_{q+1}\dots i_k}=1$ and $t=s-k\le r-1$. Hence,

$$\Big\langle \ell, \sum_{\substack{i_{q+1},\dots,i_k\\ j_{t+1},\dots,j_r}} \alpha_{i_1\dots i_k}\beta_{j_1\dots j_r}(\eta_{i_{q+1}\dots i_k}\sqcup\eta_{j_{t+1}\dots j_r})\Big\rangle = \alpha_{i_1\dots i_k}\Big\langle \ell, \sum_{j_{s-k+1},\dots,j_r} \beta_{j_1\dots j_r}\eta_{j_{s-k+1}\dots j_r}\Big\rangle = 0$$

due to Lemma 4.6, since $b\in\mathcal{J}^{\perp}_{X_1,\dots,X_m}$.

Thus, all terms of (4.9) vanish, i.e., $\langle x\,\ell, a\sqcup b\rangle=0$, which proves the lemma. ■

The following notation will be used below.

NOTATION 4.8. For any $a\in\mathcal{F}$, denote by $\widetilde{a}$ the orthoprojection of a on the subspace $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$. Analogously, for any subspace $M\subset\mathcal{F}$, denote by $\widetilde{M}$ the orthoprojection of M on $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$.

Lemma 4.9. *Let homogeneous elements $\ell_1,\ldots,\ell_n\in\mathcal{L}$ be such that* (4.2) *holds. Denote by $\{\ell_j\}_{j=n+1}^{\infty}$ a homogeneous basis of the core Lie subalgebra $\mathcal{L}_{X_1,\ldots,X_m}$. Then the set*

$$\{\widetilde{\ell}_{i_1}\sqcup\cdots\sqcup\widetilde{\ell}_{i_s}\sqcup\ell_{j_1}\sqcup\cdots\sqcup\ell_{j_t}: s+t\geq 1,\ 1\leq i_1\leq\cdots\leq i_s\leq n<j_1\leq\cdots\leq j_t\}\qquad(4.10)$$

forms a basis of $\mathcal{F}$.

Proof. Without loss of generality assume that (4.3) holds. Set $p=\operatorname{ord}(\ell_n)$ and $v_k=\dim(\mathcal{L}^1\oplus\cdots\oplus\mathcal{L}^k)-\dim(\mathcal{P}^1\oplus\cdots\oplus\mathcal{P}^k)$, $k=1,\ldots,p$ (then v is a growth vector of the corresponding system (2.1) and p is its degree of nonholonomy). Then $\operatorname{ord}(\ell_i)=k$ iff $v_{k-1}+1\leq i\leq v_k$, $k=1,\ldots,p$.

Taking into account decomposition (3.3), it is sufficient to prove that any homogeneous element $\ell\in\mathcal{L}$ can be uniquely represented as a linear combination of elements of (4.10). Notice that for any element $\ell\in\mathcal{L}_{X_1,\ldots,X_m}$ this is obvious.

We argue by induction on the order. For the elements $\ell_1,\ldots,\ell_{v_1}$ of order 1 we obviously have $\widetilde{\ell}_i=\ell_i$, $i=1,\ldots,v_1$. Hence, $\mathcal{L}^1$ is contained in the linear span of (4.10).

Suppose $\mathcal{L}^1\oplus\cdots\oplus\mathcal{L}^{k-1}$ is contained in the linear span of (4.10). Consider the subspace $\mathcal{L}^k$. It was mentioned above that $\mathcal{L}_{X_1,\ldots,X_m}\cap\mathcal{L}^k$ is contained in the linear span of (4.10). Consider any element ℓ_i with $v_{k-1}+1\leq i\leq v_k$. Then $\operatorname{ord}(\ell_i)=k$. We get

$$\ell_i=\widetilde{\ell}_i+x_i,\quad\text{where}\quad x_i\in\mathcal{J}_{X_1,\ldots,X_m}\cap\mathcal{F}^k.\qquad(4.11)$$

Due to (3.2),

$$x_i=\ell_i^*+y_i,\quad\text{where}\quad \ell_i^*\in\mathcal{L}^k,\ y_i\in\mathcal{L}^{\text{sh}}\cap\mathcal{F}^k.\qquad(4.12)$$

Thus,

$$\ell_i-\ell_i^*=\widetilde{\ell}_i+y_i.\qquad(4.13)$$

The condition $y_i\in\mathcal{L}^{\text{sh}}$ means that y_i equals a linear combination of elements of the form $\ell_{i_1}\sqcup\cdots\sqcup\ell_{i_s}$, where $\ell_{i_1},\ldots,\ell_{i_s}\in\mathcal{L}$, $s\geq 2$. Hence, $\ell_{i_1},\ldots,\ell_{i_s}\in\mathcal{L}^1\oplus\cdots\oplus\mathcal{L}^{k-1}$. Therefore, due to the induction supposition, the right hand side of (4.13) can be represented as a linear combination of elements of the form (4.10).

On the other hand, for any $i=v_{k-1}+1,\ldots,v_k$, the element $\ell_i-\ell_i^*\in\mathcal{L}^k$ is uniquely defined by formulas (4.11) and (4.12). Notice that $\widetilde{\ell}_i\in\mathcal{J}^{\perp}_{X_1,\ldots,X_m}\subset\mathcal{L}^{\perp}_{X_1,\ldots,X_m}$ and $y_i\in\mathcal{L}^{\text{sh}}=\mathcal{L}^{\perp}\subset\mathcal{L}^{\perp}_{X_1,\ldots,X_m}$. Hence, (4.13) implies $\ell_i-\ell_i^*\in\mathcal{L}^{\perp}_{X_1,\ldots,X_m}$.

Denote $\{\ell_{j_1},\ldots,\ell_{j_q}\}=\{\ell_j\}_{j=n+1}^{\infty}\cap\mathcal{L}^k$, and consider the set

$$\{\ell_{j_1},\ldots,\ell_{j_q}\}\cup\{\ell_i-\ell_i^*: v_{k-1}+1\leq i\leq v_k\}\subset\mathcal{L}^k.\qquad(4.14)$$

Let us prove that its elements are linearly independent. Taking into account that $\ell_i-\ell_i^*\in\mathcal{L}^{\perp}_{X_1,\ldots,X_m}$ and $\ell_{j_1},\ldots,\ell_{j_q}\in\mathcal{L}_{X_1,\ldots,X_m}$, it is sufficient to prove that the elements $\ell_i-\ell_i^*$, $i=v_{k-1}+1,\ldots,v_k$, are linearly independent. Assume the converse. Then

$$\sum_{i=v_{k-1}+1}^{v_k}\mu_i(\ell_i-\ell_i^*)=0$$

for some numbers μ_i such that $\sum_{i=v_{k-1}+1}^{v_k}\mu_i^2>0$. Due to (4.13), this implies

$$\sum_{i=v_{k-1}+1}^{v_k}\mu_i\widetilde{\ell}_i=-\sum_{i=v_{k-1}+1}^{v_k}\mu_i y_i\in\mathcal{L}^{\text{sh}}=\mathcal{L}^{\perp}.$$

In particular, $\sum_{i=v_{k-1}+1}^{v_k}\mu_i\widetilde{\ell}_i$ is orthogonal to $\sum_{i=v_{k-1}+1}^{v_k}\mu_i\ell_i$. Since by definition $\widetilde{\ell}_i$ is the orthoprojection of ℓ_i on the subspace $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$, we see that

$$\sum_{i=v_{k-1}+1}^{v_k}\mu_i\ell_i\in\mathcal{J}_{X_1,\dots,X_m}\cap\mathcal{L}=\mathcal{L}_{X_1,\dots,X_m},\quad\text{where}\quad\sum_{i=v_{k-1}+1}^{v_k}\mu_i^2>0,$$

which contradicts the definition of the elements ℓ_i, $i=v_{k-1}+1,\dots,v_k$.

Thus, the elements of the set (4.14) are linearly independent. Note that the number of these elements equals $\dim\mathcal{L}^k$. Hence, the set (4.14) is a basis of $\mathcal{L}^k$, and any element of this basis can be represented as a linear combination of elements of (4.10), due to (4.13) and the induction supposition.

The induction arguments show that any homogeneous element $\ell\in\mathcal{L}$ can be represented as a linear combination of elements of (4.10). As was mentioned above, the decomposition (3.3) implies that this is true for any element from $\mathcal{F}$, that is, the linear span of (4.10) coincides with $\mathcal{F}$.

However, for any $k\geq 1$, the number of elements of (4.10) of order k equals $\dim(\mathcal{F}^k)$, since it is the same as the number of elements of the Poincaré–Birkhoff–Witt basis of order k. This means that elements of (4.10) are linearly independent, which completes the proof. ∎

Theorem 4.10 (generalization of R. Ree's theorem). *Let elements $\ell_1,\dots,\ell_n\in\mathcal{L}$ be homogeneous and satisfy* (4.2). *Then the set*

$$\{\widetilde{\ell}_{i_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup\widetilde{\ell}_{i_s}: s\geq 1,\ 1\leq i_1\leq\cdots\leq i_s\leq n\}\tag{4.15}$$

is a basis of $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$.

Notice that $\widetilde{\mathcal{L}}=\operatorname{Lin}\{\widetilde{\ell}_1,\dots,\widetilde{\ell}_n\}$, where $\widetilde{\mathcal{L}}$ is the orthoprojection of $\mathcal{L}$ on $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$. Hence, Theorem 4.10 says that

$$\mathcal{J}^{\perp}_{X_1,\dots,X_m}=\widetilde{\mathcal{L}}\oplus^{\perp}(\widetilde{\mathcal{L}})^{\mathrm{sh}},$$

and therefore

$$\mathcal{F}=\mathcal{J}_{X_1,\dots,X_m}\oplus^{\perp}\widetilde{\mathcal{L}}\oplus^{\perp}(\widetilde{\mathcal{L}})^{\mathrm{sh}},$$

which generalizes R. Ree's decomposition (3.3).

Proof. Let $\{\ell_j\}_{j=n+1}^{\infty}$ be a homogeneous basis of the core Lie subalgebra $\mathcal{L}_{X_1,\dots,X_m}$. For any $k\geq 1$, consider the set

$$\{\widetilde{\ell}_{i_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup\widetilde{\ell}_{i_s}\sqcup\!\sqcup\ell_{j_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup\ell_{j_t}\in\mathcal{F}^k: s+t\geq 1,\ 1\leq i_1\leq\cdots\leq i_s\leq n<j_1\leq\cdots\leq j_t\}.\tag{4.16}$$

Due to the Poincaré–Birkhoff–Witt theorem, the number of elements in the set (4.16) equals $\dim\mathcal{F}^k$. Corollary 4.4 implies that the number of elements in (4.16) with $t\geq 1$ equals $\dim(\mathcal{J}_{X_1,\dots,X_m}\cap\mathcal{F}^k)$. Hence, the number of elements in (4.16) with $t=0$ equals $\dim(\mathcal{J}^{\perp}_{X_1,\dots,X_m}\cap\mathcal{F}^k)$. The latter elements are of the form

$$\{\widetilde{\ell}_{i_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup\widetilde{\ell}_{i_s}\in\mathcal{F}^k: s\geq 1,\ 1\leq i_1\leq\cdots\leq i_s\leq n\}.\tag{4.17}$$

Due to Lemma 4.9, these elements are linearly independent, and due to Lemma 4.7 they belong to $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$. Hence, the set (4.17) forms a basis of $\mathcal{J}^{\perp}_{X_1,\dots,X_m}\cap\mathcal{F}^k$.

Thus, the set

$$\{\widetilde{\ell}_{i_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \widetilde{\ell}_{i_s} : s \geq 1,\, 1 \leq i_1 \leq \cdots \leq i_s \leq n\}$$

is a basis of $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$, which implies the direct sum decomposition

$$\mathcal{J}^{\perp}_{X_1,\dots,X_m} = \widetilde{\mathcal{L}} \oplus (\widetilde{\mathcal{L}})^{\mathrm{sh}}.$$

It remains to prove that $\widetilde{\mathcal{L}}$ is orthogonal to $(\widetilde{\mathcal{L}})^{\mathrm{sh}}$. For any $1 \leq i \leq n$ we have $\ell_i = \widetilde{\ell}_i + x_i$, where $x_i \in \mathcal{J}_{X_1,\dots,X_m}$. Since $\widetilde{\ell}_{i_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \widetilde{\ell}_{i_s} \in \mathcal{J}^{\perp}_{X_1,\dots,X_m}$ for any $1 \leq i_1, \dots, i_s \leq n$ due to Lemma 4.7, we see that if $s \geq 2$ then

$$\langle \widetilde{\ell}_i, \widetilde{\ell}_{i_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \widetilde{\ell}_{i_s}\rangle = \langle \ell_i, \widetilde{\ell}_{i_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \widetilde{\ell}_{i_s}\rangle = 0,$$

due to R. Ree's theorem. Hence, $\widetilde{\mathcal{L}}$ is orthogonal to $(\widetilde{\mathcal{L}})^{\mathrm{sh}}$, which completes the proof. ■

Notice that (4.15) can be rewritten as

$$\{\widetilde{\ell}_1^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \widetilde{\ell}_n^{\sqcup\!\sqcup q_n} : q_1, \dots, q_n \geq 0,\, q_1 + \cdots + q_n \geq 1\}.$$

Remark 4.11. Theorem 4.10 implies that the subspace $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$ equipped with the shuffle product operation is isomorphic to the algebra of polynomials of n variables without constant term (with coefficients from $\mathbb{R}$).

4.3. Construction of privileged coordinates. Let us explain how Theorem 4.10 can be used to construct privileged coordinates.

For any $k \geq 1$, consider an orthonormal basis B_k^0 of the subspace $\mathcal{J}_{X_1,\dots,X_m} \cap \mathcal{F}^k$,

$$B_k^0 = \{b_{k,j}^0 : j = 1, \dots, r_k^0\}, \quad r_k^0 = \dim(\mathcal{J}_{X_1,\dots,X_m} \cap \mathcal{F}^k),$$

an orthonormal basis B_k^1 of the subspace $\widetilde{\mathcal{L}} \cap \mathcal{F}^k$,

$$B_k^1 = \{b_{k,j}^1 : j = 1, \dots, r_k^1\}, \quad r_k^1 = \dim(\widetilde{\mathcal{L}} \cap \mathcal{F}^k),$$

and an orthonormal basis B_k^2 of the subspace $(\widetilde{\mathcal{L}})^{\mathrm{sh}} \cap \mathcal{F}^k$,

$$B_k^2 = \{b_{k,j}^2 : j = 1, \dots, r_k^2\}, \quad r_k^2 = \dim((\widetilde{\mathcal{L}})^{\mathrm{sh}} \cap \mathcal{F}^k).$$

Then the set $\bigcup_{k\geq 1}(B_k^0 \cup B_k^1 \cup B_k^2)$ is an orthonormal basis of $\mathcal{F}$. Hence, the series on the right hand side of (2.23) can be re-expanded in this basis, which gives

$$\mathcal{E}_{X_1,\dots,X_m} = \sum_{k=1}^{\infty}\Big(\sum_{j=1}^{r_k^0} c(b_{k,j}^0)b_{k,j}^0 + \sum_{j=1}^{r_k^1} c(b_{k,j}^1)b_{k,j}^1 + \sum_{j=1}^{r_k^2} c(b_{k,j}^2)b_{k,j}^2\Big). \tag{4.18}$$

Notice that the definition of B_k^1 gives $\bigcup_{k\geq 1} B_k^1 = \{\widetilde{\ell}_1, \dots, \widetilde{\ell}_n\}$. Without loss of generality we may assume $c(\ell_i) = e_i$, $i = 1, \dots, n$; then, due to (4.3), the coordinates are linearly adapted and $B_k^1 = \{\widetilde{\ell}_{v_{k-1}+1}, \dots, \widetilde{\ell}_{v_k}\}$, $k = 1, \dots, p$, where v is the growth vector of the system. Moreover, $w_i = \mathrm{ord}(\ell_i)$ equals the weight of the coordinate x_i, $i = 1, \dots, n$.

Let us show the way of constructing privileged coordinates. Obviously, $w_1 = 1$. We have $c(b_{k,j}^0) = 0$ for $k = 1$, and $\mathrm{ord}(b_{k,j}^2) \geq 2$ for any k, j. Moreover, $\widetilde{\ell}_1 = \ell_1$. Hence, $c(\widetilde{\ell}_1) = c(\ell_1) = e_1$. Therefore,

$$(\mathcal{E}_{X_1,\dots,X_m})_1 = \widetilde{\ell}_1 + \sum_{k=w_1+1}^{\infty} \alpha^1_{i_1\dots i_k}\eta_{i_1\dots i_k}.$$

Suppose that after some change of variables for some $q \geq 1$ we have

$$\big(\mathcal{E}_{X_1^{(q)},\dots,X_m^{(q)}}\big)_i = \widetilde{\ell}_i + \sum_{k=w_i+1}^{\infty} \alpha^i_{i_1\dots i_k}\eta_{i_1\dots i_k}, \qquad i = 1,\dots,q,$$

where $X_1^{(q)},\dots,X_m^{(q)}$ are the initial vector fields expressed in the new variables (for $q = 1$ they coincide with the initial vector fields). Let us consider the $(q+1)$th coordinate,

$$\big(\mathcal{E}_{X_1^{(q)},\dots,X_m^{(q)}}\big)_{q+1} = \sum_{k=1}^{\infty}\Big(\sum_{j=1}^{d_k^0}(c^{(q)}(b^0_{k,j}))_{q+1}b^0_{k,j} + \sum_{j=1}^{d_k^1}(c^{(q)}(b^1_{k,j}))_{q+1}b^1_{k,j} + \sum_{j=1}^{d_k^2}(c^{(q)}(b^2_{k,j}))_{q+1}b^2_{k,j}\Big),$$

where the map $c^{(q)}$ corresponds to the system in the new coordinates.

Notice that the elements $b^1_{k,j}$ with $k < w_{q+1}$ are linear combinations of $\widetilde{\ell}_i$ with $i < w_{q+1}$. Hence, they can be killed by a linear change of variables.

Since the elements $b^2_{k,j}$ are shuffles of elements of the form $b^1_{q,t}$ with $q < k$, one kills all elements $b^2_{k,j}$ with $k \leq w_{q+1}$ by a polynomial change of variables.

Suppose that this has been done. Then we get

$$\begin{aligned}\big(\mathcal{E}_{X_1^{(q+1)},\dots,X_m^{(q+1)}}\big)_{q+1} &= \sum_{k=1}^{\infty}\sum_{j=1}^{d_k^0}(c^{(q+1)}(b^0_{k,j}))_{q+1}b^0_{k,j} \\ &\quad + \sum_{w_i \geq w_{q+1}}(c^{(q+1)}(\widetilde{\ell}_i))_{q+1}\widetilde{\ell}_i + \sum_{k=w_{q+1}+1}^{\infty}\sum_{j=1}^{d_k^2}(c^{(q+1)}(b^2_{k,j}))_{q+1}b^2_{k,j},\end{aligned}$$

where the map $c^{(q+1)}$ corresponds to the system in the new coordinates. Since the left ideal is invariant with respect to changes of variables, we get $(c^{(q+1)}(b^0_{k,j}))_{q+1} = 0$ for $k \leq w_{q+1}$, and $(c^{(q+1)}(\widetilde{\ell}_i))_{q+1} = (c^{(q+1)}(\ell_i))_{q+1} = \delta_{i,q+1}$ for i such that $w_i = w_{q+1}$. Thus,

$$\big(\mathcal{E}_{X_1^{(q+1)},\dots,X_m^{(q+1)}}\big)_{q+1} = \widetilde{\ell}_{q+1} + \sum_{k=w_{q+1}+1}^{\infty} \alpha^{q+1}_{i_1\dots i_k}\eta_{i_1\dots i_k}.$$

Notice that the described polynomial change of variables is of the form $y_{q+1} = x_{q+1} + p_{q+1}(x_1,\dots,x_q)$ and $y_i = x_i$, $i \neq q+1$. Hence, it is nonsingular.

By induction, there exists a polynomial nonsingular change of variables that reduces the endpoint map to the form $\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m}$ such that

$$\big(\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m}\big)_i = \widetilde{\ell}_i + \sum_{k=w_i+1}^{\infty} \alpha^i_{i_1\dots i_k}\eta_{i_1\dots i_k}, \qquad i = 1,\dots,n.$$

This means that these new coordinates are privileged. As will be shown below, the elements $\widetilde{\ell}_1,\dots,\widetilde{\ell}_n$ describe a homogeneous approximation of the system.

In the next subsection we obtain this result in another way that allows us to describe explicitly all privileged coordinates.

4.4. Dual basis. Now we are going to give another way for re-expansion of the series from (2.23), which is more convenient than the representation (4.18).

Suppose $\{\ell_i\}_{i=1}^{\infty}$ is an arbitrary homogeneous basis of $\mathcal{L}$. For our further purposes, it is convenient to rewrite the Poincaré–Birkhoff–Witt basis (4.4) in the form

$$\{\ell_{j_1}^{p_1}\cdots\ell_{j_s}^{p_s} : s\ge 1,\ 1\le j_1<\cdots<j_s,\ p_1,\dots,p_s\ge 1\}, \tag{4.19}$$

where $\ell^p=\ell\cdots\ell$ (p times), $p\ge 1$. Since all elements ℓ_i, $i\ge 1$, are homogeneous, all basis elements are homogeneous as well, and for any $k\ge 1$ the set

$$\{\ell_{j_1}^{p_1}\cdots\ell_{j_s}^{p_s}\in\mathcal{F}^k : s\ge 1,\ 1\le j_1<\cdots<j_s,\ p_1,\dots,p_s\ge 1\}$$

is a basis of $\mathcal{F}^k$. Since $\dim\mathcal{F}^k<\infty$, there exists a dual basis in $\mathcal{F}^k$. Taking into account that the subspaces $\mathcal{F}^k$ with different k are orthogonal to each other, we see that there exists a dual basis of $\mathcal{F}$. Denote this basis by

$$\{d_{i_1\dots i_r}^{q_1\dots q_r} : r\ge 1,\ 1\le i_1<\cdots<i_r,\ q_1,\dots,q_r\ge 1\}, \tag{4.20}$$

where

$$\langle\ell_{j_1}^{p_1}\cdots\ell_{j_s}^{p_s}, d_{i_1\dots i_r}^{q_1\dots q_r}\rangle=\begin{cases}1 & \text{if } s=r \text{ and } j_t=i_t,\ t=1,\dots,s,\\ 0 & \text{otherwise.}\end{cases} \tag{4.21}$$

Now we are going to use the following description of the dual basis.

Theorem 4.12 (G. Melançon and C. Reutenauer [44]). *Elements of the dual basis* (4.20) *can be found by*

$$d_{i_1\dots i_r}^{q_1\dots q_r}=\frac{1}{q_1!\cdots q_r!}\,d_{i_1}^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_{i_r}^{\sqcup\!\sqcup q_r},$$

where for brevity we set $d_q=d_q^1$, $q\ge 1$.

Below and throughout the paper we choose a basis $\{\ell_i\}_{i=1}^{\infty}$ so that (4.2) and (4.3) hold, and $\{\ell_j\}_{j=n+1}^{\infty}$ is a homogeneous basis of $\mathcal{L}_{X_1,\dots,X_m}$. Then the dual basis (4.20) can be used to describe a basis of the subspace $\mathcal{J}_{X_1,\dots,X_m}^{\perp}$.

Lemma 4.13. *Elements* $d_1^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n}$ (*where* $q_1+\cdots+q_n\ge 1$) *are orthogonal to* $\mathcal{J}_{X_1,\dots,X_m}$.

Proof. By Corollary 4.4, any element of $\mathcal{J}_{X_1,\dots,X_m}$ equals a linear combination of elements of the form $\ell_{j_1}^{p_1}\cdots\ell_{j_s}^{p_s}$, where $j_1<\cdots<j_s$ and $j_s\ge n+1$. Hence, it is orthogonal to any element $d_1^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n}$, by definition and due to the Melançon–Reutenauer theorem. ∎

Lemma 4.14. *The set*

$$\{d_1^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n} : q_1,\dots,q_n\ge 0,\ q_1+\cdots+q_n\ge 1\}$$

forms a basis of $\mathcal{J}_{X_1,\dots,X_m}^{\perp}$.

Proof. For any $k\ge 1$, let us consider the set

$$\{d_1^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n}\in\mathcal{F}^k : q_1,\dots,q_n\ge 0,\ q_1+\cdots+q_n\ge 1\}. \tag{4.22}$$

This set is contained in $\mathcal{J}_{X_1,\dots,X_m}^{\perp}$, due to Lemma 4.13. Moreover, all elements of (4.22) belong to the dual basis (4.20) (up to multipliers), hence they are linearly independent. The number of elements coincides with $\dim(\mathcal{J}_{X_1,\dots,X_m}^{\perp}\cap\mathcal{F}^k)$, since it coincides with the number of elements in the set (4.17). Hence, (4.22) is a basis of $\mathcal{J}_{X_1,\dots,X_m}^{\perp}\cap\mathcal{F}^k$. Finally, the union of sets (4.22) for all $k\ge 1$ forms a basis of $\mathcal{J}_{X_1,\dots,X_m}^{\perp}$. ∎

Corollary 4.15. *For any $i = 1, \ldots, n$, the element $\widetilde{\ell}_i$ equals a homogeneous shuffle polynomial of $d_1, \ldots, d_n$. Conversely, for any $i = 1, \ldots, n$, the element d_i equals a homogeneous shuffle polynomial of $\widetilde{\ell}_1, \ldots, \widetilde{\ell}_n$. Moreover,* $\mathrm{ord}(d_i) = \mathrm{ord}(\widetilde{\ell}_i) = \mathrm{ord}(\ell_i)$, $i = 1, \ldots, n$.

Example 4.16. Let us again consider system (2.22) from Example 2.21. In Example 2.31 we have found $\mathcal{L}_{X_1,X_2}$ and chosen three complement elements ℓ_1, ℓ_2, ℓ_3. Let us find the left ideal $\mathcal{J}_{X_1,X_2}$ and the orthoprojections of the complement elements.

We use (2.28). Obviously, $\mathcal{J}_{X_1,X_2} \cap \mathcal{F}^1 = \mathcal{L}_{X_1,X_2} \cap \mathcal{F}^1 = \mathrm{Lin}\{\eta_2\}$. Hence, all elements of the form $\eta_{i_1\ldots i_k 2}$ also belong to $\mathcal{J}_{X_1,X_2}$, which gives $\mathcal{J}_{X_1,X_2} \cap \mathcal{F}^2 = \mathrm{Lin}\{\eta_{12}, \eta_{22}\}$.

Since $[[\eta_2,\eta_1],\eta_1] = \eta_{211} - 2\eta_{121} + \eta_{112}$, $[[\eta_2,\eta_1],\eta_2] = -\eta_{221} + 2\eta_{212} - \eta_{122}$, and $\eta_{122}, \eta_{212}, \eta_{112} \in \mathcal{J}_{X_1,X_2}$, we get

$$\mathcal{J}_{X_1,X_2} \cap \mathcal{F}^3 = \mathrm{Lin}\{\eta_{112}, \eta_{122}, \eta_{212}, \eta_{222}, \eta_{211} - 2\eta_{121}, \eta_{221}\}$$

and

$$\mathcal{J}_{X_1,X_2} \cap \mathcal{F}^4 = \mathrm{Lin}\big\{\eta_{1112}, \eta_{1122}, \eta_{1212}, \eta_{1222}, \eta_{1211} - 2\eta_{1121}, \eta_{1221}, \\ \eta_{2112}, \eta_{2122}, \eta_{2212}, \eta_{2222}, \eta_{2211} - 2\eta_{2121}, \eta_{2221}\big\}. \tag{4.23}$$

Now let us find the orthoprojections of the elements (2.29) on the subspace $\mathcal{J}^{\perp}_{X_1,X_2}$. Since $\eta_2 \in \mathcal{J}_{X_1,X_2}$, we get $\widetilde{\ell}_1 = \eta_1$. Analogously, since $\eta_{12} \in \mathcal{J}_{X_1,X_2}$, we get $\widetilde{\ell}_2 = -2\eta_{21}$. Notice that the elements $\widetilde{\ell}_1^{\sqcup\!\sqcup 4} = 24\eta_{1111}$, $\widetilde{\ell}_1^{\sqcup\!\sqcup 2} \sqcup\!\sqcup \widetilde{\ell}_2 = -12\eta_{2111} - 8\eta_{1211} - 4\eta_{1121}$, and $\widetilde{\ell}_2^{\sqcup\!\sqcup 2} = 8\eta_{2121} + 16\eta_{2211}$ are orthogonal to all elements from (4.23).

Finally, notice that $[[[\eta_2,\eta_1],\eta_2],\eta_2] \in \mathcal{J}_{X_1,X_2}$ and

$$[[[\eta_2,\eta_1],\eta_1],\eta_1] = \eta_{2111} - 3\eta_{1211} + 3\eta_{1121} - \eta_{1112},$$

where $\eta_{1112} \in \mathcal{J}_{X_1,X_2}$. Obviously, the element $\eta_{2111} - 3\eta_{1211} + 3\eta_{1121}$ is orthogonal to all elements from (4.23) except $\eta_{1211} - 2\eta_{1121}$. Hence, its orthoprojection on $\mathcal{J}^{\perp}_{X_1,X_2}$ equals $\eta_{2111} - 3\eta_{1211} + 3\eta_{1121} + \alpha(\eta_{1211} - 2\eta_{1121})$ where α is such that

$$\big\langle \eta_{2111} - 3\eta_{1211} + 3\eta_{1121} + \alpha(\eta_{1211} - 2\eta_{1121}), \eta_{1211} - 2\eta_{1121} \big\rangle = 0,$$

which gives $\alpha = \frac{9}{5}$. Finally, we get $\widetilde{\ell}_3 = 3\eta_{2111} - \frac{18}{5}\eta_{1211} - \frac{9}{5}\eta_{1121}$.

Now let us find the elements of the dual basis. For definiteness, choose $\ell_4 = \eta_2$. Then d_1 is found from the equalities $\langle d_1, \ell_1\rangle = 1$ and $\langle d_1, \ell_4\rangle = 0$, which gives $d_1 = \eta_1 = \widetilde{\ell}_1$. Analogously, d_2 is found from the equalities $\langle d_2, \ell_2\rangle = 1$ and $\langle d_2, \ell_1\ell_1\rangle = \langle d_2, \ell_1\ell_4\rangle = \langle d_2, \ell_4\ell_4\rangle = 0$, which gives $d_2 = -\frac{1}{2}\eta_{21} = \frac{1}{4}\widetilde{\ell}_2$.

Also, choose $\ell_5 = [[\eta_2,\eta_1],\eta_1]$, $\ell_6 = [[\eta_2,\eta_1],\eta_2]$, $\ell_7 = [[[\eta_2,\eta_1],\eta_1],\eta_2]$, and $\ell_8 = [[[\eta_2,\eta_1],\eta_2],\eta_2]$. The element d_3 satisfies the equality $\langle d_3, \ell_3\rangle = 1$ and is orthogonal to

$$\ell_1\ell_1\ell_1\ell_1, \quad \ell_1\ell_1\ell_1\ell_4, \quad \ell_1\ell_1\ell_4\ell_4, \quad \ell_1\ell_4\ell_4\ell_4, \quad \ell_4\ell_4\ell_4\ell_4,$$
$$\ell_1\ell_1\ell_2, \quad \ell_1\ell_2\ell_4, \quad \ell_2\ell_2, \quad \ell_2\ell_4\ell_4, \quad \ell_1\ell_5, \quad \ell_1\ell_6, \quad \ell_4\ell_5, \quad \ell_4\ell_6, \quad \ell_7, \quad \ell_8.$$

These conditions give $d_3 = \frac{1}{3}\eta_{2111}$. Notice that $\eta_1 \sqcup\!\sqcup \eta_1 \sqcup\!\sqcup \eta_{21} = 6\eta_{2111} + 4\eta_{1211} + 2\eta_{1121}$, hence

$$d_3 = \tfrac{5}{126}\widetilde{\ell}_3 - \tfrac{1}{56}\widetilde{\ell}_1^{\sqcup\!\sqcup 2} \sqcup\!\sqcup \widetilde{\ell}_2.$$

4.5. Expansion of the endpoint map in the dual basis. Let us apply properties of a dual basis to the series representation of the endpoint map.

First, for any $k \geq 1$ consider any element $a \in \mathcal{F}^k$. Definition 3.2 of the inner product implies that

$$a = \sum_{1 \leq i_1, \ldots, i_k \leq m} \langle a, \eta_{i_1 \ldots i_k} \rangle \eta_{i_1 \ldots i_k}. \tag{4.24}$$

Re-expanding this element with respect to the dual basis (4.20) and taking into account Theorem 4.12 we get the representation

$$a = \sum{}' \frac{1}{q_1! \cdots q_r!} \langle a, \ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} \rangle \, d_{i_1}^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d_{i_r}^{\sqcup\!\sqcup q_r}, \tag{4.25}$$

where the sum $\sum'$ is taken over all indices $1 \leq i_1 < \cdots < i_r$ and $q_1, \ldots, q_r \geq 1$ such that $\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} \in \mathcal{F}^k$.

Now let us turn to the series $\mathcal{E}_{X_1,\ldots,X_m}$ and consider such a representation for its components $(\mathcal{E}_{X_1,\ldots,X_m})_j$, $j = 1, \ldots, n$. More specifically, let us fix $k \geq 1$ and consider any basis element $\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} \in \mathcal{F}^k$. Set $a = \sum'' (c(\eta_{i_1 \ldots i_k}))_j \eta_{i_1 \ldots i_k}$, where the sum $\sum''$ is taken over all indices $1 \leq i_1, \ldots, i_k \leq m$. Since c is a linear map, using (4.24) we get

$$\begin{aligned}\langle a, \ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} \rangle &= \sum{}'' (c(\eta_{i_1 \ldots i_k}))_j \langle \eta_{i_1 \ldots i_k}, \ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} \rangle = \sum{}'' \big(c(\langle \eta_{i_1 \ldots i_k}, \ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} \rangle \eta_{i_1 \ldots i_k}) \big)_j \\ &= \Big(c\Big(\sum{}'' \langle \eta_{i_1 \ldots i_k}, \ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} \rangle \eta_{i_1 \ldots i_k} \Big) \Big)_j = (c(\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r}))_j.\end{aligned}$$

Hence, (4.25) implies

$$a = \sum{}' \frac{1}{q_1! \cdots q_r!} (c(\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r}))_j \, d_{i_1}^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d_{i_r}^{\sqcup\!\sqcup q_r}.$$

Applying these arguments for all $k \geq 1$ and all $j = 1, \ldots, n$, we get the following result.

THEOREM 4.17. *Suppose $\{\ell_1, \ldots, \ell_n\}$ is a set of homogeneous elements of $\mathcal{L}$ such that* (4.2) *holds, and $\{\ell_j\}_{j=n+1}^{\infty}$ is a homogeneous basis of $\mathcal{L}_{X_1,\ldots,X_m}$. Then the series on the right hand side of* (2.23) *can be represented in the form*

$$\mathcal{E}_{X_1,\ldots,X_m} = \sum_{\substack{1 \leq i_1 < \cdots < i_r \\ q_1, \ldots, q_r \geq 1}} \frac{1}{q_1! \cdots q_r!} c(\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r}) \, d_{i_1}^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d_{i_r}^{\sqcup\!\sqcup q_r}, \tag{4.26}$$

where $d_j = d_j^1$ are elements of the dual basis (4.20).

Now we separate terms containing only $\ell_1, \ldots, \ell_n$. So, we get

$$\mathcal{E}_{X_1,\ldots,X_m} = \mathcal{S} + \mathcal{T}, \tag{4.27}$$

where

$$\mathcal{S} = \sum_{\substack{q_1, \ldots, q_n \geq 0 \\ q_1 + \cdots + q_n \geq 1}} \frac{1}{q_1! \cdots q_n!} c(\ell_1^{q_1} \cdots \ell_n^{q_n}) \, d_1^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n}, \tag{4.28}$$

$$\mathcal{T} = \sum_{\substack{1 \leq i_1 < \cdots < i_r,\, i_r \geq n+1 \\ q_1, \ldots, q_r \geq 1}} \frac{1}{q_1! \cdots q_r!} c(\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r}) \, d_{i_1}^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d_{i_r}^{\sqcup\!\sqcup q_r}, \tag{4.29}$$

where we set $\ell^0 = 1$. If $i_r \geq n+1$ then $\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} \in \mathcal{J}_{X_1,\dots,X_m}$, which means that all coefficients in the series $\mathcal{T}$ belong to $c(\mathcal{J}_{X_1,\dots,X_m})$. This implies the following lemma.

LEMMA 4.18. *Suppose $i = 1, \dots, n$ is fixed and $(\mathcal{S})_i$ contains only terms of order no less than k. Then $(\mathcal{T})_i$ contains only terms of order greater than k.*

Proof. It is sufficient to prove that $(c(\mathcal{J}_{X_1,\dots,X_m} \cap \mathcal{F}^j))_i = 0$ for any $j = 1, \dots, k$.

The proof is by induction on j. For $j = 1$, there is nothing to prove, since $c(\mathcal{J}_{X_1,\dots,X_m} \cap \mathcal{F}^1) = 0$ due to Lemma 4.2.

Suppose that for some $1 \leq j < k$ one has $(c(\mathcal{J}_{X_1,\dots,X_m} \cap (\mathcal{F}^1 \oplus \cdots \oplus \mathcal{F}^j)))_i = 0$. Consider any $a \in \mathcal{J}_{X_1,\dots,X_m} \cap \mathcal{F}^{j+1}$. Due to Lemma 4.2, $c(a) \in c(\mathcal{F}^1 \oplus \cdots \oplus \mathcal{F}^j)$. We have

$$c(\mathcal{F}^1 \oplus \cdots \oplus \mathcal{F}^j) = c(M^j) + c(N^j),$$

where we denote temporarily

$$\begin{aligned} M^j &= \mathrm{Lin}\{\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} : i_1 \leq \cdots \leq i_r \leq n\} \cap (\mathcal{F}^1 \oplus \cdots \oplus \mathcal{F}^j), \\ N^j &= \mathrm{Lin}\{\ell_{i_1}^{q_1} \cdots \ell_{i_r}^{q_r} : i_1 \leq \cdots \leq i_r,\ i_r \geq n+1\} \cap (\mathcal{F}^1 \oplus \cdots \oplus \mathcal{F}^j). \end{aligned}$$

However, $(c(M^j))_i = 0$ since, due to the condition of the lemma, $(\mathcal{S})_i$ contains only terms of order no less than k (recall that $j < k$), and $(c(N^j))_i = 0$ due to the induction supposition. Thus, $(c(\mathcal{F}^1 \oplus \cdots \oplus \mathcal{F}^j))_i = 0$, which gives $(c(a))_i = 0$. The induction arguments complete the proof. ∎

4.6. Weight, order, and privileged coordinates again. Now let us return to the concepts of the weight, the order, and privileged coordinates, and reformulate them taking into account the representation (4.27)–(4.29).

As before, suppose that (4.2) and (4.3) hold. Due to Corollary 2.30, the vectors $c(\ell_1), \dots, c(\ell_n)$ are linearly independent. Without loss of generality assume $c(\ell_i) = e_i$, $i = 1, \dots, n$. Then the coordinates are linearly adapted.

The weight of the coordinate x_i equals $w_i = \mathrm{ord}(\ell_i)$, $i = 1, \dots, n$.

The order of the coordinate function $f_i(x) = x_i$ equals the minimal order of an element that enters $(\mathcal{S})_i$ or $(\mathcal{T})_i$ with a nonzero coefficient.

Lemma 4.18 says that if $(\mathcal{S})_i$ contains terms of order w_i or more, then $(\mathcal{T})_i$ contains terms of order greater than w_i. Hence, we are led to the following reformulation.

The order of the coordinate function $f_i(x) = x_i$ equals the minimal order of an element that enters $(\mathcal{S})_i$ with a nonzero coefficient.

Therefore, we get a new "definition" of privileged coordinates.

Privileged coordinates are those for which

$$\text{if } \mathrm{ord}(\ell_1^{q_1} \cdots \ell_n^{q_n}) < \mathrm{ord}(\ell_i) \quad \text{then} \quad (c(\ell_1^{q_1} \cdots \ell_n^{q_n}))_i = 0,\ i = 1, \dots, n.$$

Hence, to construct privileged coordinates, we should reduce $\mathcal{S}$ to a "triangular form", i.e., to the form

$$(\mathcal{S})_i = d_i + \text{"elements of order} \geq \mathrm{ord}(\ell_i)\text{"}, \qquad i = 1, \dots, n. \tag{4.30}$$

In other words, we should exclude the elements

$$\{d_1^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n} : \mathrm{ord}(d_1^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n}) < \mathrm{ord}(\ell_i)\}$$

from the ith component of $\mathcal{S}$. Suppose a change of variables $y = Q(x)$ in the system is applied. Then the endpoint map $\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m} = Q(\mathcal{E}_{X_1,\dots,X_m})$ takes the form (2.21). Taking into account (4.27)–(4.29), we get

$$\begin{aligned}\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m} &= Q(\mathcal{E}_{X_1,\dots,X_m})\\ &= \sum_{q=1}^{\infty}\sum_{j_1+\dots+j_n=q}\frac{1}{j_1!\dots j_n!}\,\frac{\partial^{j_1+\dots+j_n}Q(0)}{\partial x_1^{j_1}\cdots\partial x_n^{j_n}}\,(\mathcal{S}+\mathcal{T})_1^{\sqcup\!\sqcup j_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup(\mathcal{S}+\mathcal{T})_n^{\sqcup\!\sqcup j_n}\\ &= Q(\mathcal{S})+\mathcal{T}',\end{aligned}\tag{4.31}$$

where

$$\mathcal{T}' = \sum_{q=1}^{\infty}\sum_{j_1+\dots+j_n=q}\sum_{\substack{0\le k_i\le j_i\\ k_1+\dots+k_n\ge 1}}\alpha_{j_1\dots j_n}^{k_1\dots k_n}\mathcal{S}_1^{\sqcup\!\sqcup(j_1-k_1)}\sqcup\!\sqcup\mathcal{T}_1^{\sqcup\!\sqcup k_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup\mathcal{S}_n^{\sqcup\!\sqcup(j_n-k_n)}\sqcup\!\sqcup\mathcal{T}_n^{\sqcup\!\sqcup k_n}$$

and

$$\alpha_{j_1\dots j_n}^{k_1\dots k_n} = \frac{\partial^{j_1+\dots+j_n}Q(0)}{\partial x_1^{j_1}\cdots\partial x_n^{j_n}}\,\frac{1}{(j_1-k_1)!k_1!\cdots(j_n-k_n)!k_n!}.$$

In particular, each term of the series $\mathcal{T}'$ necessarily includes a multiplier $\mathcal{T}_j$ for some $j = 1,\dots,n$.

On the other hand, a representation of the form (4.26) gives

$$\mathcal{E}_{\widehat{X}_1,\dots,\widehat{X}_m} = \sum_{\substack{1\le i_1<\dots<i_r\\ q_1,\dots,q_r\ge 1}}\frac{1}{q_1!\cdots q_r!}\widehat{c}(\ell_{i_1}^{q_1}\cdots\ell_{i_r}^{q_r})\,d_{i_1}^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_{i_r}^{\sqcup\!\sqcup q_r} = \widehat{\mathcal{S}}+\widehat{\mathcal{T}},\tag{4.32}$$

where $\widehat{c}$ denotes the linear operator $\widehat{c}:\mathcal{F}\to\mathbb{R}^n$ defined as $\widehat{c}(\eta_{i_1\dots i_k}) = \widehat{X}_{i_k}\cdots\widehat{X}_{i_1}E(0)$ and

$$\begin{aligned}\widehat{\mathcal{S}} &= \sum_{q_1,\dots,q_n\ge 0}\frac{1}{q_1!\cdots q_n!}\widehat{c}(\ell_1^{q_1}\cdots\ell_n^{q_n})\,d_1^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n},\\ \widehat{\mathcal{T}} &= \sum_{\substack{1\le i_1<\dots<i_r,\ i_r\ge n+1\\ q_1,\dots,q_r\ge 1}}\frac{1}{q_1!\cdots q_r!}\widehat{c}(\ell_{i_1}^{q_1}\cdots\ell_{i_r}^{q_r})\,d_{i_1}^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_{i_r}^{\sqcup\!\sqcup q_r}.\end{aligned}$$

Let us compare the expressions (4.31) and (4.32). We see that all terms of the form $d_1^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_n^{\sqcup\!\sqcup q_n}$ are included in $Q(\mathcal{S})$, while all terms of the form $d_{i_1}^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_{i_r}^{\sqcup\!\sqcup q_r}$ with $i_1<\dots<i_r$ and $i_r\ge n+1$ are included in $\mathcal{T}'$. Hence,

$$\widehat{\mathcal{S}} = Q(\mathcal{S})\quad\text{and}\quad\widehat{\mathcal{T}} = \mathcal{T}'.$$

Thus, the conclusion is: The change of variables $y = Q(x)$ that gives privileged coordinates is such that the series $Q(\mathcal{S})$ is of triangular form. Hence, in practice, when constructing privileged coordinates, we operate only with the series $\mathcal{S}$.

4.7. Description of all privileged coordinates. Along with $\mathcal{S}$, let us consider the vector function $\Phi:\mathbb{R}^n\to\mathbb{R}^n$ of the form

$$\Phi(z) = \sum_{\substack{q_1,\dots,q_n\ge 0\\ q_1+\dots+q_n\ge 1}}\frac{1}{q_1!\cdots q_n!}c(\ell_1^{q_1}\cdots\ell_n^{q_n})\,z_1^{q_1}\cdots z_n^{q_n},\qquad z\in\mathbb{R}^n.$$

Notice that $\Phi(0) = 0$. Since the vector fields $X_1, \ldots, X_m$ are real analytic in a neighborhood of the origin, $\Phi(z)$ is also real analytic in a neighborhood of the origin. Moreover, $\frac{\partial \Phi(0)}{\partial z_i} = c(\ell_i) = e_i$, $i = 1, \ldots, n$. Hence, $\Phi(z)$ is locally invertible in a neighborhood of the origin.

Recall that $w_i = \operatorname{ord}(\ell_i)$, $i = 1, \ldots, n$ (weights of coordinates), and $w_1 \leq \cdots \leq w_n$.

THEOREM 4.19. *A nonsingular real analytic change of variables $y = Q(x)$ gives privileged coordinates if and only if it reduces the vector function $\Phi(z)$ to a triangular form, i.e.,*

$$(Q(\Phi(z)))_i = \sum_{w_1 r_1 + \cdots + w_n r_n \geq w_i} \alpha_i^{r_1 \ldots r_n} z_1^{r_1} \cdots z_n^{r_n}, \qquad i = 1, \ldots, n,$$

where $\alpha_i^{r_1 \ldots r_n} \in \mathbb{R}$.

Proof. The proof is clear: Since the map Q acts on $\Phi(z)$ and on $\mathcal{S}$ similarly, we get

$$(Q(\mathcal{S}))_i = \sum_{w_1 r_1 + \cdots + w_n r_n \geq w_i} \alpha_i^{r_1 \ldots r_n} d_1^{\sqcup r_1} \sqcup \cdots \sqcup d_n^{\sqcup r_n}, \qquad i = 1, \ldots, n,$$

which coincides with (4.30) up to a linear map. ∎

In particular, the map $Q(z) = \Phi^{-1}(z)$ defines privileged coordinates; the corresponding approximation was described in [24]. However, it is not easy to find the explicit form of this transformation. On the other hand, in order to reduce $\Phi(z)$ to a triangular form, we need to transform only terms of order no greater than w_n. Hence, any map that reduces $\Phi(z)$ to a triangular form, also reduces the *polynomial* vector function

$$\widehat{\Phi}(z) = \sum_{w_1 q_1 + \cdots + w_n q_n \leq w_n} \frac{1}{q_1! \cdots q_n!} c(\ell_1^{q_1} \cdots \ell_n^{q_n}) z_1^{q_1} \cdots z_n^{q_n}, \qquad z \in \mathbb{R}^n, \tag{4.33}$$

to a triangular form, and vice versa. As a consequence, we obtain the following "finite" description of all privileged coordinates.

THEOREM 4.20. *A nonsingular real analytic change of variables $y = Q(x)$ gives privileged coordinates if and only if it reduces the polynomial vector function* (4.33) *to a triangular form.*

Thus, $Q(x)$ can be chosen in a polynomial form; in essence, such a way is described in [6]. Notice that in Subsection 4.3 we describe a close polynomial transformation. For practical purposes, it is convenient to construct a change of variables componentwise, so that at the ith step we transform the ith component excluding all terms of order less than w_i.

4.8. Representation theorem and a principal part of the series. Let us summarize the obtained results. In this section we have proved that, for any bracket generating system of the form (2.1), there exists a nonsingular polynomial change of variables that reduces the endpoint map to the form

$$(\mathcal{E}_{\widehat{X}_1, \ldots, \widehat{X}_m})_i = a_i + \text{"elements of order} > w_i\text{"},$$

where $w_i = \operatorname{ord}(a_i)$ is the weight of the coordinate x_i, $i = 1, \ldots, n$. In this sense the set of elements $(a_1, \ldots, a_n)$ is the principal part of the series for the endpoint map $\mathcal{E}_{\widehat{X}_1, \ldots, \widehat{X}_m}$.

As we have proved, a_i can be chosen as elements of the dual basis described above, $a_i = d_i$, $i = 1, \ldots, n$.

THEOREM 4.21. *For any bracket generating* (*real analytic*) *system of the form* (2.1), *there exists a nonsingular polynomial change of variables* $y = Q(x)$ *such that the endpoint map of the system in the new coordinates is represented as a series of the form*

$$(\mathcal{E}_{\widehat{X}_1,\ldots,\widehat{X}_m})_i = d_i + \rho_i, \quad i = 1, \ldots, n, \tag{4.34}$$

where $\rho_i \in \bigoplus_{j=w_i+1}^{\infty} \mathcal{F}^j$, $w_i = \mathrm{ord}(d_i)$, $i = 1, \ldots, n$. *Here* $d_1, \ldots, d_n$ *are elements of the basis* (4.20) *dual to the Poincaré–Birkhoff–Witt basis* (4.4), *where the homogeneous elements* $\ell_1, \ldots, \ell_n \in \mathcal{L}$ *are such that* (4.2) *and* (4.3) *hold, and* $\{\ell_j\}_{j=n+1}^{\infty}$ *is a homogeneous basis of* $\mathcal{L}_{X_1,\ldots,X_m}$.

However, the principal part of the series (and therefore privileged coordinates) is not uniquely defined; for example, it can be chosen also as $a_i = d_i + P_i(d_1, \ldots, d_{i-1})$, where $P_i(d_1, \ldots, d_{i-1}) \in \mathcal{F}^{w_i}$ are homogeneous polynomials without linear terms. Using Corollary 4.15, we get another convenient form for the principal part.

THEOREM 4.22. *For any bracket generating* (*real analytic*) *system of the form* (2.1), *there exists a nonsingular polynomial change of variables* $y = \Psi(x)$ *such that the endpoint map of the system in the new coordinates is represented as a series of the form*

$$(\mathcal{E}_{\widehat{X}_1,\ldots,\widehat{X}_m})_i = \widetilde{\ell}_i + \widehat{\rho}_i, \quad i = 1, \ldots, n, \tag{4.35}$$

where $\widehat{\rho}_i \in \bigoplus_{j=w_i+1}^{\infty} \mathcal{F}^j$, $w_i = \mathrm{ord}(\ell_i)$, $i = 1, \ldots, n$. *Here the homogeneous elements* $\ell_i \in \mathcal{L}$ *are such that* (4.2) *and* (4.3) *hold, and* $\widetilde{\ell}_i$ *denotes the orthogonal projection of* ℓ_i *on the subspace* $\mathcal{J}^{\perp}_{X_1,\ldots,X_m}$.

Proof. Suppose a change of variables $y = Q(x)$ reduces the series for the endpoint map $\mathcal{E}_{X_1,\ldots,X_m}$ to the form (4.34). Due to Corollary 4.15, any $\widetilde{\ell}_i$, $i = 1, \ldots, n$, can be expressed as a shuffle polynomial of $d_1, \ldots, d_n$, and vice versa. More specifically,

$$\widetilde{\ell}_i = P_i(d_1, \ldots, d_n) = \sum_{w_j = w_i} \alpha^i_j d_j + \sum_{\substack{q_1+\cdots+q_n \geq 2 \\ q_1 w_1 + \cdots + q_n w_n = w_i}} \alpha^i_{q_1\ldots q_n} d_1^{ш q_1} ш \cdots ш d_n^{ш q_n}, \quad i = 1, \ldots, n,$$

where the matrix $\{\alpha^i_j\}$ is nonsingular. Since $P_i(d_1 + \rho_1, \ldots, d_n + \rho_n) = \widetilde{\ell}_i + \widehat{\rho}_i$, where $\widehat{\rho}_i$ contains terms of order greater than w_i, $i = 1, \ldots, n$, the nonsingular change of variables $y = \Psi(x) = P(Q(x))$ reduces the series to the form (4.35). Obviously, the coordinates y are privileged. ■

Thus, the principal part of the series for the endpoint map can be constructed in a purely algebraic way, by the "standard" procedure of finding the orthogonal projection of elements $\ell_1, \ldots, \ell_n$ satisfying (4.2) and (4.3) on the subspace $\mathcal{J}^{\perp}_{X_1,\ldots,X_m}$.

EXAMPLE 4.23. Consider the system

$$\begin{aligned} \dot{x}_1 &= u_1, \\ \dot{x}_2 &= u_2 + x_1 u_2, \\ \dot{x}_3 &= x_2 u_1 + x_1 u_2 + x_1^2 u_2 + \tfrac{1}{6} x_1^3 u_2. \end{aligned} \tag{4.36}$$

Analogously to Example 2.21, we find the series representation of the endpoint map

$$\mathcal{E}_{X_1,X_2} = \begin{pmatrix} \eta_1 \\ \eta_2 + \eta_{21} \\ \eta_{12} + \eta_{21} + \eta_{121} + 2\eta_{211} + \eta_{2111} \end{pmatrix}.$$

Therefore,

$$c(\eta_1) = c_1 = e_1, \quad c(\eta_2) = c_2 = e_2, \quad c([\eta_2,\eta_1]) = c_{21} - c_{12} = e_2 \in \mathrm{Lin}\{e_1, e_2\},$$
$$c([[\eta_2,\eta_1],\eta_1]) = c_{211} - 2c_{121} + c_{112} = 0, \quad c([[\eta_2,\eta_1],\eta_2]) = 2c_{212} - c_{122} - c_{221} = 0,$$
$$c\big([[[\eta_2,\eta_1],\eta_1],\eta_1]\big) = c_{2111} - 3c_{1211} + 3c_{1121} - c_{1112} = e_3 \notin \mathrm{Lin}\{e_1, e_2\},$$

and all other brackets vanish. Hence,

$$\mathcal{P}^1 = \{0\}, \quad \mathcal{P}^2 = \{[\eta_2,\eta_1]\} = \mathcal{L}^2, \quad \mathcal{P}^3 = \mathrm{Lin}\big\{[[\eta_2,\eta_1],\eta_1], [[\eta_2,\eta_1],\eta_2]\big\} = \mathcal{L}^3,$$
$$\mathcal{P}^4 = \mathrm{Lin}\big\{[[[\eta_2,\eta_1],\eta_1],\eta_2], [[[\eta_2,\eta_1],\eta_2],\eta_2]\big\},$$

and $\mathcal{P}^k = \mathcal{L}^k$ for $k \geq 5$. Then $\mathcal{L}_{X_1,X_2} = \sum_{k=1}^{\infty} \mathcal{P}^k$. We may choose

$$\ell_1 = \eta_1, \quad \ell_2 = \eta_2, \quad \ell_3 = [[[\eta_2,\eta_1],\eta_1],\eta_1].$$

Then $\mathcal{L} = \mathrm{Lin}\{\ell_1, \ell_2, \ell_3\} + \mathcal{L}_{X_1,X_2}$. Thus, for any choice of a basis of $\mathcal{L}_{X_1,X_2}$,

$$d_1 = \eta_1, \quad d_2 = \eta_2, \quad d_3 = \eta_{2111}.$$

For definiteness, set $\ell_4 = [\eta_2, \eta_1]$. Then $d_4 = \eta_{21}$. Rewriting $\mathcal{E}_{X_1,X_2}$ in the form (4.26), we get

$$\mathcal{E}_{X_1,X_2} = \begin{pmatrix} d_1 \\ d_2 + d_4 \\ d_1 \sqcup\!\sqcup\, d_2 + d_1 \sqcup\!\sqcup\, d_4 + d_3 \end{pmatrix}.$$

Since $c(\ell_i) = e_i$, $i = 1, 2, 3$, the initial coordinates are linearly adapted. It is explained in Subsection 4.6 that for the first and second coordinates, the order equals the weight ($\mathrm{ord}(d_1) = \mathrm{ord}(d_2) = 1$). However, the weight of the third coordinate equals $\mathrm{ord}(d_3) = 4$ while its order equals $\mathrm{ord}(d_1 \sqcup\!\sqcup\, d_2) = 2$. Hence, these coordinates are not privileged.

Let us find privileged coordinates, following the way proposed in Subsection 4.7. Let us rewrite $\mathcal{E}_{X_1,X_2}$ in the form (4.27)–(4.29). We get $\mathcal{E}_{X_1,X_2} = \mathcal{S} + \mathcal{T}$, where

$$\mathcal{S} = \begin{pmatrix} d_1 \\ d_2 \\ d_1 \sqcup\!\sqcup\, d_2 + d_3 \end{pmatrix}, \quad \mathcal{T} = \begin{pmatrix} 0 \\ d_4 \\ d_1 \sqcup\!\sqcup\, d_4 \end{pmatrix}.$$

Though $\mathrm{ord}(d_1 \sqcup\!\sqcup\, d_4) < \mathrm{ord}(d_3)$, we ignore the term containing d_4 in the third line of $\mathcal{E}_{X_1,X_2}$ and find privileged coordinates only by use of the form of $\mathcal{S}$, which includes only d_1, d_2, and d_3. Namely, any change of variables that reduces the vector function

$$\Phi(z) = \begin{pmatrix} z_1 \\ z_2 \\ z_1 z_2 + z_3 \end{pmatrix}$$

to a triangular form gives privileged coordinates. For example, we may choose

$$y = Q(x) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 - x_1x_2 \end{pmatrix}.$$

Then

$$\mathcal{E}_{\widehat{X}_1,\widehat{X}_2} = Q(\mathcal{E}_{X_1,X_2}) = \begin{pmatrix} d_1 \\ d_2 + d_4 \\ d_3 \end{pmatrix}.$$

In this case the principal part is defined by the elements d_1, d_2, d_3.

The map

$$y = Q(x) = \begin{pmatrix} x_1 + x_3 \\ x_2 + x_2^2 \\ x_3 - x_1x_2 + x_1^4 \end{pmatrix}$$

also reduces the vector function Φ to a triangular form. It also defines privileged coordinates; in this case we get

$$\mathcal{E}_{\widehat{X}_1,\widehat{X}_2} = Q(\mathcal{E}_{X_1,X_2}) = \begin{pmatrix} d_1 + d_1 ш d_2 + d_1 ш d_4 + d_3 \\ d_2 + d_4 + (d_2 + d_4)^{ш2} \\ d_3 + d_1^{ш4} \end{pmatrix},$$

and the principal part is defined by the elements d_1, d_2, $d_3 + d_1^{ш4}$.

5. Realization problem and algebraic definition of homogeneous approximation

5.1. Approximating system and realizability conditions. In the previous section we obtained the descriptions (4.34) and (4.35) of a principal part of the series representing the endpoint map $\mathcal{E}_{X_1,\dots,X_m}$. Let us consider the "series" $\mathcal{E}$ containing only the principal part, i.e., $(\mathcal{E})_i = d_i$ (or $(\mathcal{E})_i = \widetilde{\ell}_i$), $i = 1, \dots, n$. The question is whether there exists a system (3.1) such that $\mathcal{E} = \mathcal{E}_{Z_1,\dots,Z_m}$, that is, $\mathcal{E}$ is realizable as the endpoint map of some control-linear system. If so, then such a system can be considered as a homogeneous approximation of the initial system (2.1) (see Definition 3.1). More specifically, we are interested in the following version of the realization problem: Given a linear map $c : \mathcal{F} \to \mathbb{R}^n$, determine whether there exists a system of the form (2.1) such that the equalities

$$c(\eta_{i_1\dots i_k}) = X_{i_k} \cdots X_{i_1} E(0) \tag{5.1}$$

hold for any $k \geq 1$ and any $1 \leq i_1, \dots, i_k \leq m$. If this is the case, the series

$$\mathcal{E} = \sum_{k=1}^{\infty} \sum_{1\leq i_1,\dots,i_k\leq m} c(\eta_{i_1\dots i_k})\eta_{i_1\dots i_k}$$

is realized as the endpoint map of this system.

The realization problem in a more general formulation was carefully studied [18, 28, 29, 30], and realizability conditions are well known. Following [50] we formulate a particular case suitable for our purpose.

Theorem 5.1. *Suppose a linear map* $c : \mathcal{F} \to \mathbb{R}^n$ *is such that* $\dim c(\mathcal{L}) = \mathbb{R}^n$. *The realization problem is solvable (i.e., there exists a system of the form* (2.1) *such that equalities* (5.1) *are satisfied) if and only if*

(a) *there exist positive constants* C_1 *and* C_2 *such that*

$$\|c(\eta_{i_1\dots i_k})\| \le k! C_1 C_2^k$$

for any $k \ge 1$ *and any* $1 \le i_1, \dots, i_k \le m$;

(b) *for any* $\ell \in \mathcal{L}$ *such that* $c(\ell) = 0$, *one has* $c(a\ell) = 0$ *for all* $a \in \mathcal{F}$.

Moreover, in this case such a system is unique.

Let us return to our realization problem and consider the series $\mathcal{E}$ such that $(\mathcal{E})_i = d_i$, $i = 1, \dots, n$.

Lemma 5.2. *Suppose* $\mathcal{L}_{X_1,\dots,X_m} \subset \mathcal{L}$ *is a Lie subalgebra corresponding to system* (2.1). *Then the series* $\mathcal{E}$ *such that* $(\mathcal{E})_i = d_i$, $i = 1, \dots, n$, *is realizable, i.e., there exists a system* (3.1) *such that* $\mathcal{E} = \mathcal{E}_{Z_1,\dots,Z_m}$. *Here* $d_1, \dots, d_n$ *are the elements of the basis* (4.20) *dual to the Poincaré–Birkhoff–Witt basis* (4.4), *where the homogeneous elements* $\ell_1, \dots, \ell_n \in \mathcal{L}$ *are such that* (4.2) *and* (4.3) *hold, and* $\{\ell_j\}_{j=n+1}^{\infty}$ *is a homogeneous basis of* $\mathcal{L}_{X_1,\dots,X_m}$.

Proof. Taking into account the representation (4.26), we see that the series $\mathcal{E}$ defines a map $c : \mathcal{F} \to \mathbb{R}^n$ by

$$\begin{aligned} &c(\ell_i) = e_i, && i = 1, \dots, n,\\ &c(\ell_j) = 0, && j \ge n+1,\\ &c(\ell_{j_1} \cdots \ell_{j_r}) = 0, && j_1 \le \dots \le j_r,\ r \ge 2. \end{aligned} \tag{5.2}$$

Condition (a) of Theorem 5.1 is obviously satisfied. Let us prove that condition (b) also holds.

Suppose $\ell \in \mathcal{L}$ is such that $c(\ell) = 0$. Taking into account (5.2), we conclude that $\ell = \sum_{k=1}^{q} \alpha_k \ell_{p_k}$, where $p_1, \dots, p_q \ge n+1$. Now let us choose any $a \in \mathcal{F}$; since (4.4) is a basis of $\mathcal{F}$, it is sufficient to consider $a = \ell_{i_1} \cdots \ell_{i_s}$, where $i_1 \le \dots \le i_s$.

Thus, consider the element $a\ell_{p_k} = (\ell_{i_1} \cdots \ell_{i_s})\,\ell_{p_k}$ with $p_k \ge n+1$. Obviously, $a\ell_{p_k}$ is in $\mathcal{J}_{X_1,\dots,X_m}$. Due to Corollary 4.4, it equals a linear combination of elements (4.7), i.e., elements $\ell_{j_1} \cdots \ell_{j_r}$ with $r \ge 2$ and elements ℓ_{j_1} with $j_1 \ge n+1$. Then (5.2) implies $c(a\ell_{p_k}) = 0$, therefore $c(a\ell) = 0$.

Hence, the map (5.2) is realizable, which means that the series $(\mathcal{E})_i = d_i$, $i = 1, \dots, n$, is realizable as an endpoint map for a certain system. ■

Lemmas 4.14 and 5.2 imply the following corollary.

Corollary 5.3. *Suppose* $\mathcal{L}_{X_1,\dots,X_m} \in \mathcal{L}$ *is a Lie subalgebra corresponding to system* (2.1). *Then the series* $\mathcal{E}$ *such that* $(\mathcal{E})_i = \widetilde{\ell}_i$, $i = 1, \dots, n$, *is realizable, i.e., there exists a system* (3.1) *such that* $\mathcal{E} = \mathcal{E}_{Z_1,\dots,Z_m}$. *Here the homogeneous elements* $\ell_1, \dots, \ell_n \in \mathcal{L}$ *are such that* (4.2) *and* (4.3) *hold, and* $\widetilde{\ell}_i$ *denotes the orthogonal projection of* ℓ_i *on the subspace* $\mathcal{J}^{\perp}_{X_1,\dots,X_m}$.

Now we are ready to describe homogeneous approximations in the sense of Definition 3.1.

LEMMA 5.4. *Let system* (3.1) *be such that* $(\mathcal{E}_{Z_1,\ldots,Z_m})_i = \widetilde{\ell}_i$, $i = 1,\ldots,n$. *This system is a homogeneous approximation for* (2.1) *in the sense of Definition* 3.1.

Proof. Let us check properties (i) and (ii) of Definition 3.1.

(i) The property (4.1) implies $\mathcal{F}^k = (\mathcal{J}_{X_1,\ldots,X_m} \cap \mathcal{F}^k) \oplus^\perp (\mathcal{J}^\perp_{X_1,\ldots,X_m} \cap \mathcal{F}^k)$ for any $k \geq 1$. Set $w_i = \mathrm{ord}(\ell_i)$, $i = 1,\ldots,n$. Then $\ell_i \in \mathcal{F}^{w_i}$ gives $\widetilde{\ell}_i \in \mathcal{F}^{w_i}$, $i = 1,\ldots,n$. Thus, the elements $\widetilde{\ell}_1,\ldots,\widetilde{\ell}_n$ are homogeneous. Hence, the endpoint map $(\mathcal{E}_{Z_1,\ldots,Z_m}(\theta,u))_i = \widetilde{\ell}_i(\theta,u)$, $i = 1,\ldots,n$, satisfies property (i).

(ii) Suppose $y = Q(x)$ defines privileged coordinates such that (4.35) holds. Then, due to (4.35),

$$(Q(\mathcal{E}_{X_1,\ldots,X_m}) - \mathcal{E}_{Z_1,\ldots,Z_m})_i = \widehat{\rho}_i, \qquad i = 1,\ldots,n,$$

where $\widehat{\rho}_i \in \bigoplus_{j=w_i+1}^{\infty} \mathcal{F}^j$, $w_i = \mathrm{ord}(\ell_i)$, $i = 1,\ldots,n$. Hence,

$$\widehat{\rho}_i(\theta, u^{1/\theta}) = \sum_{k=w_i+1}^{\infty} \sum_{1\leq i_1,\ldots,i_k\leq m} (\widehat{c}_{i_1\ldots i_k})_i \eta_{i_1\ldots i_k}(\theta, u^{1/\theta}), \qquad i = 1,\ldots,n,$$

where $\widehat{c}_{i_1\ldots i_k} = \widehat{X}_{i_k}\cdots\widehat{X}_{i_1}E(0)$ and $\widehat{X}_1,\ldots,\widehat{X}_m$ are vector fields in the new coordinates. Taking into account analyticity of $\widehat{X}_i(y)$ and the requirement $u \in B^1$, analogously to (2.6) we get the estimates

$$|\widehat{\rho}_i(\theta, u^{1/\theta})| \leq \widehat{C}_1 \widehat{C}_2^{w_i+1}\theta^{w_i+1}, \qquad i = 1,\ldots,n,$$

for some positive $\widehat{C}_1$, $\widehat{C}_2$ and sufficiently small θ, which implies condition (ii). ∎

COROLLARY 5.5. *System* (3.1) *is a homogeneous approximation for system* (2.1) *in the sense of Definition* 3.1 *if and only if its series is of the form* $(\mathcal{E}_{Z_1,\ldots,Z_m})_i = P_i(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_n)$, $i = 1,\ldots,n$, *where* P *is a polynomial vector function with nonsingular linear part and* $P_i(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_n) \in \mathcal{F}^{w_i}$ (*where* $w_i = \mathrm{ord}(\ell_i)$).

Proof. Suppose $(\mathcal{E}_{Z_1,\ldots,Z_m})_i = P_i(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_n)$, $i = 1,\ldots,n$, where P has nonsingular linear part and $P_i(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_n) \in \mathcal{F}^{w_i}$. As follows from Lemma 5.4, there exists a nonsingular change of variables $Q(x)$ such that $(Q(\mathcal{E}_{X_1,\ldots,X_m}))_i - \widetilde{\ell}_i = \rho_i$, where $\rho_i \in \bigoplus_{j=w_i+1}^{\infty} \mathcal{F}^j$. Then

$$P_i(Q(\mathcal{E}_{X_1,\ldots,X_m})) = P_i(\widetilde{\ell}_1 + \rho_1,\ldots,\widetilde{\ell}_n + \rho_n) = P_i(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_n) + \widehat{\rho}_i = (\mathcal{E}_{Z_1,\ldots,Z_m})_i + \widehat{\rho}_i,$$

where $\widehat{\rho}_i \in \bigoplus_{j=w_i+1}^{\infty} \mathcal{F}^j$, which proves that system (3.1) is a homogeneous approximation for (2.1).

Let now a system $\dot{z} = \sum_{i=1}^m u_i \widehat{Z}_i(z)$ with the series $\widehat{\mathcal{E}} = \mathcal{E}_{\widehat{Z}_1,\ldots,\widehat{Z}_m}$ be another homogeneous approximation of (2.1). Then condition (i) of Definition 3.1 implies that $(\widehat{\mathcal{E}})_i$ are homogeneous; set $\widehat{w}_i = \mathrm{ord}((\widehat{\mathcal{E}})_i)$, $i = 1,\ldots,n$, $\widehat{w}_1 \leq \cdots \leq \widehat{w}_n$.

Now consider condition (ii). It implies that there exists a nonsingular change of variables $\widehat{Q}(x)$ such that $(\widehat{Q}(\mathcal{E}_{X_1,\ldots,X_m}))_i = (\widehat{\mathcal{E}})_i + \widehat{\rho}_i$, where $\widehat{\rho}_i$ contains terms of order greater than $\widehat{w}_i$. On the other hand, Lemma 5.4 implies that there exists a nonsingular change of variables $Q(x)$ such that $(Q(\mathcal{E}_{X_1,\ldots,X_m}))_i - \widetilde{\ell}_i = \rho_i$, where $\rho_i \in \bigoplus_{j=w_i+1}^{\infty} \mathcal{F}^j$.

Denote $\Phi(z) = \widehat{Q}(Q^{-1}(z))$. Then $\Phi(z)$ is a nonsingular change of variables and

$$\Phi_i(\widetilde{\ell}_1 + \rho_1,\ldots,\widetilde{\ell}_n + \rho_n) = \widehat{\mathcal{E}}_i + \widehat{\rho}_i, \qquad i = 1,\ldots,n.$$

Thus,

$$\sum_{r=1}^{\infty}\sum_{j_1+\cdots+j_n=r}\frac{1}{j_1!\cdots j_n!}\frac{\partial^{j_1+\cdots+j_n}\Phi_i(0)}{\partial x_1^{j_1}\cdots\partial x_n^{j_n}}(\widetilde{\ell}_1+\rho_1)^{\sqcup\!\sqcup j_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup(\widetilde{\ell}_n+\rho_n)^{\sqcup\!\sqcup j_n}=\widehat{\mathcal{E}}_i+\widehat{\rho}_i.$$

Hence, the smallest order of elements on both sides of this equality equals $\widehat{w}_i$. Separating elements of this order, we get

$$\sum_{w_1j_1+\cdots+w_nj_n=\widehat{w}_i}\frac{1}{j_1!\cdots j_n!}\frac{\partial^{j_1+\cdots+j_n}\Phi_i(0)}{\partial x_1^{j_1}\cdots\partial x_n^{j_n}}\widetilde{\ell}_1^{\sqcup\!\sqcup j_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup\widetilde{\ell}_n^{\sqcup\!\sqcup j_n}=\widehat{\mathcal{E}}_i,\qquad i=1,\ldots,n.$$

Thus, $\widehat{\mathcal{E}}_i$ is a shuffle polynomial of $\widetilde{\ell}_1,\ldots,\widetilde{\ell}_n$ of order $\widehat{w}_i$. However, $\Phi(z)$ is nonsingular, i.e., the matrix $\frac{\partial\Phi_i(0)}{\partial x_j}$ is nonsingular. Hence, the sets of orders of $\widehat{\mathcal{E}}_i$ and $\widetilde{\ell}_i$ coincide. Taking into account that $w_1\le\cdots\le w_n$ and $\widehat{w}_1\le\cdots\le\widehat{w}_n$, we see that $w_i=\widehat{w}_i$, $i=1,\ldots,n$, which completes the proof. ■

Corollary 5.6. *System* (3.1) *is a homogeneous approximation for* (2.1) *in the sense of Definition* 3.1 *if and only if its series is of the form* $(\mathcal{E}_{Z_1,\ldots,Z_m})_i=P_i(d_1,\ldots,d_n)$, $i=1,\ldots,n$, *where* P *is a polynomial vector function with nonsingular linear part and* P_i *are such that* $P_i(d_1,\ldots,d_n)\in\mathcal{F}^{w_i}$ *(where* $w_i=\mathrm{ord}(d_i)$*).*

Remark 5.7. Corollaries 5.5–5.6 directly imply that if (3.1) is a homogeneous approximation for (2.1) in the sense of Definition 3.1, then $\mathcal{L}_{X_1,\ldots,X_m}=\mathcal{L}_{Z_1,\ldots,Z_m}$.

Thus, the series of a system which is a homogeneous approximation is defined, in essence, uniquely, up to a homogeneous polynomial change of variables. Since the series satisfies conditions (a) and (b) of Theorem 5.1, the approximating system is also defined uniquely. We get the following corollary.

Corollary 5.8. *For a system of the form* (2.1), *the homogeneous approximation exists and is unique, up to a polynomial homogeneous change of variables.*

Finally, let us discuss a connection between two definitions of homogeneous approximation, namely Definition 3.1 and the definition from [6]. Recall that in [6] the concept of homogeneous approximation is introduced in the following way. Suppose system (2.1) is written in privileged coordinates. Then $X_i(x)=X_i^{(-1)}(x)+Y_i(x)$, $i=1,\ldots,m$, where the vector fields $X_i^{(-1)}(x)$ are of order -1, and Y_i consist of terms of order greater than -1. It turns out that in privileged coordinates this is the same as

$$(X_i^{(-1)}(x))_j=\sum_{k_1w_1+\cdots+k_{j-1}w_{j-1}=w_j-1}\mu^{j,i}_{k_1\ldots k_{j-1}}x_1^{k_1}\cdots x_{j-1}^{k_{j-1}},\qquad j=1,\ldots,n,$$

$$(Y_i(x))_j=\sum_{k_1w_1+\cdots+k_nw_n\ge w_j}\nu^{j,i}_{k_1\ldots k_n}x_1^{k_1}\cdots x_n^{k_n},\qquad j=1,\ldots,n.$$

Then the system $\dot z=\sum_{i=1}^m u_iX_i^{(-1)}(z)$ is called a *homogeneous approximation of* (2.1).

It can be shown that this system satisfies Definition 3.1. Let us consider $\mathcal{E}=\mathcal{E}_{X_1,\ldots,X_m}$ and $\widehat{\mathcal{E}}=\mathcal{E}_{X_1^{(-1)},\ldots,X_m^{(-1)}}$. Since

$$(\widehat{\mathcal{E}}(\theta,u))_j = z_j(\theta) = \sum_{i=1}^m \int_0^\theta u_i(\tau)(X_i^{(-1)}(z(\tau)))_j\,d\tau,$$
$$(\mathcal{E}(\theta,u))_j = x_j(\theta) = \sum_{i=1}^m \int_0^\theta u_i(\tau)\big((X_i^{(-1)}(x(\tau)))_j + (Y_i(x(\tau)))_j\big)\,d\tau,$$

we have

$$(\widehat{\mathcal{E}})_j = \sum_{i=1}^m \eta_i\Big(\sum_{k_1w_1+\cdots+k_{j-1}w_{j-1}=w_j-1} \mu^{j,i}_{k_1\ldots k_{j-1}}(\widehat{\mathcal{E}})_1^{\sqcup\!\sqcup k_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup(\widehat{\mathcal{E}})_{j-1}^{\sqcup\!\sqcup k_{j-1}}\Big),$$
$$(\mathcal{E})_j = \sum_{i=1}^m \eta_i\Big(\sum_{k_1w_1+\cdots+k_{j-1}w_{j-1}=w_j-1} \mu^{j,i}_{k_1\ldots k_{j-1}}(\mathcal{E})_1^{\sqcup\!\sqcup k_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup(\mathcal{E})_{j-1}^{\sqcup\!\sqcup k_{j-1}}$$
$$+\sum_{k_1w_1+\cdots+k_nw_n\geq w_j} \nu^{j,i}_{k_1\ldots k_n}(\mathcal{E})_1^{\sqcup\!\sqcup k_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup(\mathcal{E})_n^{\sqcup\!\sqcup k_n}\Big).$$

Using induction on j, it is easy to show that $(\widehat{\mathcal{E}})_j$ is homogeneous and contains elements of order w_j only while $(\mathcal{E})_j$ contains elements of order no less than w_j, and, moreover, elements of order w_j in $(\widehat{\mathcal{E}})_j$ and $(\mathcal{E})_j$ coincide. This implies Definition 3.1. As follows from Corollary 5.8, a homogeneous approximation in the sense of Definition 3.1 is unique (up to a polynomial homogeneous change of variables). Hence, Definition 3.1 and the definition of homogeneous approximation in [6] define the same concept.

5.2. Algebraic definition of homogeneous approximation. The definition of homogeneous approximation used above (see Definition 3.1) is coordinate dependent. Now we are ready to reformulate it in a coordinate-free manner.

As was noticed in Remark 5.7, if system (3.1) is a homogeneous approximation for (2.1) then $\mathcal{L}_{X_1,\ldots,X_m} = \mathcal{L}_{Z_1,\ldots,Z_m}$. In turn, this property provides condition (ii) of Definition 3.1. In fact, suppose $\mathcal{L}_{X_1,\ldots,X_m} = \mathcal{L}_{Z_1,\ldots,Z_m}$ and elements $\widetilde{\ell}_1,\ldots,\widetilde{\ell}_n$ are chosen as in Lemma 5.4. Arguing as in the proof of Lemma 5.4, for both systems we see that there exist Q_1 and Q_2 (privileged coordinates for these systems) such that

$$(Q_1(\mathcal{E}_{X_1,\ldots,X_m}) - Q_2(\mathcal{E}_{Z_1,\ldots,Z_m}))_i = \rho_i,\quad i=1,\ldots,n,$$

where $\rho_i\in\bigoplus_{j=w_i+1}^\infty \mathcal{F}^j$, $w_i=\mathrm{ord}(\ell_i)$, and moreover ρ_i satisfies the estimate $|\rho_i(\theta,u^{1/\theta})| \leq C_1C_2^{w_i+1}\theta^{w_i+1}$, $i=1,\ldots,n$. Hence,

$$\big(Q_2^{-1}(Q_1(\mathcal{E}_{X_1,\ldots,X_m})) - \mathcal{E}_{Z_1,\ldots,Z_m}\big)_i = \overline{\rho}_i,$$

where $\overline{\rho}_i\in\bigoplus_{j=w_i+1}^\infty \mathcal{F}^j$ also satisfies the estimate $|\overline{\rho}_i(\theta,u^{1/\theta})| \leq \overline{C}_1\overline{C}_2^{w_i+1}\theta^{w_i+1}$, $i=1,\ldots,n$. This obviously gives condition (ii) of Definition 3.1.

Now let us turn to condition (i) of Definition 3.1. It can be interpreted in the following way. Denote by $c_{Z_1,\ldots,Z_m}:\mathcal{F}\to\mathbb{R}^n$ the linear map defined as $c_{Z_1,\ldots,Z_m}(\eta_{i_1\ldots i_k}) = Z_{i_k}\cdots Z_{i_1}E(0)$. Suppose $\mathcal{E}_{Z_1,\ldots,Z_m} = \mathcal{S}+\mathcal{T}$ is a decomposition considered in Subsection 4.5. If $\mathcal{T}$ is nontrivial then $\mathcal{E}_{Z_1,\ldots,Z_m}$ is not homogeneous due to Lemma 4.18. On the other hand, $\mathcal{T}$ is trivial if and only if $c_{Z_1,\ldots,Z_m}(\mathcal{J}_{Z_1,\ldots,Z_m})=0$ or, what is the same, $c_{Z_1,\ldots,Z_m}(\mathcal{L}_{Z_1,\ldots,Z_m})=0$. If this is the case, then $\mathcal{E}_{Z_1,\ldots,Z_m}=\mathcal{S}$, and therefore $\mathcal{E}_{Z_1,\ldots,Z_m}$ can be reduced to the form $(Q(\mathcal{E}_{Z_1,\ldots,Z_m}))_i = d_i$, $i=1,\ldots,n$, which satisfies condition (i).

Hence, $c_{Z_1,\dots,Z_m}(\mathcal{L}_{Z_1,\dots,Z_m}) = 0$ if and only if after some change of variables system (3.1) satisfies condition (i) of Definition 3.1.

Thus, we get the following coordinate-free definition, which is equivalent to Definition 3.1.

Definition 5.9. Consider a bracket generating control-linear system of the form (2.1). Let (3.1) be a (bracket generating) system; denote by $c_{Z_1,\dots,Z_m} : \mathcal{F} \to \mathbb{R}^n$ the linear map defined as $c_{Z_1,\dots,Z_m}(\eta_{i_1\dots i_k}) = Z_{i_k}\cdots Z_{i_1}E(0)$. System (3.1) is called a *homogeneous approximation* for (2.1) if

(i) $c_{Z_1,\dots,Z_m}(\mathcal{L}_{Z_1,\dots,Z_m}) = 0$;
(ii) $\mathcal{L}_{X_1,\dots,X_m} = \mathcal{L}_{Z_1,\dots,Z_m}$.

Remark 5.10. Conditions (i) and (ii) of Definition 5.9 can be replaced by the equivalent conditions

(i′) $c_{Z_1,\dots,Z_m}(\mathcal{J}_{Z_1,\dots,Z_m}) = 0$;
(ii′) $\mathcal{J}_{X_1,\dots,X_m} = \mathcal{J}_{Z_1,\dots,Z_m}$.

5.3. Construction of approximating systems. In this subsection we give a convenient method for constructing an approximating system. For example, let us construct system (3.1) so that $(\mathcal{E}_{Z_1,\dots,Z_m})_i = \widetilde{\ell}_i$, $i = 1,\dots,n$.

We act by induction on $i = 1,\dots,n$. For $i = 1$, consider the element $\widetilde{\ell}_1 = \sum_{j=1}^m \alpha_j^1\eta_j$. Define the first component of the vector fields $Z_1,\dots,Z_m$ as follows:

$$(Z_j)_1(z) = \alpha_j^1, \qquad j = 1,\dots,m.$$

Then the function

$$z_1(t) = \widetilde{\ell}_1(t,u) = \sum_{j=1}^m \alpha_j^1\eta_j(t,u) = \sum_{j=1}^m \alpha_j^1 \int_0^t u_j(\tau)\,d\tau$$

satisfies

$$\dot z_1(t) = \sum_{j=1}^m \alpha_j^1 u_j(t) = \sum_{j=1}^m u_j(t)(Z_j)_1.$$

Suppose $2 \le i \le n$. Then after $i-1$ steps all components $(Z_j)_1,\dots,(Z_j)_{i-1}$ are chosen so that the functions

$$z_q(t) = \widetilde{\ell}_q(t,u), \quad q = 1,\dots,i-1,$$

satisfy the differential equalities

$$\dot z_q(t) = \sum_{j=1}^m u_j(t)(Z_j)_q\big(z_1(t),\dots,z_{q-1}(t)\big), \qquad q = 1,\dots,i-1.$$

At the ith step we consider the element $\widetilde{\ell}_i$. Since $\ell_i \in \mathcal{F}^{w_i}$, we get

$$\widetilde{\ell}_i = \sum_{1\le i_1,\dots,i_k\le m} \alpha^i_{i_1\dots i_k}\eta_{i_1\dots i_k}, \qquad \alpha^i_{i_1\dots i_k} \in \mathbb{R},\ k = w_i.$$

If $k = 1$ then $\widetilde{\ell}_i = \sum_{j=1}^m \alpha_j^i\eta_j$. Then we define the ith component of the vector fields $Z_1,\dots,Z_m$ as follows:

$$(Z_j)_i(z) = \alpha_j^i, \qquad j = 1,\dots,m.$$

Thus the function

$$z_i(t) = \widetilde{\ell}_i(t,u) = \sum_{j=1}^m \alpha_j^i \eta_j(t,u) = \sum_{j=1}^m \alpha_j^i \int_0^t u_j(\tau)\,d\tau$$

satisfies

$$\dot{z}_i(t) = \sum_{j=1}^m \alpha_j^i u_j(t) = \sum_{j=1}^m u_j(t)(Z_j)_i.$$

Suppose $k \geq 2$. Then rewrite $\widetilde{\ell}_i$ as

$$\widetilde{\ell}_i = \sum_{1\leq i_1,\ldots,i_k\leq m} \alpha^i_{i_1\ldots i_k}\eta_{i_1\ldots i_k} = \sum_{1\leq i_1,\ldots,i_k\leq m} \alpha^i_{i_1\ldots i_k}\eta_{i_1}\eta_{i_2\ldots i_k} = \sum_{j=1}^m \eta_j a_j,$$

where

$$a_j = \sum_{1\leq i_2,\ldots,i_k\leq m} \alpha^i_{ji_2\ldots i_k}\eta_{i_2\ldots i_k} \in \mathcal{F}^{k-1}.$$

Let us show that $a_j \in \mathcal{J}^{\perp}_{X_1,\ldots,X_m}$. In fact, if $\langle a_j, a\rangle \neq 0$ for some $a \in \mathcal{J}_{X_1,\ldots,X_m}$ then

$$\langle a_j, a\rangle = \langle \eta_j a_j, \eta_j a\rangle = \langle \widetilde{\ell}_i, \eta_j a\rangle \neq 0,$$

where $\eta_j a \in \mathcal{J}_{X_1,\ldots,X_m}$, while $\widetilde{\ell}_i \in \mathcal{J}^{\perp}_{X_1,\ldots,X_m}$. Hence, $a_j \in \mathcal{J}^{\perp}_{X_1,\ldots,X_m}$.

Notice that $\operatorname{ord}(a_j) < \operatorname{ord}(\widetilde{\ell}_i)$. Then taking into account Theorem 4.10, we express a_j as a (homogeneous) shuffle polynomial of $\widetilde{\ell}_1,\ldots,\widetilde{\ell}_{i-1}$,

$$a_j = P_j(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_{i-1}) = \sum_{w_1q_1+\cdots+w_{i-1}q_{i-1}=k-1} \gamma_j^{q_1\cdots q_{i-1}} \widetilde{\ell}_1^{\,\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \widetilde{\ell}_{i-1}^{\,\sqcup\!\sqcup q_{i-1}}.$$

Then we define the ith component of the vector fields $Z_1,\ldots,Z_m$ as follows:

$$(Z_j)_i(z) = P_j(z_1,\ldots,z_{i-1}) = \sum_{w_1q_1+\cdots+w_{i-1}q_{i-1}=k-1} \gamma_j^{q_1\cdots q_{i-1}} z_1^{q_1}\cdots z_{i-1}^{q_{i-1}}, \qquad j=1,\ldots,m.$$

Therefore, we get

$$z_i(t) = \widetilde{\ell}_i(t,u) = \sum_{j=1}^m \big(\eta_j P_j(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_{i-1})\big)(t,u) = \sum_{j=1}^m \int_0^t u_j(\tau)P_j(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_{i-1})(\tau,u)\,d\tau.$$

Recall that, due to the definition of shuffle product,

$$P_j(\widetilde{\ell}_1,\ldots,\widetilde{\ell}_{i-1})(\tau,u) = P_j\big(\widetilde{\ell}_1(\tau,u),\ldots,\widetilde{\ell}_{i-1}(\tau,u)\big),$$

where on the left hand side we consider P_j as a shuffle polynomial, while on the right hand side we consider P_j as a usual polynomial of $i-1$ variables. Hence,

$$z_i(t) = \widetilde{\ell}_i(t,u) = \sum_{j=1}^m \int_0^t u_j(\tau)P_j\big(\widetilde{\ell}_1(\tau,u),\ldots,\widetilde{\ell}_{i-1}(\tau,u)\big)\,d\tau.$$

Therefore, due to the induction supposition,

$$\dot z_i(t)=\sum_{j=1}^m u_j(t)P_j\big(\widetilde\ell_1(t,u),\dots,\widetilde\ell_{i-1}(t,u)\big)=\sum_{j=1}^m u_j(t)P_j\big(z_1(t),\dots,z_{i-1}(t)\big)$$
$$=\sum_{j=1}^m u_j(t)Z_j\big(z_1(t),\dots,z_{i-1}(t)\big).$$

By induction, after n steps we construct the polynomial vector fields $Z_1,\dots,Z_m$ such that the trajectory $z(t)$ satisfying the Cauchy problem $\dot z=\sum_{j=1}^m u_j(t)Z_j(z)$, $z(0)=0$ (for arbitrary fixed controls $u_1(t),\dots,u_m(t)$), is such that $z_i(t)=\widetilde\ell_i(t,u)$, $i=1,\dots,n$. Recall that we denote $z(t)=\mathcal{E}_{Z_1,\dots,Z_m}(t,u)$. Thus, the vector fields $Z_1,\dots,Z_m$ are such that $\mathcal{E}_{Z_1,\dots,Z_m}=\widetilde\ell_i$, $i=1,\dots,n$.

Analogously, the polynomial vector fields can be found such that $\mathcal{E}_{Z_1,\dots,Z_m}=d_i$, $i=1,\dots,n$.

Remark 5.11. Suppose $\mathcal{L}'\subset\mathcal{L}$ is an arbitrary graded Lie subalgebra of codimension n. Set $\mathcal{J}'=\mathrm{Lin}\{\mathcal{F}^e\mathcal{L}'\}$, choose any homogeneous elements $\ell_1,\dots,\ell_n$ such that $\mathcal{L}=\mathrm{Lin}\{\ell_1,\dots,\ell_n\}+\mathcal{L}'$, and denote by $\widetilde\ell_i$ the orthoprojection of ℓ_i on the subspace $\mathcal{J}'^\perp$. Then all the results of Subsections 4.1 and 4.2 (naturally, except Lemma 4.2) can be repeated for $\mathcal{L}'$ and $\mathcal{J}'$; in particular, the analog of Theorem 4.10 holds. Hence, following the arguments of the present subsection, we can construct a (homogeneous) system of the form (2.1) such that $\mathcal{L}_{X_1,\dots,X_m}=\mathcal{L}'$. This means that a core Lie subalgebra can be an arbitrary graded Lie subalgebra of codimension n. Along with Lemma 2.29, this gives a complete algebraic classification of possible homogeneous approximations.

Example 5.12. Suppose $\mathcal{L}$ is a free Lie algebra generated by the elements η_1 and η_2. Set $\mathcal{L}'=\sum_{k=1}^\infty\mathcal{P}^k$, where

$$\mathcal{P}^1=\mathrm{Lin}\{\eta_2\},\quad \mathcal{P}^2=\{0\},\quad \mathcal{P}^3=\mathrm{Lin}\{[[\eta_2,\eta_1],\eta_2]\},$$
$$\mathcal{P}^4=\mathrm{Lin}\{[[[\eta_2,\eta_1],\eta_2],\eta_2]\},$$

and $\mathcal{P}^k=\mathcal{L}^k$ for $k\ge5$. Then $\mathcal{L}'$ is a Lie subalgebra of codimension $n=5$. Choose

$$\ell_1=\eta_1,\quad \ell_2=[\eta_2,\eta_1],\quad \ell_3=[[\eta_2,\eta_1],\eta_1],$$
$$\ell_4=-[[[\eta_2,\eta_1],\eta_1],\eta_2],\quad \ell_5=[[[\eta_2,\eta_1],\eta_1],\eta_1].$$

Then $\mathcal{L}=\mathrm{Lin}\{\ell_1,\dots,\ell_5\}+\mathcal{L}'$. Now, set $\mathcal{J}'=\mathrm{Lin}\{\mathcal{F}^e\mathcal{L}'\}$, and find $\widetilde\ell_i$, $i=1,\dots,5$. Obviously, $\widetilde\ell_1=\eta_1$. Since

$$\mathcal{J}'\cap\mathcal{F}^2=\mathrm{Lin}\{\eta_{12},\eta_{22}\},$$

we get $\widetilde\ell_2=\eta_{21}$. The subspace $\mathcal{J}'\cap\mathcal{F}^3$ is defined by all elements of the form $\eta_{i_1i_22}$ and $[[\eta_2,\eta_1],\eta_2]$, hence

$$\mathcal{J}'\cap\mathcal{F}^3=\mathrm{Lin}\{\eta_{112},\eta_{122},\eta_{212},\eta_{222},\eta_{221}\};$$

this implies $\widetilde\ell_3=\eta_{211}-2\eta_{121}$. Finally,

$$\mathcal{J}'\cap\mathcal{F}^4=\mathrm{Lin}\{\eta_{1112},\eta_{1122},\eta_{1212},\eta_{1222},\eta_{1221},\eta_{2112},\eta_{2122},\eta_{2212},\eta_{2222},\eta_{2221}\},$$

which gives $\widetilde\ell_4=\eta_{2211}-2\eta_{2121}$ and $\widetilde\ell_5=\eta_{2111}-3\eta_{1211}+3\eta_{1121}$.

Now let us construct a system

$$\dot{z} = u_1 Z_1(z) + u_2 Z_2(z)$$

such that $(\mathcal{E}_{Z_1,Z_2})_i = \widetilde{\ell}_i$, $i = 1, \ldots, 5$, i.e.,

$$\mathcal{E}_{Z_1,Z_2} = \begin{pmatrix} \eta_1 \\ \eta_{21} \\ \eta_{211} - 2\eta_{121} \\ \eta_{2211} - 2\eta_{2121} \\ \eta_{2111} - 3\eta_{1211} + 3\eta_{1121} \end{pmatrix}, \tag{5.3}$$

as is explained in Subsection 5.3.

Since $\widetilde{\ell}_1 = \eta_1$, we set $(Z_1)_1 = 1$ and $(Z_2)_1 = 0$.

Rewrite $\widetilde{\ell}_2$ as $\widetilde{\ell}_2 = \eta_{21} = \eta_2\eta_1 = \eta_2\widetilde{\ell}_1$. Hence, $(Z_1)_2 = 0$ and $(Z_2)_2 = z_1$.

Rewrite $\widetilde{\ell}_3$ as $\widetilde{\ell}_3 = \eta_{211} - 2\eta_{121} = \eta_2\eta_{11} - 2\eta_1\eta_{21}$. Since $\eta_{11} = \frac{1}{2}\eta_1^{\sqcup\!\sqcup 2} = \frac{1}{2}\widetilde{\ell}_1^{\sqcup\!\sqcup 2}$ and $\eta_{21} = \widetilde{\ell}_2$, we set $(Z_1)_3 = -2z_2$ and $(Z_2)_3 = \frac{1}{2}z_1^2$.

Analogously, $\widetilde{\ell}_4 = \eta_{2211} - 2\eta_{2121} = \eta_2(\eta_{211} - 2\eta_{121}) = \eta_2\widetilde{\ell}_3$. Hence, $(Z_1)_4 = 0$ and $(Z_2)_4 = z_3$.

Finally, $\widetilde{\ell}_5 = \eta_{2111} - 3\eta_{1211} + 3\eta_{1121} = \eta_2\eta_{111} - 3\eta_1(\eta_{211} - \eta_{121})$. Notice $\eta_{111} = \frac{1}{6}\widetilde{\ell}_1^{\sqcup\!\sqcup 3}$ and $\eta_{211} - \eta_{121} = \frac{1}{5}(\eta_{21} \sqcup\!\sqcup \eta_1) + \frac{3}{5}(\eta_{211} - 2\eta_{121}) = \frac{1}{5}\widetilde{\ell}_1 \sqcup\!\sqcup \widetilde{\ell}_2 + \frac{3}{5}\widetilde{\ell}_3$. Hence, $(Z_1)_5 = -\frac{3}{5}z_1z_2 - \frac{9}{5}z_3$ and $(Z_2)_5 = \frac{1}{6}z_1^3$.

Thus, we get

$$Z_1(z) = \begin{pmatrix} 1 \\ 0 \\ -2z_2 \\ 0 \\ -\frac{3}{5}z_1z_2 - \frac{9}{5}z_3 \end{pmatrix}, \quad Z_2(z) = \begin{pmatrix} 0 \\ z_1 \\ \frac{1}{2}z_1^2 \\ z_3 \\ \frac{1}{6}z_1^3 \end{pmatrix},$$

i.e., the system is of the form

$$\begin{aligned} \dot{z}_1 &= u_1, \\ \dot{z}_2 &= z_1u_2, \\ \dot{z}_3 &= -2z_2u_1 + \tfrac{1}{2}z_1^2u_2, \\ \dot{z}_4 &= z_3u_2, \\ \dot{z}_5 &= -\tfrac{3}{5}z_1z_2u_1 - \tfrac{9}{5}z_3u_1 + \tfrac{1}{6}z_1^3u_2. \end{aligned}$$

This system seems to be rather complicated. Let us try to find a simplifying change of variables. Again, consider the endpoint map (5.3). By the change of variables

$$y = Q(z) = \begin{pmatrix} z_1 \\ z_2 \\ \frac{1}{5}(z_3 + 2z_1z_2) \\ \frac{1}{5}(z_4 + z_2^2) \\ \frac{1}{19}\left(z_5 + \frac{21}{10}z_1^2z_2 + \frac{18}{10}z_1z_3\right) \end{pmatrix}$$

the series representation is reduced to the form

$$Q(\mathcal{E}_{Z_1,Z_2}) = \begin{pmatrix} \eta_1 \\ \eta_{21} \\ \eta_{211} \\ \eta_{2211} \\ \eta_{2111} \end{pmatrix}.$$

The system corresponding to this endpoint map can be easily found by use of the described procedure; it is of the form

$$\begin{aligned} \dot y_1 &= u_1, \\ \dot y_2 &= y_1 u_2, \\ \dot y_3 &= \tfrac{1}{2} y_1^2 u_2, \\ \dot y_4 &= y_3 u_2, \\ \dot y_5 &= \tfrac{1}{6} y_1^3 u_2. \end{aligned}$$

6. Homogeneous approximation in a neighborhood

6.1. Coproduct operation and concatenation of trajectories. In this section we deal with the algebra $\mathcal{F}^e = \mathcal{F} + \mathbb{R}$. As before, assume $1 \cdot a = a \cdot 1 = a$ and $1 \sqcup\!\sqcup a = a \sqcup\!\sqcup 1 = a$, for any $a \in \mathcal{F}^e$. Let us extend the inner product to $\mathcal{F}^e$ assuming $\langle 1, 1\rangle = 1$ and $\langle 1, a\rangle = 0$, for any $a \in \mathcal{F}$.

Introduce the tensor product $\mathcal{F}^e \otimes \mathcal{F}^e$ with the basis

$$\{\eta_{i_1\ldots i_k} \otimes \eta_{j_1\ldots j_s} : k, s \geq 0,\ 1 \leq i_1, \ldots, i_k, j_1, \ldots, j_s \leq m\}$$

(as before, we assume $\eta_{q_1\ldots q_r} = 1$ if $r = 0$). Introduce the inner product in $\mathcal{F}^e \otimes \mathcal{F}^e$ assuming this basis is orthonormal. Hence, if $\{b'_q\}_{q=1}^{\infty}$ and $\{b''_q\}_{q=1}^{\infty}$ are dual bases in $\mathcal{F}^e$ then $\{b'_i \otimes b'_j\}_{i,j=1}^{\infty}$ and $\{b''_i \otimes b''_j\}_{i,j=1}^{\infty}$ are dual bases in $\mathcal{F}^e \otimes \mathcal{F}^e$. Therefore, for any $a \in \mathcal{F}^e \otimes \mathcal{F}^e$ one has

$$a = \sum_{i,j=1}^{\infty} \langle a, b'_i \otimes b'_j\rangle b''_i \otimes b''_j. \tag{6.1}$$

Moreover, this identity can be extended to any formal power series a of elements of $\mathcal{F}^e \otimes \mathcal{F}^e$ with vector coefficients.

Now let us introduce the following helpful definition.

Definition 6.1. We say that the linear map $\Delta : \mathcal{F}^e \to \mathcal{F}^e \otimes \mathcal{F}^e$ defined on the basis elements by the rule

$$\Delta(\eta_{i_1\ldots i_k}) = \sum_{j=0}^{k} \eta_{i_1\ldots i_j} \otimes \eta_{i_{j+1}\ldots i_k} \tag{6.2}$$

is the *coproduct* in $\mathcal{F}^e$.

In fact, Δ can be interpreted as a coproduct in the Hopf algebra (see [34], where this operation is denoted by Δ'). By linearity, Δ is naturally extended to formal power series of elements of $\mathcal{F}^e$.

One can easily get the following property of Δ: for any $a, a_1, a_2 \in \mathcal{F}^e$,

$$\langle \Delta(a), a_1 \otimes a_2 \rangle = \langle a, a_1 a_2 \rangle. \tag{6.3}$$

Consequently, if $\{b'_q\}_{q=1}^{\infty}$ and $\{b''_q\}_{q=1}^{\infty}$ are dual bases in $\mathcal{F}^e$, then for any $a \in \mathcal{F}^e$,

$$\Delta(a) = \sum_{i,j=1}^{\infty} \langle a, b'_i b'_j \rangle b''_i \otimes b''_j, \tag{6.4}$$

and this property can be extended to any formal power series a of elements of $\mathcal{F}^e$.

In the following lemma we use the notation of concatenation of controls (2.8).

Lemma 6.2. *Suppose $\{b'_q\}_{q=1}^{\infty}$ and $\{b''_q\}_{q=1}^{\infty}$ are dual bases in $\mathcal{F}^e$. Then for any pair of controls $u^1 \in B^{\theta^1}$, $u^2 \in B^{\theta^2}$ and any $a \in \mathcal{F}^e$ one has*

$$a(\theta^1 + \theta^2, u^1 \circ u^2) = \sum_{i,j=1}^{\infty} \langle a, b'_i b'_j \rangle b''_i(\theta^2, u^2)\, b''_j(\theta^1, u^1), \tag{6.5}$$

and this property can be extended to any formal power series a of elements of $\mathcal{F}^e$.

Proof. Let us consider any pair of controls $u^1 \in B^{\theta^1}$, $u^2 \in B^{\theta^2}$; below for the sake of brevity we denote it as P. For a pair P, let us introduce the linear map $m_P : \mathcal{F}^e \otimes \mathcal{F}^e \to \mathbb{R}$ defined on basis elements $\eta_{i_1\ldots i_k} \otimes \eta_{j_1\ldots j_s}$ by

$$m_P(\eta_{i_1\ldots i_k} \otimes \eta_{j_1\ldots j_s}) = \eta_{i_1\ldots i_k}(\theta^2, u^2)\eta_{j_1\ldots j_s}(\theta^1, u^1).$$

Due to Lemma 2.10, for any $\eta_{i_1\ldots i_k} \in \mathcal{F}$ the following identity holds

$$\eta_{i_1\ldots i_k}(\theta^1 + \theta^2, u^1 \circ u^2) = \sum_{j=0}^{k} \eta_{i_1\ldots i_j}(\theta^2, u^2)\eta_{i_{j+1}\ldots i_k}(\theta^1, u^1), \tag{6.6}$$

where we assume $\eta_{i_p\ldots i_q}(\theta, u) = 1$ for any θ and u if $p > q$. The definitions of Δ and m_P allow us to rewrite (6.6) as

$$\eta_{i_1\ldots i_k}(\theta^1 + \theta^2, u^1 \circ u^2) = m_P(\Delta(\eta_{i_1\ldots i_k})),$$

which, by linearity, implies

$$a(\theta^1 + \theta^2, u^1 \circ u^2) = m_P(\Delta(a))$$

for any $a \in \mathcal{F}^e$. Then (6.4) gives

$$a(\theta^1 + \theta^2, u^1 \circ u^2) = \sum_{i,j=1}^{\infty} \langle a, b'_i b'_j \rangle m_P(b''_i \otimes b''_j) = \sum_{i,j=1}^{\infty} \langle a, b'_i b'_j \rangle b''_i(\theta^2, u^2) b''_j(\theta^1, u^1),$$

which proves the lemma. ■

Let us outline the next step of analysis. Consider a system of the form (2.1). As above, let φ be the natural anti-homomorphism $\varphi : \mathcal{F} \to F$ defined by

$$\varphi(\eta_{i_1\ldots i_k}) = X_{i_k} \cdots X_{i_1}, \qquad k \geq 1,\ 1 \leq i_1, \ldots, i_k \leq m.$$

For any $z \in U(0)$, introduce a linear map $c^z : \mathcal{F} \to \mathbb{R}^n$ defined as

$$c^z(a) = \varphi(a)E(z), \qquad a \in \mathcal{F};$$

in particular, $c^z(\eta_{i_1\dots i_k}) = X_{i_k}\cdots X_{i_1}E(z)$. Analogously to Theorem 2.2, the end point $x(\theta)$ of the solution of the Cauchy problem

$$\dot{x} = \sum_{i=1}^{m} u_i(t)X_i(x), \qquad x(0) = z,$$

can be written as

$$x(\theta) = z + \mathcal{E}^z_{X_1,\dots,X_m}(\theta, u),$$

where the endpoint map from z is expressed as a series of the form

$$\mathcal{E}^z_{X_1,\dots,X_m}(\theta, u) = \sum_{k=1}^{\infty} \sum_{1\le i_1,\dots,i_k\le m} c^z(\eta_{i_1\dots i_k})\eta_{i_1\dots i_k}(\theta, u).$$

Below we mainly deal with the corresponding formal power series

$$\mathcal{E}^z_{X_1,\dots,X_m} = \sum_{k=1}^{\infty} \sum_{1\le i_1,\dots,i_k\le m} c^z(\eta_{i_1\dots i_k})\eta_{i_1\dots i_k}.$$

Suppose the Rashevsky–Chow condition (2.24) holds at the origin. Then without loss of generality it holds at any $z \in U(0)$. This means that

$$\sum_{k=1}^{\infty} c^z(\mathcal{L}^k) = \mathbb{R}^n, \qquad z \in U(0). \tag{6.7}$$

Let us find a connection between coefficients of the series $\mathcal{E}^z_{X_1,\dots,X_m}$ and $\mathcal{E}_{X_1,\dots,X_m}$. Consider an arbitrary point $x \in U(0)$ and a trajectory of system (2.1) going from the origin to x through z. Namely, suppose $u^1 \in B^{\theta^1}$ steers the origin to z and $u^2 \in B^{\theta^2}$ steers z to x. Then $u^1 \circ u^2$ steers the origin to x (at the time $\theta^1 + \theta^2$). This means that

$$x = \mathcal{E}_{X_1,\dots,X_m}(\theta^1 + \theta^2, u^1 \circ u^2) = z + \mathcal{E}^z_{X_1,\dots,X_m}(\theta^2, u^2), \tag{6.8}$$

where

$$z = \mathcal{E}_{X_1,\dots,X_m}(\theta^1, u^1). \tag{6.9}$$

Below we assume z is fixed whereas x is arbitrary.

The question arises whether coefficients of $\mathcal{E}^z_{X_1,\dots,X_m}$ (i.e., $c^z_{i_1\dots i_k}$) can be expressed directly via coefficients of $\mathcal{E}_{X_1,\dots,X_m}$ (i.e., $c_{i_1\dots i_k}$). The answer is "yes" for a class of systems described in Subsection 6.3 below.

In the rest of this section we study core Lie subalgebras and left ideals for $z \in U(0)$. Namely, consider the subspaces

$$\mathcal{P}^k(z) = \{\ell \in \mathcal{L}^k : c^z(\ell) \in c^z(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k-1})\}, \qquad k \ge 1,$$

and set

$$\mathcal{L}^z_{X_1,\dots,X_m} = \bigoplus_{k=1}^{\infty} \mathcal{P}^k(z).$$

Set also

$$\mathcal{J}^z_{X_1,\dots,X_m} = \mathrm{Lin}\{\mathcal{F}^e \mathcal{L}^z_{X_1,\dots,X_m}\}.$$

For $z = 0$ we, as a rule, omit the reference to the point, i.e., write $\mathcal{L}_{X_1,\dots,X_m}$ instead of $\mathcal{L}^0_{X_1,\dots,X_m}$, $\mathcal{J}_{X_1,\dots,X_m}$ instead of $\mathcal{J}^0_{X_1,\dots,X_m}$, etc.

6.2. Regular systems. The simplest approximate characteristic of a system in a neighborhood is the behavior of its growth vector. Namely, let p^z be the degree of nonholonomy of the system at the point z. Set

$$v_k^z = \dim\, c^z(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^k), \qquad k = 1, \ldots, p^z.$$

Then the sequence $v^z = (v_1^z, \ldots, v_{p^z}^z)$ is the growth vector of the system at z. Denote by p and $v = (v_1, \ldots, v_p)$ the degree of nonholonomy and the growth vector at the origin. Obviously, there exists a neighborhood $U(0)$ such that for any $z \in U(0)$,

$$p^z \le p \quad \text{and} \quad v_k^z \ge v_k,\ k = 1, \ldots, p^z.$$

Definition 6.3. System (2.1) is called *regular at the origin* if its growth vector is constant in a certain neighborhood $U(0)$, i.e., $p^z = p$ and $v_k^z = v_k$, $k = 1, \ldots, p$, for any $z \in U(0)$. In the opposite case the system is called *nonregular at the origin.*

Lemma 6.4. *Suppose system* (2.1) *is regular at the origin. Then its core Lie subalgebra* $\mathcal{L}_{X_1,\ldots,X_m}$ *is a Lie ideal in* $\mathcal{L}$*, i.e., for any* $a \in \mathcal{L}$ *and any* $\ell \in \mathcal{L}_{X_1,\ldots,X_m}$ *one has* $[a, \ell] \in \mathcal{L}_{X_1,\ldots,X_m}$.

Proof. Suppose elements $\ell_1, \ldots, \ell_n$ are such that

$$\mathcal{L} = \mathrm{Lin}\{\ell_1, \ldots, \ell_n\} + \mathcal{L}_{X_1,\ldots,X_m}, \tag{6.10}$$

and without loss of generality assume $\ell_1, \ldots, \ell_n$ are homogeneous and

$$\mathrm{ord}(\ell_i) \le \mathrm{ord}(\ell_j) \qquad \text{for } i < j. \tag{6.11}$$

As follows from Corollary 2.30, the vectors $c(\ell_1), \ldots, c(\ell_n)$ are linearly independent, therefore vectors $c^x(\ell_1), \ldots, c^x(\ell_n)$ are linearly independent for any x from a certain neighborhood $U(0)$. Without loss of generality assume that the growth vector is constant in $U(0)$, i.e., $p^x = p$ and $v_k^x = v_k$, $k = 1, \ldots, p$. Then for any $x \in U(0)$,

$$c^x(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^k) = \mathrm{Lin}\{c^x(\ell_1), \ldots, c^x(\ell_{v_k})\}, \qquad 1 \le k \le p.$$

Let us consider any $k = 1, \ldots, p$ and any $\ell \in \mathcal{L}_{X_1,\ldots,X_m} \cap \mathcal{L}^k$. The vector $c^x(\ell)$ depends linearly on $c^x(\ell_1), \ldots, c^x(\ell_{v_k})$, i.e., there exist scalar functions $\alpha_i(x)$, $i = 1, \ldots, v_k$, such that

$$c^x(\ell) = \sum_{i=1}^{v_k} \alpha_i(x) c^x(\ell_i).$$

However, $\ell \in \mathcal{L}_{X_1,\ldots,X_m} \cap \mathcal{L}^k$, which implies $c(\ell) \in \mathrm{Lin}\{c(\ell_1), \ldots, c(\ell_{v_{k-1}})\}$. Therefore,

$$\alpha_i(0) = 0, \qquad i = v_{k-1} + 1, \ldots, v_k. \tag{6.12}$$

Since the vectors $c^x(\ell_1), \ldots, c^x(\ell_{v_k})$ are linearly independent, the functions $\alpha_i(x)$, $i = 1, \ldots, v_k$, are smooth.

Now let us consider an arbitrary $a \in \mathcal{L}^q$, $q \geq 1$. We have

$$\begin{aligned}
c^x([a,\ell]) &= (c^x(\ell))'_x c^x(a) - (c^x(a))'_x c^x(\ell) \\
&= \sum_{i=1}^{v_k} \big(\alpha_i(x) c^x(\ell_i)\big)'_x c^x(a) - \sum_{i=1}^{v_k} (c^x(a))'_x \alpha_i(x) c^x(\ell_i) \\
&= \sum_{i=1}^{v_k} \big(\alpha'_i(x) c^x(a)\big) c^x(\ell_i) + \sum_{i=1}^{v_k} \alpha_i(x) \big((c^x(\ell_i))'_x c^x(a) - (c^x(a))'_x c^x(\ell_i)\big) \\
&= \sum_{i=1}^{v_k} \widetilde{\alpha}_i(x) c^x(\ell_i) + \sum_{i=1}^{v_k} \alpha_i(x) c^x([a,\ell_i]),
\end{aligned}$$

where $\widetilde{\alpha}_i(x) = \alpha'_i(x) c^x(a)$, $i = 1, \ldots, v_k$. Taking into account (6.12), at $x = 0$ we get

$$c([a,\ell]) = \sum_{i=1}^{v_k} \widetilde{\alpha}_i(0) c(\ell_i) + \sum_{i=1}^{v_{k-1}} \alpha_i(0) c([a,\ell_i]).$$

However, $\ell_i \in \mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^k$ for $i = 1, \ldots, v_k$, and $[a, \ell_i] \in \mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k+q-1}$ for $i = 1, \ldots, v_{k-1}$. Hence, $c([a,\ell]) \in c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k+q-1})$ whereas $[a,\ell] \in \mathcal{L}^{k+q}$. Therefore, $[a,\ell] \in \mathcal{P}^{k+q} \subset \mathcal{L}_{X_1,\ldots,X_m}$. ■

If a system is regular at the origin, then it is obviously regular at any point from a certain neighborhood of the origin. Hence, we get the following corollary.

Corollary 6.5. *Suppose system* (2.1) *is regular at the origin. Then there exists a neighborhood $U(0)$ such that for any $z \in U(0)$ the core Lie subalgebra $\mathcal{L}^z_{X_1,\ldots,X_m}$ is a Lie ideal in $\mathcal{L}$.*

The condition on $\mathcal{L}_{X_1,\ldots,X_m}$ to be a Lie ideal can be expressed in terms of the left ideal $\mathcal{J}_{X_1,\ldots,X_m}$.

Lemma 6.6. *The core Lie subalgebra $\mathcal{L}_{X_1,\ldots,X_m}$ of system* (2.1) *is a Lie ideal in $\mathcal{L}$ if and only if the left ideal $\mathcal{J}_{X_1,\ldots,X_m}$ is two-sided, i.e., for any $a \in \mathcal{F}$ and any $b \in \mathcal{J}_{X_1,\ldots,X_m}$ one has $ba \in \mathcal{J}_{X_1,\ldots,X_m}$.*

Proof. Suppose $\mathcal{J}_{X_1,\ldots,X_m}$ is two-sided. Choose any $\ell \in \mathcal{L}_{X_1,\ldots,X_m} \subset \mathcal{J}_{X_1,\ldots,X_m}$ and any $a \in \mathcal{L}$. Then $a\ell \in \mathcal{J}_{X_1,\ldots,X_m}$ and $\ell a \in \mathcal{J}_{X_1,\ldots,X_m}$. Hence, using Corollary 4.5, we get $[a,\ell] = a\ell - \ell a \in \mathcal{J}_{X_1,\ldots,X_m} \cap \mathcal{L} = \mathcal{L}_{X_1,\ldots,X_m}$. Therefore, $\mathcal{L}_{X_1,\ldots,X_m}$ is a Lie ideal.

Now, let $\mathcal{L}_{X_1,\ldots,X_m}$ be a Lie ideal. Let us prove that the left ideal $\mathcal{J}_{X_1,\ldots,X_m}$ is two-sided. Obviously, it is sufficient to prove that $\ell a \in \mathcal{J}_{X_1,\ldots,X_m}$ for any $\ell \in \mathcal{L}_{X_1,\ldots,X_m}$ and any $a \in \mathcal{F}$. Moreover, denote

$$M_k = \{\ell \ell_{i_1} \cdots \ell_{i_k} : \ell \in \mathcal{L}_{X_1,\ldots,X_m},\ \ell_{i_1}, \ldots, \ell_{i_k} \in \mathcal{L}\}, \qquad k \geq 1.$$

Due to the Poincaré–Birkhoff–Witt theorem, it is sufficient to prove that $M_k \subset \mathcal{J}_{X_1,\ldots,X_m}$ for all $k \geq 1$.

We argue by induction on k. For $k = 1$ one has $\ell\ell_{i_1} = [\ell, \ell_{i_1}] + \ell_{i_1}\ell$. Since $\mathcal{L}_{X_1,\ldots,X_m}$ is a Lie ideal, $[\ell, \ell_{i_1}] \in \mathcal{L}_{X_1,\ldots,X_m} \subset \mathcal{J}_{X_1,\ldots,X_m}$; since $\mathcal{J}_{X_1,\ldots,X_m}$ is a left ideal, $\ell_{i_1}\ell \in \mathcal{J}_{X_1,\ldots,X_m}$. Hence, $\ell\ell_{i_1} \in \mathcal{J}_{X_1,\ldots,X_m}$, and therefore $M_1 \subset \mathcal{J}_{X_1,\ldots,X_m}$.

Suppose $M_k \subset \mathcal{J}_{X_1,\ldots,X_m}$ for some $k \geq 1$. Choose any element $a \in M_{k+1}$. Then $a = b\,\ell_{i_{k+1}}$, where $\ell_{i_{k+1}} \in \mathcal{L}$ and $b \in M_k$. Hence, $b \in \mathcal{J}_{X_1,\ldots,X_m}$, and therefore it can be written as $b = \sum b_q \ell_{j_q}$, where $b_q \in \mathcal{F}^e$ and $\ell_{j_q} \in \mathcal{L}_{X_1,\dots,X_m}$. Then, analogously to the case $k = 1$, we get

$$a = b\,\ell_{i_{k+1}} = \sum b_q \ell_{j_q} \ell_{i_{k+1}} = \sum b_q [\ell_{j_q}, \ell_{i_{k+1}}] + \sum (b_q \ell_{i_{k+1}}) \ell_{j_q} \in \mathcal{J}_{X_1,\dots,X_m}.$$

Hence, $M_{k+1} \subset \mathcal{J}_{X_1,\dots,X_m}$. ■

Corollary 6.7. *Suppose system* (2.1) *is regular at the origin. Then there exists a neighborhood $U(0)$ such that for any $z \in U(0)$ the left ideal $\mathcal{J}^z_{X_1,\dots,X_m}$ is two-sided, i.e., for any $a \in \mathcal{F}$ and any $b \in \mathcal{J}^z_{X_1,\dots,X_m}$ one has $ba \in \mathcal{J}^z_{X_1,\dots,X_m}$.*

The following example shows that the Lie ideal $\mathcal{L}^z_{X_1,\dots,X_m}$ of a regular system can depend on the point z.

Example 6.8. Consider the system in a neighborhood of the origin

$$\begin{aligned} \dot{x}_1 &= u_1, \\ \dot{x}_2 &= u_2 + x_1^2 u_2, \\ \dot{x}_3 &= x_1 u_2, \\ \dot{x}_4 &= x_1^2 u_2 + x_1 x_2 u_2. \end{aligned}$$

We have

$$X_1(x) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad X_2(x) = \begin{pmatrix} 0 \\ 1 + x_1^2 \\ x_1 \\ x_1^2 + x_1 x_2 \end{pmatrix}, \quad [X_1, X_2](x) = \begin{pmatrix} 0 \\ 2x_1 \\ 1 \\ 2x_1 + x_2 \end{pmatrix},$$

$$[X_1, [X_1, X_2]](x) = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \quad [X_2, [X_1, X_2]](x) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 - x_1^2 \end{pmatrix}.$$

Hence, the growth vector equals $v^x = (2, 3, 4)$ in a neighborhood of the origin, i.e., the system is regular. It is easy to check that

$$[X_2, [X_1, X_2]](x) - \frac{(1 - x_1^2)^2}{2} [X_1, [X_1, X_2]](x) = -(1 - x_1^2)\big(X_2(x) - x_1 [X_1, X_2](x)\big).$$

Thus,

$$\mathcal{P}^1(x) = \mathcal{P}^2(x) = \{0\}, \qquad \mathcal{P}^3(x) = \mathrm{Lin}\big\{[[\eta_2, \eta_1], \eta_2] - \alpha(x)[[\eta_2, \eta_1], \eta_1]\big\}$$

(where $\alpha(x) = (1 - x_1^2)^2/2$ depends on the point x), and $\mathcal{P}^k(x) = \mathcal{L}^k$, $k \geq 4$. Hence, the system is regular (and obviously $\mathcal{L}^x_{X_1,X_2}$ is a Lie ideal) but $\mathcal{L}^x_{X_1,X_2}$ depends on x.

Thus, a core Lie subalgebra of a regular system is not necessarily constant in a neighborhood of the origin.

In the next example we consider a nonregular system whose core Lie subalgebra is a Lie ideal.

Example 6.9. Consider the system in a neighborhood of the origin

$$\begin{aligned}\dot x_1 &= u_1,\\ \dot x_2 &= u_2,\\ \dot x_3 &= x_1u_2,\\ \dot x_4 &= x_1^2u_2,\\ \dot x_5 &= x_1^3u_2 + x_3x_1^2u_2.\end{aligned}$$

We have

$$X_1(x) = \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix},\quad X_2(x) = \begin{pmatrix}0\\1\\x_1\\x_1^2\\x_1^3+x_3x_1^2\end{pmatrix},\quad [X_1,X_2](x) = \begin{pmatrix}0\\0\\1\\2x_1\\3x_1^2+2x_1x_3\end{pmatrix},$$

$$[X_1,[X_1,X_2]](x) = \begin{pmatrix}0\\0\\0\\2\\6x_1+2x_3\end{pmatrix},\quad [X_2,[X_1,X_2]](x) = \begin{pmatrix}0\\0\\0\\0\\x_1^2\end{pmatrix},$$

$$[X_1,[X_1,[X_1,X_2]]](x) = \begin{pmatrix}0\\0\\0\\0\\6\end{pmatrix},\quad [X_1,[X_2,[X_1,X_2]]](x) = \begin{pmatrix}0\\0\\0\\0\\2x_1\end{pmatrix},$$

and $[X_2,[X_2,[X_1,X_2]]](x) = 0$. At $x = 0$, we have

$$X_1(0) = e_1,\quad X_1(0) = e_2,\quad [X_1,X_2](0) = e_3,\quad [X_1,[X_1,X_2]](0) = 2e_4,$$
$$[X_2,[X_1,X_2]](0) = 0,\quad [X_1,[X_1,[X_1,X_2]]](0) = e_5.$$

Hence, the growth vector at the origin equals $v^0 = (2,3,4,5)$. However, $[X_2,[X_1,X_2]](x) = x_1^2e_5$. Hence, for $x_1 \neq 0$ the growth vector equals $v^x = (2,3,5)$. Thus, the system is not regular at the origin.

Let us find its core Lie subalgebra $\mathcal{L}^x_{X_1,X_2}$. Since the system is not regular, $\mathcal{L}^x_{X_1,X_2}$ cannot be constant.

If $x_1 = 0$ (including $x = 0$) then $\mathcal{P}^1(x) = \mathcal{P}^2(x) = \{0\}$, $\mathcal{P}^3(x) = \mathrm{Lin}\{[[\eta_2,\eta_1],\eta_2]\}$, $\mathcal{P}^4(x) = \mathrm{Lin}\{[[[\eta_2,\eta_1],\eta_2],\eta_1],[[[\eta_2,\eta_1],\eta_2],\eta_2]\}$, and $\mathcal{P}^k(x) = \mathcal{L}^k$, $k \geq 5$. Obviously, $\mathcal{L}^x_{X_1,X_2}$ is a Lie ideal.

If $x_1 \neq 0$ then $\mathcal{P}^1(x) = \mathcal{P}^2(x) = \mathcal{P}^3(x) = \{0\}$ and $\mathcal{P}^k(x) = \mathcal{L}^k$, $k \geq 4$. Hence, $\mathcal{L}^x_{X_1,X_2}$ is also a Lie ideal.

Hence, $\mathcal{L}^x_{X_1,X_2}$ is a Lie ideal at any point from a neighborhood of the origin. Thus, even if a core Lie subalgebra is a Lie ideal in a neighborhood, the system can be nonregular.

In the next subsection we show that for homogeneous systems the property of the core Lie subalgebra $\mathcal{L}_{X_1,\ldots,X_m}$ to be a Lie ideal is sufficient for regularity, and moreover implies that the core Lie subalgebra is the same for all points from a neighborhood of the origin.

6.3. Re-expanding the series and regular homogeneous systems. In this subsection we consider homogeneous systems from the point of view of properties of their core Lie subalgebras and series $\mathcal{E}_{X_1,\dots,X_m}$.

Following Definition 5.9, we adopt the following definition of a homogeneous system.

DEFINITION 6.10. A (bracket generating) system of the form (2.1) is called *homogeneous at the origin* if $c(\mathcal{L}_{X_1,\dots,X_m}) = 0$.

As follows from the discussion in Subsection 5.2, a system is homogeneous at the origin in the sense of Definition 6.10 iff there exists a nonsingular mapping $Q(x)$ ($Q(0) = 0$) such that $(Q(\mathcal{E}_{X_1,\dots,X_m}))_k$ is homogeneous for any $k = 1,\dots,n$, i.e., $(Q(\mathcal{E}_{X_1,\dots,X_m}))_k \in \mathcal{F}^{w_k}$, $k = 1,\dots,n$. Suppose the change of variables $y = Q(x)$ is already applied. It follows from Theorem 4.21 that for a homogeneous system without loss of generality we may assume

$$(\mathcal{E}_{X_1,\dots,X_m})_k = d_k, \qquad k = 1,\dots,n, \tag{6.13}$$

where d_k are elements of the dual basis (4.20). Below we have in mind that a homogeneous system can be considered in the whole $\mathbb{R}^n$ rather than in a neighborhood of the origin.

LEMMA 6.11. *Let system* (2.1) *be homogeneous at the origin. Then* $\mathcal{E}^z_{X_1,\dots,X_m}$ *can be found directly, without evaluating nonholonomic derivatives* $X_{i_k}\cdots X_{i_1}E(z)$.

Proof. For brevity, let us denote $\mathcal{E} = \mathcal{E}_{X_1,\dots,X_m}$ and $\mathcal{E}^z = \mathcal{E}^z_{X_1,\dots,X_m}$. Suppose $\{\ell_i\}_{i=1}^\infty$ is a homogeneous basis of $\mathcal{L}$ satisfying (6.10) and (6.11). Let d_k be elements of the dual basis (4.20). Then (4.21) holds. Moreover,

$$\langle d_k, \ell_{j_1}\cdots\ell_{j_r}\rangle = 0 \quad \text{ if } r \ge 2 \text{ and } j_r \ge k. \tag{6.14}$$

In fact, if $j_r \ge n+1$ then $\ell_{j_1}\cdots\ell_{j_r} \in \mathcal{J}_{X_1,\dots,X_m}$. Hence, (6.14) holds due to Lemma 4.13. If $k \le j_r \le n$ then (6.11) implies $\operatorname{ord}(\ell_{j_r}) \ge \operatorname{ord}(\ell_k) = \operatorname{ord}(d_k)$. Since $r \ge 2$, we get $\operatorname{ord}(\ell_{j_1}\cdots\ell_{j_r}) > \operatorname{ord}(d_k)$, which gives (6.14).

Now let us apply Lemma 6.2. Without loss of generality assume (6.13) holds. Taking into account (4.21), (6.5), (6.8), (6.9), and (6.14), we get

$$\begin{aligned}\mathcal{E}^z_k(\theta^2,u^2) &= \mathcal{E}(\theta^1+\theta^2,u^1\circ u^2) - \mathcal{E}(\theta^1,u^1) = d_k(\theta^1+\theta^2,u^1\circ u^2) - d_k(\theta^1,u^1)\\ &= d_k(\theta^2,u^2) + \sum \left\langle d_k, (\ell_{i_1}^{q_1}\cdots\ell_{i_j}^{q_j})(\ell_1^{r_1}\cdots\ell_{k-1}^{r_{k-1}})\right\rangle \frac{\prod_{s=1}^{j} d_{i_s}^{q_s}(\theta^2,u^2)\prod_{s=1}^{k-1} d_s^{r_s}(\theta^1,u^1)}{q_1!\cdots q_j! r_1!\cdots r_{k-1}!},\end{aligned}$$

where the sum is taken over all $j \ge 1$, $i_1 < \cdots < i_j$, $q_1,\dots,q_j \ge 1$, $r_1+\cdots+r_{k-1} \ge 1$ such that

$$\sum_{s=1}^{j}\operatorname{ord}(\ell_{i_s})q_s + \sum_{s=1}^{k-1}\operatorname{ord}(\ell_s)r_s = \operatorname{ord}(\ell_k).$$

Due to (6.9), $d_i(\theta^1,u^1)=\mathcal{E}_i(\theta^1,u^1)=z_i$, $i=1,\dots,n$, hence

$$\mathcal{E}_k^z(\theta^2,u^2)=d_k(\theta^2,u^2)+\sum_{\substack{j\geq1,\, i_1<\cdots<i_j\\ q_1,\dots,q_j\geq1}} P_k^{q_1\dots q_j i_1\dots i_j}(z)\prod_{s=1}^{j} d_{i_s}^{q_s}(\theta^2,u^2), \tag{6.15}$$

where $P_k^{q_1\dots q_j i_1\dots i_j}(z)$ are polynomials of the form

$$P_k^{q_1\dots q_j i_1\dots i_j}(z)=\sum\frac{\big\langle d_k,(\ell_{i_1}^{q_1}\cdots\ell_{i_j}^{q_j})(\ell_1^{r_1}\cdots\ell_{k-1}^{r_{k-1}})\big\rangle}{q_1!\cdots q_j!r_1!\cdots r_{k-1}!}\prod_{s=1}^{k-1}z_s^{r_s}, \tag{6.16}$$

and the sum is taken over all $r_1,\dots,r_{k-1}\geq0$ such that

$$r_1+\cdots+r_{k-1}\geq1\quad\text{and}\quad\sum_{s=1}^{k-1}\operatorname{ord}(\ell_s)r_s=\operatorname{ord}(\ell_k)-\sum_{s=1}^{j}\operatorname{ord}(\ell_{i_s})q_s. \tag{6.17}$$

In particular, if $\operatorname{ord}(\ell_k)-\sum_{s=1}^{j}\operatorname{ord}(\ell_{i_s})q_s\leq0$ then $P_k^{q_1\dots q_j i_1\dots i_j}(z)\equiv0$. The polynomials (6.16) can be explicitly found in the following way. Let us consider any element of the form

$$a=(\ell_{i_1}^{q_1}\cdots\ell_{i_j}^{q_j})(\ell_1^{r_1}\cdots\ell_{k-1}^{r_{k-1}})$$

such that (6.17) holds, and expand it with respect to the Poincaré–Birkhoff–Witt basis. Then $\langle d_k,a\rangle$ equals the coefficient of ℓ_k in this expansion.

Finally, notice that (6.15) holds for all $u^2\in B^{\theta^2}$, which gives the explicit representation of the formal power series $\mathcal{E}^z$,

$$\mathcal{E}_k^z=d_k+\sum_{\substack{j\geq1,\, i_1<\cdots<i_j\\ q_1,\dots,q_j\geq1}} P_k^{q_1\dots q_j i_1\dots i_j}(z)\, d_{i_1}^{\sqcup\!\sqcup q_1}\sqcup\!\sqcup\cdots\sqcup\!\sqcup d_{i_j}^{\sqcup\!\sqcup q_j},\qquad k=1,\dots,n, \tag{6.18}$$

where $P_k^{q_1\dots q_j i_1\dots i_j}(z)$ are defined by (6.16)–(6.17). ■

Below we describe the case when the right hand side of (6.18) includes only the elements $d_1,\dots,d_k$ for any $k=1,\dots,n$.

Lemma 6.12. *Let system* (2.1) *be homogeneous at the origin and $\mathcal{L}_{X_1,\dots,X_m}$ be a Lie ideal. Then the right hand side of* (6.18) *includes only shuffle polynomials of $d_1,\dots,d_k$* (*with coefficients depending on z*).

Proof. As before, without loss of generality assume $\mathcal{E}_k=d_k$, $k=1,\dots,n$. Due to Lemma 6.6, the ideal $\mathcal{J}_{X_1,\dots,X_m}$ is two-sided, hence

$$\langle d_k,a\ell_i b\rangle=0\quad\text{ for any } a,b\in\mathcal{F}^e\quad\text{ if } i\geq n+1.$$

In particular,

$$\big\langle d_k,(\ell_{i_1}^{q_1}\cdots\ell_{i_j}^{q_j})(\ell_1^{r_1}\cdots\ell_{k-1}^{r_{k-1}})\big\rangle=0\quad\text{ if } i_j\geq n+1.$$

Moreover, (6.11) implies

$$\big\langle d_k,(\ell_{i_1}^{q_1}\cdots\ell_{i_j}^{q_j})(\ell_1^{r_1}\cdots\ell_{k-1}^{r_{k-1}})\big\rangle=0\quad\text{ if } r_1+\cdots+r_{k-1}\geq1 \text{ and } k\leq i_j\leq n.$$

Hence,

$$P_k^{q_1\dots q_j i_1\dots i_j}(z)=0\qquad\text{if } i_j\geq k.$$

Taking into account (6.16) and (6.17), we rewrite (6.18) in the form

$$\mathcal{E}_k^z = d_k + \sum_{q_1+\cdots+q_{k-1}\geq 1} \widehat{P}_k^{q_1\cdots q_{k-1}}(z)\, d_1^{\sqcup\!\sqcup q_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d_{k-1}^{\sqcup\!\sqcup q_{k-1}}, \qquad k = 1,\ldots,n, \tag{6.19}$$

where

$$\widehat{P}_k^{q_1\cdots q_{k-1}}(z) = \sum \frac{\langle d_k, (\ell_1^{q_1}\cdots\ell_{k-1}^{q_{k-1}})(\ell_1^{r_1}\cdots\ell_{k-1}^{r_{k-1}})\rangle}{q_1!\cdots q_{k-1}!r_1!\cdots r_{k-1}!} \prod_{s=1}^{k-1} z_s^{r_s}, \tag{6.20}$$

and the sum is taken over all $r_1,\ldots,r_{k-1}\geq 0$ such that

$$r_1+\cdots+r_{k-1}\geq 1 \quad \text{and} \quad \sum_{s=1}^{k-1}\operatorname{ord}(\ell_s)r_s = \operatorname{ord}(\ell_k) - \sum_{s=1}^{k-1}\operatorname{ord}(\ell_s)q_s. \tag{6.21}$$

Hence, $\mathcal{E}_k^z$ equals a shuffle polynomial of $d_1,\ldots,d_k$. ■

The following result was suggested by Igor Zelenko.

Theorem 6.13. *Let system* (2.1) *be homogeneous at the origin. This system is regular if and only if $\mathcal{L}_{X_1,\ldots,X_m}$ is a Lie ideal. Moreover, in this case the core Lie subalgebra of the system is constant, that is, $\mathcal{L}^z_{X_1,\ldots,X_m} = \mathcal{L}_{X_1,\ldots,X_m}$ for any $z\in\mathbb{R}^n$* (*hence, the system has the same homogeneous approximation at any point*). *Moreover, for any $z\in\mathbb{R}^n$ there exists a polynomial change of variables* (*depending on z*) *that transforms the system to a homogeneous form at z.*

Proof. Due to Lemma 6.4, if a system is regular then its core Lie subalgebra is a Lie ideal. Let us prove the converse statement for a homogeneous system.

Consider a homogeneous system of the form (2.1) and suppose $\mathcal{L}_{X_1,\ldots,X_m}$ is a Lie ideal. Then, due to Lemma 6.12, we get the representation (6.19)–(6.21). Introduce the polynomial mapping $\Phi:\mathbb{R}^n\to\mathbb{R}^n$ (depending on the parameter z) of the form $\Phi = (\Phi_1,\ldots,\Phi_n)$, where

$$\Phi_k(x_1,\ldots,x_n) = x_k + \sum_{q_1+\cdots+q_{k-1}\geq 1} \widehat{P}_k^{q_1\cdots q_{k-1}}(z)\prod_{s=1}^{k-1} x_s^{q_s}.$$

Obviously, it is of triangular form, namely $\Phi_k = x_k + \widetilde{\Phi}_k(x_1,\ldots,x_{k-1})$. Therefore, Φ^{-1} is also a nontrivial polynomial mapping. Therefore, the change of variables $x=\Phi^{-1}(y)$ (depending on z) satisfies $(\Phi^{-1}(\mathcal{E}^z))_k = d_k$ for $k=1,\ldots,n$. This means that the system in the new variables is homogeneous at z and $\mathcal{L}^z_{X_1,\ldots,X_m} = \mathcal{L}_{X_1,\ldots,X_m}$, i.e., $c^z(\mathcal{L}^z_{X_1,\ldots,X_m}) = c^z(\mathcal{L}_{X_1,\ldots,X_m}) = 0$. ■

Remark 6.14. Regular homogeneous systems can be thought of as homogeneous approximations of regular systems. Notice that the representation (6.19) is, in essence, constructed in [6]. It is used there to obtain distance estimates in a neighborhood of a regular point for the original system and for a homogeneous approximation of the system [6, Section 7]. We emphasize, however, that algebraic methods allow us to obtain the precise formula (6.20) for the polynomial coefficients $\widehat{P}_k^{q_1\cdots q_{k-1}}(z)$.

Recall (see Subsection 2.5) that $L=\sum_{k=1}^\infty L^k$ denotes a (filtered) Lie algebra of vector fields generated by the set $X_1,\ldots,X_m$. As a consequence of Theorem 6.13, we get the well-known property of the Lie algebra L for the case of a regular and homogeneous system.

Corollary 6.15. *Let system* (2.1) *be regular and homogeneous at the origin. Then the Lie algebra of vector fields L generated by the set $X_1, \dots, X_m$ is n-dimensional.*

Proof. Suppose elements $\ell_1, \dots, \ell_n$ satisfy (6.10) and (6.11). Then, in particular,

$$c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^k) = \mathrm{Lin}\{c(\ell_1), \dots, c(\ell_{v_k})\}, \qquad k = 1, \dots, p,$$

where p is the degree of nonholonomy and $v = (v_1, \dots, v_p)$ is the growth vector; by the supposition, they are the same for all z. Introduce the vector fields $Y_i = \varphi(\ell_i)$, $i = 1, \dots, n$. Let us show that $Y_1, \dots, Y_n$ form a basis for L.

It is sufficient to prove that any vector field $Y = \varphi(\ell)$, where $\ell \in \mathcal{L}^k$, $k \geq 1$, equals a linear combination of $Y_1, \dots, Y_n$ with constant coefficients.

First, suppose $k \leq p$. Since $Y(0) = c(\ell) \in c(\mathcal{L}^k)$, we get

$$Y(0) = \sum_{i=1}^{v_k} \alpha_i Y_i(0),$$

where α_i are constants. Denote

$$\widehat{\ell} = \ell - \sum_{i=v_{k-1}+1}^{v_k} \alpha_i \ell_i \in \mathcal{L}^k \quad \text{and} \quad \widehat{Y} = \varphi(\widehat{\ell}) = Y - \sum_{i=v_{k-1}+1}^{v_k} \alpha_i Y_i.$$

Then

$$c(\widehat{\ell}) = \widehat{Y}(0) = \sum_{i=1}^{v_{k-1}} \alpha_i Y_i(0) = \sum_{i=1}^{v_{k-1}} \alpha_i c(\ell_i) \in c(\mathcal{L}^1 \oplus \cdots \oplus \mathcal{L}^{k-1}),$$

that is, $\widehat{\ell} \in \mathcal{P}^k \subset \mathcal{L}_{X_1,\dots,X_m}$. Since the system is regular and homogeneous at the origin, Theorem 6.13 implies $\mathcal{L}_{X_1,\dots,X_m} = \mathcal{L}^z_{X_1,\dots,X_m}$ and, moreover, $c^z(\mathcal{L}^z_{X_1,\dots,X_m}) = c^z(\mathcal{L}_{X_1,\dots,X_m}) = 0$, for any $z \in \mathbb{R}^n$. Hence, $c^z(\widehat{\ell}) = \widehat{Y}(z) = 0$ for any $z \in \mathbb{R}^n$, i.e.,

$$\widehat{Y}(z) = Y(z) - \sum_{i=v_{k-1}+1}^{v_k} \alpha_i Y_i(z) = 0.$$

This means that

$$Y(z) = \sum_{i=v_{k-1}+1}^{v_k} \alpha_i Y_i(z), \qquad z \in \mathbb{R}^n.$$

If $k \geq p+1$ then automatically $\ell \in \mathcal{L}^z_{X_1,\dots,X_m}$, and hence $c^z(\ell) = Y(z) = 0$, for any $z \in \mathbb{R}^n$.

Since $L = \varphi(\bigoplus_{k=1}^{\infty} \mathcal{L}^k)$, an arbitrary vector field $Y(z) \in L$ equals a linear combination of vector fields $Y_1(z), \dots Y_n(z)$ with constant coefficients. In other words, $Y_1, \dots, Y_n$ is a basis for the Lie algebra of vector fields L (over $\mathbb{R}$). Thus, L is n-dimensional. ∎

7. Time optimality

7.1. Time-optimal controls. In this section we return to general control-linear systems of the form (2.1), where the vector fields $X_1, \dots, X_m$ are real analytic in a neighborhood of the origin. Moreover, we assume they satisfy the Rashevsky–Chow condition (2.24). Then there exists a neighborhood $U(0)$ of the origin such that any point from this neighborhood can be reached from any other point from this neighborhood.

In this subsection we consider the time-optimal control problem for system (2.1) of the form

$$\dot{x} = \sum_{i=1}^{m} u_i(t)X_i(x), \qquad x(0) = s^1,\ x(\theta) = s^2,\ \sum_{i=1}^{m} u_i^2(t) \le 1 \text{ a.e., } t \in [0,\theta], \qquad \theta \to \min, \tag{7.1}$$

where we assume $s^1, s^2 \in U(0)$ and $s^1 \neq s^2$.

Our first observation concerns the character of the optimal control.

Theorem 7.1. *Suppose θ^* is the optimal time and $u^*(t) \in B^{\theta^*}$ is an optimal control in the problem* (7.1)*. Then*

$$\sum_{i=1}^{m} u_i^{*\,2}(t) = 1 \qquad a.e.,\ t \in [0,\theta^*]. \tag{7.2}$$

Proof. Notice that the existence of the time-optimal control follows from the Filippov theorem [15, 16]; however, it is not necessarily unique. Denote by $x^*(t)$ the optimal trajectory corresponding to the control $u^*(t)$.

For any $\varepsilon > 0$, consider a reparameterization of the curve $x^*(t)$ of the form

$$\tau = \psi(t) = \int_0^t \sqrt{\sum_{i=1}^{m} u_i^{*2}(\sigma)}\, d\sigma + \varepsilon t, \qquad t \in [0,\theta^*].$$

In other words, $\tau = \psi(t)$ is a change of time in (7.1); it is well defined since $\dot{\psi}(t) > 0$. With respect to this new time, the optimal trajectory $\widehat{x}(\tau) = x^*(\psi^{-1}(\tau))$ satisfies the differential equality

$$\frac{d\widehat{x}(\tau)}{d\tau} = \frac{dx^*(t)}{dt}\bigg|_{t=\psi^{-1}(\tau)} \cdot \frac{d\psi^{-1}(\tau)}{d\tau} = \sum_{i=1}^{m} \widehat{u}_i(\tau)X_i(\widehat{x}(\tau)), \qquad \tau \in [0,\psi(\theta^*)],$$

where

$$\widehat{u}_i(\tau) = \frac{u_i^*(t)}{\dot{\psi}(t)}\bigg|_{t=\psi^{-1}(\tau)} = \frac{u_i^*(t)}{\sqrt{\sum_{i=1}^{m} u_i^{*2}(t)} + \varepsilon}\bigg|_{t=\psi^{-1}(\tau)}, \qquad i = 1,\dots,m,$$

and the conditions

$$\widehat{x}(0) = x^*(0) = s^1, \qquad \widehat{x}(\psi(\theta^*)) = x^*(\theta^*) = s^2.$$

Moreover,

$$\sum_{i=1}^{m} \widehat{u}_i^2(\tau) = \frac{\sum_{i=1}^{m} u_i^{*2}(t)}{\left(\sqrt{\sum_{i=1}^{m} u_i^{*2}(t)} + \varepsilon\right)^2}\Bigg|_{t=\psi^{-1}(\tau)} \le 1, \qquad \tau \in [0,\psi(\theta^*)].$$

Thus, the control $\widehat{u}(\tau) \in B^{\psi(\theta^*)}$ steers the origin to the point s in time $\psi(\theta^*)$ via system (7.1). Hence, the time of movement $\psi(\theta^*)$ is greater than or equal to the optimal time θ^*, that is,

$$\psi(\theta^*) = \int_0^{\theta^*} \sqrt{\sum_{i=1}^m u_i^{*2}(t)}\, dt + \varepsilon\theta^* \geq \theta^*.$$

Since this inequality holds for any $\varepsilon > 0$, we get

$$\int_0^{\theta^*} \sqrt{\sum_{i=1}^m u_i^{*2}(t)}\, dt \geq \theta^*.$$

Taking into account the constraint $u^* \in B^{\theta^*}$, we obtain (7.2). ■

COROLLARY 7.2. *Suppose θ^* is the optimal time and $u^*(t) \in B^{\theta^*}$ is an optimal control in the problem* (7.1). *Denote $\widehat{u}(t) = \theta^* u^*(t\theta^*)$, $t \in [0,1]$. Then*

(i) *the control $\widehat{u}(t)$ minimizes the "length functional", i.e., solves the optimal control problem*

$$\dot{x} = \sum_{i=1}^m u_i X_i(x), \qquad x(0) = s^1,\ x(1) = s^2, \qquad \ell(u) = \int_0^1 \sqrt{\sum_{i=1}^m u_i^2(t)}\, dt \to \min, \tag{7.3}$$

and $\min \ell(u) = \ell(\widehat{u}) = \theta^*$;

(ii) *the control $\widehat{u}(t)$ minimizes the "energy functional", i.e., solves the optimal control problem*

$$\dot{x} = \sum_{i=1}^m u_i X_i(x), \qquad x(0) = s^1,\ x(1) = s^2, \qquad J(u) = \int_0^1 \sum_{i=1}^m u_i^2(t)\, dt \to \min, \tag{7.4}$$

and $\min J(u) = J(\widehat{u}) = \theta^{*2}$.

Proof. (i) Let us consider an arbitrary control $u(t)$, $t \in [0,1]$, steering s^1 to s^2, and use the arguments analogous to those applied in the proof of Theorem 7.1. Considering the reparameterization

$$\tau = \psi(t) = \int_0^t \sqrt{\sum_{i=1}^m u_i^2(\sigma)}\, d\sigma + \varepsilon t, \qquad t \in [0,1],$$

we see that the control

$$\widetilde{u}_i(\tau) = \frac{u_i(t)}{\sqrt{\sum_{i=1}^m u_i^2(t)} + \varepsilon}\Bigg|_{t=\psi^{-1}(\tau)}, \qquad i = 1, \ldots, m,$$

steers s^1 to s^2 in time $\widetilde{\theta} = \psi(1)$ via system (7.1) and satisfies the constraints. Hence,

$$\widetilde{\theta} = \psi(1) = \int_0^1 \sqrt{\sum_{i=1}^m u_i^2(t)}\, dt + \varepsilon = \ell(u) + \varepsilon \geq \theta^*.$$

Since $\varepsilon > 0$ is arbitrary, we have $\ell(u) \geq \theta^*$.

On the other hand, due to condition (7.2) we get

$$\sum_{i=1}^m \widehat{u}_i^2(t) = \theta^{*2} \sum_{i=1}^m u_i^{*2}(t\theta^*) \equiv \theta^{*2}, \tag{7.5}$$

and hence

$$\ell(\widehat{u}) = \int_0^1 \sqrt{\sum_{i=1}^m \widehat{u}_i^2(t)}\,dt = \theta^*.$$

This means that $\widehat{u}(t)$ minimizes the length functional and, moreover, $\min \ell(u) = \theta^*$.

(ii) The Cauchy–Bunyakovsky inequality gives $\ell(u) \le \sqrt{J(u)}$. Hence, taking into account (i), we see that if u steers s^1 to s^2 then $\theta^* = \ell(\widehat{u}) \le \ell(u) \le \sqrt{J(u)}$.

On the other hand, due to (7.5), we have $\sqrt{J(\widehat{u})} = \theta^*$. This means that $\widehat{u}(t)$ minimizes the energy functional and $\min J(u) = \theta^{*2}$. ■

Recall that the length functional is closely connected with a concept of sub-Riemannian metrics [6]. Namely, the sub-Riemannian metric is defined as

$$\rho(s^1, s^2) = \inf \ell(u), \quad \text{where} \quad \ell(u) = \int_0^1 \sqrt{\sum_{i=1}^m u_i^2(t)}\,dt,$$

and infimum is taken over all $u_i(t) \in L_2[0,1]$ satisfying

$$\dot{x} = \sum_{i=1}^m u_i(t) X_i(x), \qquad x(0) = s^1,\ x(1) = s^2.$$

Thus, the solution $\widehat{u}(t)$ of (7.3), which exists due to Corollary 7.2, gives $\rho(s^1, s^2) = \ell(\widehat{u})$.

For the sake of completeness, we prove the analogous property for the energy minimization problem.

Proposition 7.3. *Suppose a control $\widehat{u}(t)$ minimizes the energy functional, i.e., solves* (7.4). *Then*

$$\sum_{i=1}^m \widehat{u}_i^2(t) \equiv \text{const},$$

where the constant obviously coincides with $\min J(u) = J(\widehat{u})$. *As a consequence,*

(i) $\widehat{u}(t)$ *minimizes the length functional, i.e., solves* (7.3), *and* $\min \ell(u) = \ell(\widehat{u}) = \sqrt{J(\widehat{u})}$;
(ii) $\theta^* = \sqrt{J(\widehat{u})}$ *is the optimal time and* $u^*(t) = (1/\theta^*)\widehat{u}(t/\theta^*)$ *is an optimal control for the time-optimal control problem* (7.1).

Proof. Let $\widehat{x}(t)$ be the optimal trajectory corresponding to the control $\widehat{u}(t)$. Consider any invertible smooth reparameterization $\tau = \psi(t)$ such that $\psi(0) = 0$, $\psi(1) = 1$. Then, analogously to the proof of Theorem 7.1, the curve $\widetilde{x}(\tau) = \widehat{x}(\psi^{-1}(\tau))$ is a trajectory of the system from s^1 to s^2 corresponding to the control

$$\widetilde{u}_i(\tau) = \left.\frac{\widehat{u}_i(t)}{\dot{\psi}(t)}\right|_{t=\psi^{-1}(\tau)}, \qquad i = 1, \dots, m.$$

Then

$$J(\widetilde{u}) = \int_0^1 \sum_{i=1}^m \widetilde{u}_i^2(\tau)\,d\tau = \int_0^1 \sum_{i=1}^m \frac{\widehat{u}_i^2(t)}{\dot{\psi}(t)}\,dt.$$

By supposition, $\widehat{u}(t)$ minimizes the energy functional. Hence, $\psi(t)=t$ is a solution of the variational problem

$$F(\psi)=\int_0^1\sum_{i=1}^m\frac{\widehat{u}_i^2(t)}{\dot\psi(t)}\,dt\to\min,\qquad \psi(0)=0,\ \psi(1)=1.$$

Thus, $\psi(t)=t$ satisfies the Euler equation, i.e.,

$$\sum_{i=1}^m\frac{\widehat{u}_i^2(t)}{\dot\psi^2(t)}=\mathrm{const}.$$

Substituting $\psi(t)=t$, we get $\sum_{i=1}^m\widehat{u}_i^2(t)\equiv\mathrm{const}$. More specifically, we obviously get $\sum_{i=1}^m\widehat{u}_i^2(t)\equiv J(\widehat{u})$.

(i) Let us prove that $\widehat{u}$ minimizes the length functional. Assume the converse. Then there exists a control $\bar u(t)$ such that $\ell(\bar u)<\ell(\widehat{u})$.

Denote $\bar\ell=\ell(\bar u)>0$ and consider a reparameterization of the form

$$\tau=\psi(t)=\left(\int_0^t\sqrt{\sum_{i=1}^m\bar u_i^2(\sigma)}\,d\sigma+\varepsilon t\right)\frac{1}{\bar\ell+\varepsilon},$$

where $\varepsilon>0$. Then $\dot\psi(t)>0$, $\psi(0)=0$, and $\psi(1)=1$. Set

$$\widetilde{\bar u}_i(\tau)=\left.\frac{\bar u_i(t)}{\dot\psi(t)}\right|_{t=\psi^{-1}(\tau)},\qquad i=1,\ldots,m.$$

Then

$$J(\widetilde{\bar u})=\int_0^1\sum_{i=1}^m\widetilde{\bar u}_i^2(\tau)\,d\tau=\int_0^1\sum_{i=1}^m\frac{\bar u_i^2(t)}{\dot\psi(t)}\,dt=\int_0^1\sum_{i=1}^m\frac{\bar u_i^2(t)}{\sqrt{\sum_{i=1}^m\bar u_i^2(t)}+\varepsilon}(\bar\ell+\varepsilon)\,dt$$
$$\le\int_0^1\sqrt{\sum_{i=1}^m\bar u_i^2(t)}\,dt\,(\bar\ell+\varepsilon)=\bar\ell^2+\bar\ell\varepsilon.$$

By supposition, $\widehat{u}$ minimizes the functional J. Also recall that due to the Cauchy–Bunyakovsky inequality, $\ell(\widehat{u})\le\sqrt{J(\widehat{u})}$. Hence,

$$\bar\ell=\ell(\bar u)<\ell(\widehat{u})\le\sqrt{J(\widehat{u})}\le\sqrt{J(\widetilde{\bar u})}\le\sqrt{\bar\ell^2+\bar\ell\varepsilon}.$$

Passing to the limit as $\varepsilon\to0$, we obtain a contradiction. This proves that $\widehat{u}$ minimizes the length functional, and $\min\ell(u)=\sqrt{J(\widehat{u})}$.

(ii) Set

$$\theta^*=\ell(\widehat{u})\quad\text{and}\quad u^*(t)=\frac{1}{\theta^*}\widehat{u}\left(\frac{t}{\theta^*}\right).$$

Since $\sum_{i=1}^m\widehat{u}_i^2(t)\equiv\ell^2(\widehat{u})=\theta^{*2}$, we get

$$\sum_{i=1}^m u_i^{*2}(t)=\frac{1}{\theta^{*2}}\sum_{i=1}^m\widehat{u}_i^2\left(\frac{t}{\theta^*}\right)\equiv1,$$

i.e., $u^*\in B^{\theta^*}$. Moreover, u^* steers s^1 to s^2 in time θ^*. Denote by θ_0 the optimal time for (7.1). Then $\theta_0\le\theta^*$. However, as proved in Corollary 7.2, $\theta_0=\min\ell(u)$. Hence,

$\min \ell(u) = \theta_0 \le \theta^* = \ell(\widehat{u}) = \min \ell(u)$, which implies that θ^* is the optimal time for (7.1). Therefore, u^* is an optimal control. ■

Theorem 7.1 and Proposition 7.3 mean that the optimal control problems (7.1) and (7.4) are equivalent. Namely, θ^* is the optimal time and $u^*(t)$ is an optimal control for (7.1) iff $\theta^* u^*(t\theta^*)$ is an optimal control for (7.4). It is commonly accepted that the problem (7.3) is equivalent to both of them [41], but we could not find a complete and rigorous proof in the literature. We emphasize that Corollary 7.2 and Proposition 7.3 give only a one-way implication.

7.2. Weak continuity property of iterated integrals and weak convergence of optimal controls. Let $T_0 > 0$ be such that the series (2.3) converges absolutely for any $0 \le \theta \le T_0$ and any $u \in B^\theta$. Since the origin is an equilibrium of (2.1), we have $\mathcal{E}_{X_1,\dots,X_m}(\theta, B^\theta) \subset \mathcal{E}_{X_1,\dots,X_m}(T_0, B^{T_0})$ if $0 \le \theta \le T_0$. Notice that the accessibility set $\mathcal{E}_{X_1,\dots,X_m}(T_0, B^{T_0})$ is a neighborhood of the origin, due to (2.24).

From now on, we consider the time-optimal control problem for system (2.1) of the form

$$\dot{x} = \sum_{i=1}^m u_i(t) X_i(x), \qquad x(0) = 0,\ x(\theta) = s,\ \sum_{i=1}^m u_i^2(t) \le 1 \text{ a.e., } t \in [0,\theta], \quad \theta \to \min. \tag{7.6}$$

Definition 7.4. Let $s \in \mathcal{E}_{X_1,\dots,X_m}(T_0, B^{T_0})$. We say that a pair (θ_s^*, u_s^*) is a *solution* of (7.6) if θ_s^* is the optimal time and $u_s^*(t)$, $t \in [0, \theta_s^*]$, is an optimal control for problem (7.6). The set of all optimal controls is denoted by U_s^*.

Remark 7.5. It follows from [36] that θ_s^* is continuous with respect to s.

In this subsection we consider controls as elements of the Hilbert space $L_2([0,1], \mathbb{R}^m)$. Below we say that a sequence $u_{(q)}$ weakly converges to u (in $L_2([0,1], \mathbb{R}^m)$), written

$$u_{(q)} \xrightarrow{w} u \quad \text{ as } q \to \infty,$$

if for any $f \in L_2([0,1], \mathbb{R}^m)$,

$$\int_0^1 \sum_{i=1}^m f_i(t) u_{(q)i}(t)\, dt \to \int_0^1 \sum_{i=1}^m f_i(t) u_i(t)\, dt.$$

This is the same as saying that $u_{(q)i} \xrightarrow{w} u_i$ in $L_2[0,1]$ for any $i = 1, \dots, m$.

Also, we denote by $\|\cdot\|_{L_2}$ the norm in $L_2[0,1]$, i.e., $\|v\|_{L_2} = \sqrt{\int_0^1 v^2(t)\, dt}$.

Remark 7.6. Suppose $z(t) \in L_2[0,1]$ satisfies the condition $|z(t)| \le C$ a.e. Then, as is well known, $(A(v))(t) = \int_0^t z(\tau) v(\tau)\, d\tau : L_2[0,1] \to L_2[0,1]$ is a compact linear operator. This implies the following property: If a sequence $v_{(q)} \in L_2[0,1]$ is weakly convergent, $v_{(q)} \xrightarrow{w} v$, then the sequence $A(v_{(q)})$ strongly converges to $A(v)$ in $L_2[0,1]$, i.e.,

$$\|A(v_{(q)}) - A(v)\|_{L_2}^2 = \int_0^1 \left| \int_0^t z(\tau)\big(v_{(q)}(\tau) - v(\tau)\big)\, d\tau \right|^2 dt \to 0 \quad \text{ as } q \to \infty.$$

Lemma 7.7. *Let* $u_{(q)} \stackrel{w}{\to} u$. *Then* $\eta_{i_1\dots i_k}(\cdot,u_{(q)}) \to \eta_{i_1\dots i_k}(\cdot,u)$ *in* $L_2[0,1]$ *for all* $k\ge 1$ *and all* $1\le i_1,\dots,i_k\le m$, *i.e.*,

$$\|\eta_{i_1\dots i_k}(\cdot,u_{(q)})-\eta_{i_1\dots i_k}(\cdot,u)\|_{L_2}^2=\int_0^1\left|\eta_{i_1\dots i_k}(t,u_{(q)})-\eta_{i_1\dots i_k}(t,u)\right|^2dt\to 0\qquad \textit{as } q\to\infty.$$

Proof. We argue by induction on k. For $k=1$, the proof follows from Remark 7.6. Suppose $j\ge 1$ and the statement of the lemma holds for all $k\le j$. Fix any $1\le i_1,\dots,i_{j+1}\le m$ and denote

$$z_{(q)}(t)=\eta_{i_2\dots i_{j+1}}(t,u_{(q)}),\qquad z(t)=\eta_{i_2\dots i_{j+1}}(t,u).$$

Then the induction supposition implies that $z_{(q)}\to z$, i.e.,

$$\|z_{(q)}-z\|_{L_2}^2=\int_0^1|z_{(q)}(t)-z(t)|^2\,dt\to 0\qquad\text{as } q\to\infty.$$

Notice that $|z(t)|\le C$ a.e. In fact,

$$\begin{aligned}|z(t)|&\le\int_0^t\int_0^{\tau_2}\cdots\int_0^{\tau_j}|u_{i_2}(\tau_2)|\,|u_{i_3}(\tau_3)|\cdots|u_{i_{j+1}}(\tau_{j+1})|\,d\tau_{j+1}\cdots d\tau_3\,d\tau_2\\&\le\prod_{r=2}^{j+1}\Big(\int_0^1|u_{i_r}(\tau)|\,d\tau\Big)\le\prod_{r=2}^{j+1}\|u_{i_r}\|_{L_2}=C.\end{aligned}$$

Hence,

$$\begin{aligned}\eta_{i_1\dots i_{j+1}}(t,u_{(q)})-\eta_{i_1\dots i_{j+1}}(t,u)&=\int_0^t u_{(q)i_1}(\tau_1)z_{(q)}(\tau_1)\,d\tau_1-\int_0^t u_{i_1}(\tau_1)z(\tau_1)\,d\tau_1\\&=\int_0^t u_{(q)i_1}(\tau_1)\big(z_{(q)}(\tau_1)-z(\tau_1)\big)\,d\tau_1+\int_0^t\big(u_{(q)i_1}(\tau_1)-u_{i_1}(\tau_1)\big)z(\tau_1)\,d\tau_1.\end{aligned}\tag{7.7}$$

Then Remark 7.6 implies that the second term (strongly) converges to zero. Let us estimate the first term:

$$\begin{aligned}&\int_0^1\left|\int_0^t u_{(q)i_1}(\tau_1)\big(z_{(q)}(\tau_1)-z(\tau_1)\big)\,d\tau_1\right|^2dt\\&\qquad\le\int_0^1\int_0^1|u_{(q)i_1}(\tau_1)|^2\,d\tau_1\int_0^1|z_{(q)}(\tau_2)-z(\tau_2)|^2\,d\tau_2\,dt\\&\qquad=\|u_{(q)i_1}\|_{L_2}^2\|z_{(q)}-z\|_{L_2}^2\le C\|z_{(q)}-z\|_{L_2}^2\to 0,\end{aligned}$$

due to the induction supposition and the fact that the weakly convergent sequence $u_{(q)i_1}$ is bounded. Thus, $\eta_{i_1\dots i_{j+1}}(\cdot,u_{(q)})-\eta_{i_1\dots i_{j+1}}(\cdot,u)$ strongly converges to zero. ■

Corollary 7.8. *Any functional* $\eta_{i_1\dots i_k}(1,u):L_2([0,1],\mathbb{R}^m)\to\mathbb{R}^1$ *is weakly continuous, i.e., if* $u_{(q)}\stackrel{w}{\to}u$ *then* $\eta_{i_1\dots i_k}(1,u_{(q)})\to\eta_{i_1\dots i_k}(1,u)$ *as* $q\to\infty$.

Proof. For $k=1$ the statement is clear. Suppose $k\ge 2$. Analogously to (7.7), we get

$$\begin{aligned}\eta_{i_1\dots i_k}(1,u_{(q)})-\eta_{i_1\dots i_k}(1,u)&=\int_0^1 u_{(q)i_1}(\tau_1)z_{(q)}(\tau_1)\,d\tau_1-\int_0^1 u_{i_1}(\tau_1)z(\tau_1)\,d\tau_1\\&=\int_0^1 u_{(q)i_1}(\tau_1)\big(z_{(q)}(\tau_1)-z(\tau_1)\big)\,d\tau_1+\int_0^1\big(u_{(q)i_1}(\tau_1)-u_{i_1}(\tau_1)\big)z(\tau_1)\,d\tau_1,\end{aligned}$$

where $z_{(q)}(t) = \eta_{i_2\dots i_k}(t, u_{(q)})$ and $z(t) = \eta_{i_2\dots i_k}(t, u)$. The second term tends to zero since $u_{(q)i_1} \xrightarrow{w} u_{i_1}$, and the first term tends to zero since $u_{(q)i_1}$ is bounded and $z_{(q)} - z$ strongly converges to zero due to Lemma 7.7. ∎

Below we use Notation 2.7. In particular, for any $\theta > 0$ and any $u(t) \in B^\theta$ we denote $u^\theta(t) = u(t\theta) \in B^1$, as well as for any $\theta > 0$ and any $u(t) \in B^1$ we denote $u^{1/\theta}(t) = u(t/\theta) \in B^\theta$.

Corollary 7.9. *For system* (2.1), *set*

$$\mathcal{E}^k(\theta, u) = \sum_{1\le i_1,\dots,i_k\le m} c_{i_1\dots i_k}\eta_{i_1\dots i_k}(\theta, u), \qquad k \ge 1. \tag{7.8}$$

Suppose $\theta_q \to \theta_0$, where $\theta_0 \le T_0$, $\theta_q \le T_0$, and $u_{(q)} \in B^{\theta_q}$, $u_0 \in B^{\theta_0}$ are such that $u_{(q)}^{\theta_q}(t) \xrightarrow{w} u_0^{\theta_0}(t)$ as $q \to \infty$. Then for any $N \ge 0$,

$$\sum_{k=N+1}^{\infty} \mathcal{E}^k(\theta_q, u_{(q)}) \to \sum_{k=N+1}^{\infty} \mathcal{E}^k(\theta_0, u_0) \qquad \text{as } q \to \infty.$$

Proof. Recall that $\|c_{i_1\dots i_k}\| \le k!C_1C_2^k$ for some $C_1, C_2 > 0$ such that $mC_2T_0 < 1$ (see Remark 2.5). Hence, if $u \in B^\theta$ then $\|\mathcal{E}^k(\theta, u)\| \le C_1(mC_2\theta)^k \le C_1(mC_2T_0)^k$.

Now, for any $\varepsilon > 0$ let us find $r \ge N$ such that $\frac{C_1}{1-mC_2T_0}(mC_2T_0)^{r+1} < \frac{1}{4}\varepsilon$. Then

$$\sum_{k=r+1}^{\infty} \|\mathcal{E}^k(\theta, u)\| < \frac{1}{4}\varepsilon \qquad \text{for any } 0 \le \theta \le T_0,\ u \in B^\theta.$$

Using the supposition of this corollary and Corollary 7.8, for any $k = N+1, \dots, r$ we get

$$\begin{aligned}\mathcal{E}^k(\theta_q, u_{(q)}) - \mathcal{E}^k(\theta_0, u_0) &= \theta_q^k\mathcal{E}^k(1, u_{(q)}^{\theta_q}) - \theta_0^k\mathcal{E}^k(1, u_0^{\theta_0})\\ &= \theta_q^k\big(\mathcal{E}^k(1, u_{(q)}^{\theta_q}) - \mathcal{E}^k(1, u_0^{\theta_0})\big) + (\theta_q^k - \theta_0^k)\mathcal{E}^k(1, u_0^{\theta_0}) \to 0 \qquad \text{as } q \to \infty.\end{aligned}$$

Hence, there exists q_0 such that

$$\sum_{k=N+1}^{r} \|\mathcal{E}^k(\theta_q, u_{(q)}) - \mathcal{E}^k(\theta_0, u_0)\| < \frac{1}{2}\varepsilon \qquad \text{for all } q > q_0.$$

As a result,

$$\begin{aligned}&\Big\|\sum_{k=N+1}^{\infty} \mathcal{E}^k(\theta_q, u_{(q)}) - \sum_{k=N+1}^{\infty} \mathcal{E}^k(\theta_0, u_0)\Big\|\\ &\qquad \le \sum_{k=N+1}^{r} \|\mathcal{E}^k(\theta_q, u_{(q)}) - \mathcal{E}^k(\theta_0, u_0)\| + \sum_{k=r+1}^{\infty} \|\mathcal{E}^k(\theta_q, u_{(q)})\| + \sum_{k=r+1}^{\infty} \|\mathcal{E}^k(\theta_0, u_0)\| < \varepsilon\end{aligned}$$

for all $q > q_0$, which completes the proof. ∎

Lemma 7.10. *Suppose $u_{(q)} \xrightarrow{w} u$ as $q \to \infty$, and $u_{(q)} \in B^1$. Then $u \in B^1$.*

Proof. The lemma states that the unit ball of $L_\infty[0, 1]$ is a weakly closed subset of $L_2[0, 1]$. For the sake of completeness, we prove this fact.

Suppose this is not true. Then there exists $E \subset [0,1]$ such that $\mu(E) > 0$ and $\sum_{i=1}^m u_i^2(t) > 1$, $t \in E$. Set $v(t) = u(t)$ if $t \in E$ and $v(t) = 0$ otherwise. Then

$$\int_E \sum_{i=1}^m u_i(t)u_{(q)i}(t)\,dt = \int_0^1 \sum_{i=1}^m v_i(t)u_{(q)i}(t)\,dt \to \int_0^1 \sum_{i=1}^m v_i(t)u_i(t)\,dt = \int_E \sum_{i=1}^m u_i^2(t)\,dt.$$

On the other hand,

$$\Big|\sum_{i=1}^m u_i(t)u_{(q)i}(t)\Big| \le \sqrt{\sum_{i=1}^m u_i^2(t)}\sqrt{\sum_{i=1}^m u_{(q)i}^2(t)} \le \sqrt{\sum_{i=1}^m u_i^2(t)},$$

and hence

$$\int_E \sum_{i=1}^m u_i^2(t)\,dt \le \int_E \sqrt{\sum_{i=1}^m u_i^2(t)}\,dt.$$

However, by supposition, $\sum_{i=1}^m u_i^2(t) > 1$, hence $\sum_{i=1}^m u_i^2(t) > \sqrt{\sum_{i=1}^m u_i^2(t)}$, $t \in E$. Since $\mu(E) > 0$, we get

$$\int_E \sum_{i=1}^m u_i^2(t)\,dt > \int_E \sqrt{\sum_{i=1}^m u_i^2(t)}\,dt.$$

This contradiction proves the lemma. ∎

Lemma 7.11. *Let $s_{(q)} \in \mathcal{E}_{X_1,\dots,X_m}(\theta_{s_{(q)}}, B^{\theta_{s_{(q)}}})$ be such that $s_{(q)} \to s$ as $q \to \infty$, where $0 < \theta_{s_{(q)}} \le T_0$ and $\theta_{s_{(q)}} \to \theta_0$. Then $s \in \mathcal{E}_{X_1,\dots,X_m}(\theta_0, B^{\theta_0})$, i.e., s can be achieved from the origin in time θ_0 by a control from B^{θ_0}.*

Moreover, assume $s_{(q)} = \mathcal{E}_{X_1,\dots,X_m}(\theta_{s_{(q)}}, u_{s_{(q)}})$. Then $s = \mathcal{E}_{X_1,\dots,X_m}(\theta_0, u_0^{1/\theta_0})$, where $u_0(t)$ is an arbitrary weak partial limit of the sequence $u_{s_{(q)}}(t\theta_{s_{(q)}})$.

Proof. Denote $v_q(t) = u_{s_{(q)}}(t\theta_{s_{(q)}})$, $t \in [0,1]$. Then $v_q \in B^1$, and hence v_q are elements of the unit ball of the space $L_2([0,1];\mathbb{R}^m)$. Since the unit ball of $L_2([0,1];\mathbb{R}^m)$ is weakly compact, the set of partial weak limits of the sequence v_q is nonempty. Let v_0 be an arbitrary partial weak limit of v_q, i.e., $v_{q_r} \overset{w}{\to} v_0$ as $r \to \infty$. Due to Lemma 7.10, $v_0 \in B^1$.

By assumption, $s_{(q)} = \mathcal{E}_{X_1,\dots,X_m}(\theta_{s_{(q)}}, u_{s_{(q)}})$ and $\theta_{s_{(q)}} \to \theta_0$ as $q \to \infty$. Hence, due to Corollary 7.9,

$$\mathcal{E}_{X_1,\dots,X_m}(\theta_{s_{(q_r)}}, u_{s_{(q_r)}}) \to \mathcal{E}_{X_1,\dots,X_m}(\theta_0, v_0^{1/\theta_0}) \quad \text{as } r \to \infty. \tag{7.9}$$

By assumption, $s_{(q_r)} \to s$. Hence,

$$s_{(q_r)} = \mathcal{E}_{X_1,\dots,X_m}(\theta_{s_{(q_r)}}, u_{s_{(q_r)}}) \to \mathcal{E}_{X_1,\dots,X_m}(\theta_0, v_0^{1/\theta_0}) = s,$$

that is, the control $v_0^{1/\theta_0}(t) \in B^{\theta_0}$ steers the origin to s in time θ_0. ∎

Corollary 7.12. *The set $\mathcal{E}_{X_1,\dots,X_m}(T_0, B^{T_0})$ is closed.*

Proof. Apply Lemma 7.11 with $\theta_{s_{(q)}} = T_0$. ∎

Now, let us return to the time-optimal control problem (7.6). Recall that we denote by (θ_s^*, u_s^*) a solution of problem (7.6), i.e., θ_s^* is the optimal time and $u_s^* \in B^{\theta_s^*}$ is an optimal control steering the origin to s.

Corollary 7.13. *Suppose* $s_{(q)} = \mathcal{E}_{X_1,\dots,X_m}(\theta_{s_{(q)}}, u_{s_{(q)}})$, *where* $0 < \theta_{s_{(q)}} \le T_0$ *and* $u_{s_{(q)}}$ *is in* $B^{\theta_{s_{(q)}}}$. *Assume* $s_{(q)} \to s \neq 0$ *and* $\theta_{s_{(q)}} \to \theta_s^*$ *as* $q \to \infty$ (*where* θ_s^* *is the optimal time for* (7.6)). *Let* $v(t)$ *be a partial weak limit of the sequence* $u_{s_{(q)}}(t\theta_{s_{(q)}})$, $t \in [0,1]$. *Then* $v^{1/\theta_s^*} \in U_s^*$, *i.e.*, $v^{1/\theta_s^*}(t) = v(t/\theta_s^*)$, $t \in [0,\theta_s^*]$, *is an optimal control for* (7.6).

Proof. Applying Lemma 7.11 with $\theta_0 = \theta_s^*$, we get $s = \mathcal{E}_{X_1,\dots,X_m}(\theta_s^*, v_0^{1/\theta_s^*})$, i.e., the control $v_0^{1/\theta_s^*}(t) \in B^{\theta_s^*}$ steers the origin to s in time θ_s^*. Since θ_s^* is the optimal time, $v_0^{1/\theta_s^*}(t)$ is an optimal control. ■

Corollary 7.14. *Assume that, in addition to the suppositions of Corollary* 7.13, *problem* (7.6) *has a unique solution* (θ_s^*, u_s^*). *Then*

$$u_{s_{(q)}}(t\theta_{s_{(q)}}) \xrightarrow{w} u_s^*(t\theta_s^*). \tag{7.10}$$

Finally, we apply Theorem 7.1.

Corollary 7.15. *Assume that, in addition to the suppositions of Corollary* 7.13, *problem* (7.6) *has a unique solution* (θ_s^*, u_s^*). *Then componentwise*

$$\int_0^1 |u_{s_{(q)}i}(t\theta_{s_{(q)}}) - u^*_{s\,i}(t\theta_s^*)|\,dt \to 0, \qquad i = 1,\dots,m, \qquad \text{as } q \to \infty. \tag{7.11}$$

Proof. Since the optimal control u_s^* satisfies (7.2), $u_s^*(t\theta_s^*)$ belongs to the unit sphere of the Hilbert space $L_2([0,1];\mathbb{R}^m)$, while the sequence $u_{s_{(q)}}(t\theta_{s_{(q)}})$ belongs to the unit ball of $L_2([0,1];\mathbb{R}^m)$. Hence, the weak convergence of this sequence implies strong convergence. This means that

$$u_{s_{(q)}}(t\theta_{s_{(q)}}) \to u_s^*(t\theta_s^*) \quad \text{in } L_2([0,1];\mathbb{R}^m),$$

i.e.,

$$\int_0^1 \sum_{i=1}^m |u_{s_{(q)}i}(t\theta_{s_{(q)}}) - u^*_{s\,i}(t\theta_s^*)|^2\,dt \to 0 \quad \text{as } q \to \infty,$$

which implies (7.11). ■

7.3. Approximation in the sense of time optimality. In nonlinear approximation theory, different approximation concepts may be adopted. One possible approach leads to the homogeneous approximation discussed above, which is connected with the properties of the endpoint map $\mathcal{E}_{X_1,\dots,X_m}$ (see Definition 3.1). In this section we introduce the concept of approximation in the sense of time optimality, following the ideas of [49, 51].

Definition 7.16. Consider the time-optimal control problems:

$$\dot{x} = \sum_{i=1}^m u_i(t)\widehat{X}_i(x), \qquad x(0) = 0,\ x(\theta) = s,\ \sum_{i=1}^m u_i^2(t) \le 1, \qquad \theta \to \min, \tag{7.12}$$

$$\dot{x} = \sum_{i=1}^m u_i(t)X_i(x), \qquad x(0) = 0,\ x(\theta) = s,\ \sum_{i=1}^m u_i^2(t) \le 1, \qquad \theta \to \min, \tag{7.13}$$

where the vector fields $\widehat{X}_1(x),\dots,\widehat{X}_m(x)$ and $X_1(x),\dots,X_m(x)$ are real analytic in a neighborhood of the origin. Suppose there exists an open domain $\Omega \subset \mathbb{R}^n \setminus \{0\}$,

$0 \in \overline{\Omega}$, such that problem (7.12) has a unique solution $(\widehat{\theta}_s^*, \widehat{u}_s^*)$ for any $s \in \Omega$. Denote by $\{(\theta_s^*, u_s^*) : u_s^* \in U_s^*\}$ the set of solutions of (7.13).

We say that the time-optimal control problem (7.12) *approximates* the time-optimal control problem (7.13) (in the domain Ω) if there exists a nonsingular transformation Φ of a neighborhood of the origin of $\mathbb{R}^n$, $\Phi(0) = 0$, such that

$$\theta^*_{\Phi(s)}/\widehat{\theta}^*_s \to 1 \quad \text{as } s \to 0, \tag{7.14}$$

and

$$\frac{1}{\theta}\int_0^\theta |u^*_{\Phi(s)i}(t) - \widehat{u}^*_{si}(t)|\,dt \to 0, \quad i = 1, \ldots, m, \quad \text{as } s \to 0, \tag{7.15}$$

for $s \in \Omega$ and any $u^*_{\Phi(s)} \in U^*_{\Phi(s)}$, where $\theta = \min\{\widehat{\theta}^*_s, \theta^*_{\Phi(s)}\}$.

In other words, after a certain change of variables in system (7.13), the optimal times and optimal controls of problems (7.12) and (7.13) become asymptotically equivalent as functions of the end point.

Our nearest goal is to prove that if system (3.1) is a homogeneous approximation of (2.1) then the time-optimal control problem for (3.1) approximates the time-optimal control problem for (2.1).

The main result of this section is the following approximation theorem; its proof complements [38], [48], [49] and [51]. (The above-mentioned papers deal with the steering problem for affine control systems with one-dimensional control, so the results obtained there are slightly different. In particular, for affine systems the analogue of Theorem 7.1 does not hold.)

Theorem 7.17. *Let a system*

$$\dot z = \sum_{i=1}^m u_i Z_i(z), \quad z \in \mathbb{R}^n,\ u_1, \ldots, u_m \in \mathbb{R}, \tag{7.16}$$

be a homogeneous approximation for (2.1). *Suppose that there exists an open domain* $\Omega \subset \mathbb{R}^n \setminus \{0\}$ *such that* $\Omega \subset \mathcal{E}_{X_1,\ldots,X_m}(T_0, B^{T_0})$, $0 \in \overline{\Omega}$, *and for any* $s \in \Omega$ *the solution* $(\widehat{\theta}^*_s, \widehat{u}^*_s)$ *of the time-optimal control problem*

$$\dot z = \sum_{i=1}^m u_i(t) Z_i(z), \quad z(0) = 0,\ z(\theta) = s,\ \sum_{i=1}^m u_i^2(t) \le 1 \text{ a.e.}, \quad \theta \to \min, \tag{7.17}$$

is unique. Then there exists a set of embedded domains $\Omega(\delta)$, $\delta > 0$, *such that* $\Omega(\delta_1) \subset \Omega(\delta_2)$ *if* $\delta_1 > \delta_2 > 0$ *and* $\Omega = \bigcup_{\delta>0} \Omega(\delta)$, *in each of which the time-optimal control problem* (7.17) *approximates the time-optimal control problem* (7.6).

Proof. Denote by $(\widehat{\theta}^*_s, \widehat{u}^*_s)$ the solution of (7.17), and by $\{(\theta^*_s, u^*_s) : u^*_s \in U^*_s\}$ the set of solutions of (7.6).

Suppose system (7.16) is written in privileged coordinates and $w_1 \le \cdots \le w_n$ are weights of the coordinates. Let H_ε denote the dilation $H_\varepsilon(y) = (\varepsilon^{w_1} y_1, \ldots, \varepsilon^{w_n} y_n)$. Then, due to homogeneity,

$$\widehat{\theta}^*_{H_\varepsilon(y)} = \varepsilon \widehat{\theta}^*_y, \quad \widehat{u}^*_{H_\varepsilon(y)}(t\varepsilon) = \widehat{u}^*_y(t),\ t \in [0, \widehat{\theta}^*_y]. \tag{7.18}$$

Hence, if some properties concerning the optimal time and control for problem (7.17) (such as existence, uniqueness, etc.) are satisfied in some domain Ω, then they are also true in any domain $H_\varepsilon(\Omega)$, $\varepsilon > 0$. Thus, without loss of generality we assume that the domain Ω satisfies the condition

$$\text{if} \quad y \in \Omega \quad \text{then} \quad H_\varepsilon(y) \in \Omega \quad \text{for any } 0 < \varepsilon \leq 1.$$

Introduce the pseudonorm $|||y||| = \max_{1\leq j\leq n}\{|y_j|^{1/w_j}\}$ in $\mathbb{R}^n$ and denote

$$V^\alpha = \{y \in \mathbb{R}^n : |||y||| \leq \alpha\}, \qquad \alpha > 0.$$

Notice that

$$H_\varepsilon(V^\alpha) = V^{\varepsilon\alpha}, \qquad \varepsilon, \alpha > 0. \tag{7.19}$$

Set

$$\omega(\delta) = \{y \in \partial V^1 : y + V^\delta \subset \Omega\}, \quad \Omega(\delta) = \bigcup_{0<\varepsilon\leq 1} H_\varepsilon(\omega(\delta)), \qquad \text{for any } 0 < \delta \leq 1/2,$$

and set $\Omega(\delta) = \Omega(1/2)$ for $\delta > 1/2$. Then $\Omega(\delta_1) \subset \Omega(\delta_2)$ if $\delta_1 > \delta_2 > 0$ and $\Omega \cap V^1 = \bigcup_{\delta>0} \Omega(\delta)$.

Suppose system (2.1) is also written in privileged coordinates. Fix any $0 < \delta \leq 1/2$ and prove that (7.17) approximates (7.6) in $\Omega(\delta)$. Without loss of generality, we assume

$$(\mathcal{E}_{X_1,\ldots,X_m})_j = P_j + \rho_j, \quad (\mathcal{E}_{Z_1,\ldots,Z_m})_j = P_j, \qquad j = 1, \ldots, n,$$

where $P_j = P_j(\theta, u)$ contains terms of order w_j, and ρ_j contains terms of order greater than w_j. Moreover, for any $0 \leq \theta \leq T_0$ and any $u \in B^\theta$,

$$|\rho_j(\theta, u)| \leq C_1 C_2^{w_j+1} \theta^{w_j+1}, \qquad j = 1, \ldots, n, \tag{7.20}$$

for some $C_1, C_2 > 0$.

Set

$$C = 2\sup\{\widehat{\theta}_y^* : y \in V^1 \cap \Omega\} > 0.$$

Below, choose $0 < \varepsilon \leq \min\{1, T_0/C\}$. Then, due to (7.18)–(7.19), we have

$$\sup\{\widehat{\theta}_y^* : y \in V^\varepsilon \cap \Omega\} \leq C\varepsilon/2 \leq T_0. \tag{7.21}$$

Fix any $s \in \Omega(\delta) \cap \partial V^\varepsilon$. Hence,

$$H_\varepsilon^{-1}(s) \in \omega(\delta), \quad \text{i.e.,} \quad H_\varepsilon^{-1}(s) \in V^1 \text{ and } H_\varepsilon^{-1}(s) + V^\delta \subset \Omega. \tag{7.22}$$

Following [51], consider the operator $G_s(y) : \Omega(\delta) \to \mathbb{R}^n$ defined as

$$G_s(y) = s - \rho(\widehat{\theta}_y^*, \widehat{u}_y^*).$$

Let us prove that, for sufficiently small ε, this operator has a fixed point in the set

$$M = s + V^{\delta\varepsilon}.$$

First, we prove that $G_s(y)$ maps M to itself.

Choose any $y \in M$. Then $y = s + \widehat{y}$, where $\widehat{y} \in V^{\delta\varepsilon}$. Hence, $|\widehat{y}_j| \leq (\delta\varepsilon)^{w_j} \leq \varepsilon^{w_j}$, and therefore $|y_j| \leq |s_j| + \varepsilon^{w_j} \leq 2\varepsilon^{w_j} \leq (2\varepsilon)^{w_j}$, i.e., $y \in V^{2\varepsilon}$.

Since $H_\varepsilon^{-1}(\widehat{y}) \in V^\delta$, we have $H_\varepsilon^{-1}(y) = H_\varepsilon^{-1}(s) + H_\varepsilon^{-1}(\widehat{y}) \in H_\varepsilon^{-1}(s) + V^\delta$. Hence, (7.22) implies $H_\varepsilon^{-1}(y) \in \Omega$, and therefore $y \in \Omega$.

Thus,

$$M = s + V^{\delta\varepsilon} \subset V^{2\varepsilon} \cap \Omega.$$

Then there exists a unique solution $(\widehat{\theta}^*_y, \widehat{u}^*_y)$ of problem (7.17). Hence, the operator G_s is defined at any $y \in M$.

Analogously to (7.21), we have $\widehat{\theta}^*_y \le C\varepsilon \le T_0$. Hence, (7.20) implies

$$|||\rho(\widehat{\theta}^*_y, \widehat{u}^*_y)||| = \max_{1\le j\le n}\{|\rho_j(\widehat{\theta}^*_y, \widehat{u}^*_y)|^{1/w_j}\} \le C_2 C\varepsilon \max_{1\le j\le n}\{(C_1C_2C\varepsilon)^{1/w_j}\} \le \delta\varepsilon \qquad (7.23)$$

if ε is sufficiently small, namely if

$$0 < \varepsilon \le \frac{1}{C_1C_2C} \min_{1\le j\le n}\left\{\left(\frac{\delta}{C_2C}\right)^{w_j}\right\}.$$

In this case,

$$G_s(y) = s - \rho(\widehat{\theta}^*_y, \widehat{u}^*_y) \in s + V^{\delta\varepsilon} = M.$$

Thus, if

$$0 < \varepsilon \le \varepsilon_0 = \min\left\{1, \frac{T_0}{C}, \frac{1}{C_1C_2C} \min_{1\le j\le n}\left\{\left(\frac{\delta}{C_2C}\right)^{w_j}\right\}\right\}$$

then for any fixed point $s \in \Omega(\delta) \cap \partial V^\varepsilon$ the operator G_s maps the set M to itself. Notice that M is convex and closed (and $0 \notin M$).

Now, we prove that G_s is continuous in M. Suppose a sequence $\{y_{(q)}\}_{q=1}^\infty \subset M$ is convergent, $y_{(q)} \to y$ as $q \to \infty$ (then $y \in M$ and $y \ne 0$). Due to Remark 7.5, we have $\widehat{\theta}^*_{y_{(q)}} \to \widehat{\theta}^*_y$. Hence, Corollary 7.14 implies $\widehat{u}^*_{y_{(q)}} \overset{w}{\to} \widehat{u}^*_y$. Thus, Corollary 7.9 yields $\rho(\widehat{\theta}^*_{y_{(q)}}, \widehat{u}^*_{y_{(q)}}) \to \rho(\widehat{\theta}^*_y, \widehat{u}^*_y)$, which means that G_s is continuous.

As a result, the continuous operator G_s maps the convex and closed set $M \subset \mathbb{R}^n$ to itself. Hence, due to the Fixed Point Theorem, G_s has a fixed point in M. Let us denote it by s^1, i.e., $G_s(s^1) = s^1$. Since $s \in \partial V^\varepsilon$, we get $\varepsilon = |||s|||$ and $M \subset V^{2\varepsilon}$. Hence, if $s \to 0$ then $\varepsilon \to 0$, and therefore $s^1 \to 0$.

For the point s^1, we have $s^1 = G_s(s^1) = s - \rho(\widehat{\theta}^*_{s^1}, \widehat{u}^*_{s^1})$. Hence,

$$s = s^1 + \rho(\widehat{\theta}^*_{s^1}, \widehat{u}^*_{s^1}).$$

However, $s^1 = P(\widehat{\theta}^*_{s^1}, \widehat{u}^*_{s^1})$. Thus,

$$s = P(\widehat{\theta}^*_{s^1}, \widehat{u}^*_{s^1}) + \rho(\widehat{\theta}^*_{s^1}, \widehat{u}^*_{s^1}).$$

This means that the control $\widehat{u}^*_{s^1} \in B^{\widehat{\theta}^*_{s^1}}$ steers the origin to the point s in time $\widehat{\theta}^*_{s^1}$ with respect to system (2.1). Hence, $\widehat{\theta}^*_{s^1}$ is greater than or equal to the optimal time, i.e., $\theta^*_s \le \widehat{\theta}^*_{s^1}$. Since $s^1 \in M \subset V^{2\varepsilon}$, we get the estimate

$$\theta^*_s \le \widehat{\theta}^*_{s^1} \le C\varepsilon. \qquad (7.24)$$

Moreover, consider the point $s^0 = s - \rho(\theta^*_s, u^*_s)$. Then

$$s = s^0 + \rho(\theta^*_s, u^*_s) = P(\theta^*_s, u^*_s) + \rho(\theta^*_s, u^*_s),$$

what implies $s^0 = P(\theta^*_s, u^*_s)$. Hence, the control $u^*_s \in B^{\theta^*_s}$ steers the origin to the point s^0 in time θ^*_s with respect to system (7.16), which gives the estimate $\widehat{\theta}^*_{s^0} \le \theta^*_s$. In addition, using (7.20), (7.23), and (7.21), we get $|||\rho(\theta^*_s, u^*_s)||| \le \delta\varepsilon$. Hence, $s^0 \in s + V^{\delta\varepsilon} = M$. Therefore, if $s \to 0$ then $s^0 \to 0$.

Thus, for any sufficiently small $\varepsilon > 0$ and any $s \in \Omega(\delta) \cap \partial V^\varepsilon$ we get the inequalities

$$\widehat{\theta}^*_{s^0} \le \theta^*_s \le \widehat{\theta}^*_{s^1}, \tag{7.25}$$

where $s^0, s^1 \to 0$ as $s \to 0$.

Now consider any sequence $\{s_{(q)}\}_{q=1}^\infty \subset \Omega(\delta)$ such that $s_{(q)} \to 0$. Set $\varepsilon_q = |||s_{(q)}||| \to 0$. For each point $s_{(q)}$, find the points $s^0_{(q)}$ and $s^1_{(q)}$ as it is explained above. Then

$$s_{(q)} = s^1_{(q)} + \rho(\widehat{\theta}^*_{s^1_{(q)}}, \widehat{u}^*_{s^1_{(q)}}), \quad s_{(q)} = s^0_{(q)} + \rho(\theta^*_{s_{(q)}}, u^*_{s_{(q)}}).$$

Consider the sequences

$$\widetilde{s}_{(q)} = H^{-1}_{\varepsilon_q}(s_{(q)}) \in \partial V^1, \quad \widetilde{s}^1_{(q)} = H^{-1}_{\varepsilon_q}(s^1_{(q)}), \quad \widetilde{s}^0_{(q)} = H^{-1}_{\varepsilon_q}(s^0_{(q)}).$$

Due to (7.20), we get

$$|(s^1_{(q)} - s_{(q)})_j| \le C_1(C_2 C\varepsilon_q)^{w_j+1}, \quad |(s^0_{(q)} - s_{(q)})_j| \le C_1(C_2 C\varepsilon_q)^{w_j+1},$$

and therefore there exists $C' > 0$ such that

$$|(\widetilde{s}^1_{(q)} - \widetilde{s}_{(q)})_j| \le C'\varepsilon_q, \quad |(\widetilde{s}^0_{(q)} - \widetilde{s}_{(q)})_j| \le C'\varepsilon_q. \tag{7.26}$$

Since ∂V^1 is a compact set, there exists a subsequence $\widetilde{s}_{(q_r)}$ such that $\widetilde{s}_{(q_r)} \to \widetilde{s} \in \partial V^1$ as $r \to \infty$. Then (7.26) implies $\widetilde{s}^1_{(q_r)} \to \widetilde{s}$ and $\widetilde{s}^0_{(q_r)} \to \widetilde{s}$. Due to Remark 7.5, this yields

$$\widehat{\theta}^*_{\widetilde{s}_{(q_r)}} = \widehat{\theta}^*_{s_{(q_r)}}/\varepsilon_{q_r} \to \widehat{\theta}^*_{\widetilde{s}}, \quad \widehat{\theta}^*_{\widetilde{s}^1_{(q_r)}} = \widehat{\theta}^*_{s^1_{(q_r)}}/\varepsilon_{q_r} \to \widehat{\theta}^*_{\widetilde{s}}, \quad \widehat{\theta}^*_{\widetilde{s}^0_{(q_r)}} = \widehat{\theta}^*_{s^0_{(q_r)}}/\varepsilon_{q_r} \to \widehat{\theta}^*_{\widetilde{s}}.$$

Hence,

$$\widehat{\theta}^*_{s^1_{(q_r)}}/\widehat{\theta}^*_{s_{(q_r)}} \to 1, \quad \widehat{\theta}^*_{s^0_{(q_r)}}/\widehat{\theta}^*_{s_{(q_r)}} \to 1.$$

Then (7.25) yields

$$\theta^*_{s_{(q_r)}}/\widehat{\theta}^*_{s_{(q_r)}} \to 1.$$

Since any subsequence of $s_{(q)}$ has a subsequence satisfying this relation, we finally get

$$\theta^*_{s_{(q)}}/\widehat{\theta}^*_{s_{(q)}} \to 1 \quad \text{as } s_{(q)} \to 0, \quad s_{(q)} \in \Omega(\delta), \tag{7.27}$$

which coincides with (7.14) (for $\Phi(s) = s$).

Now let us prove (7.15). Recall that, due to homogeneity,

$$\widetilde{s}^0_{(q_r)} = P(\theta^*_{\widetilde{s}_{(q_r)}}, u^*_{\widetilde{s}_{(q_r)}}), \quad \theta^*_{\widetilde{s}_{(q_r)}} = \theta^*_{s_{(q_r)}}/\varepsilon_{q_r}, \quad u^*_{\widetilde{s}_{(q_r)}}(t) = u^*_{s_{(q_r)}}(t\varepsilon_{q_r}),$$
$$\widetilde{s}_{(q_r)} = P(\widehat{\theta}^*_{\widetilde{s}_{(q_r)}}, \widehat{u}^*_{\widetilde{s}_{(q_r)}}), \quad \widehat{\theta}^*_{\widetilde{s}_{(q_r)}} = \widehat{\theta}^*_{s_{(q_r)}}/\varepsilon_{q_r}, \quad \widehat{u}^*_{\widetilde{s}_{(q_r)}}(t) = \widehat{u}^*_{s_{(q_r)}}(t\varepsilon_{q_r}),$$

for $t \in [0, \widehat{\theta}^*_{\widetilde{s}_{(q_r)}}]$. Recall that $\widetilde{s}^0_{(q_r)} \to \widetilde{s}$, $\theta^*_{\widetilde{s}_{(q_r)}} \to \widehat{\theta}^*_{\widetilde{s}}$ and $\widetilde{s}_{(q_r)} \to \widetilde{s}$, $\widehat{\theta}^*_{\widetilde{s}_{(q_r)}} \to \widehat{\theta}^*_{\widetilde{s}}$. Hence, Corollary 7.15 implies that

$$\int_0^1 \left|u^*_{\widetilde{s}_{(q_r)}i}(t\theta^*_{\widetilde{s}_{(q_r)}}) - \widehat{u}^*_{\widetilde{s}i}(t\widehat{\theta}^*_{\widetilde{s}})\right| dt \to 0, \quad \int_0^1 \left|\widehat{u}^*_{\widetilde{s}_{(q_r)}i}(t\widehat{\theta}^*_{\widetilde{s}_{(q_r)}}) - \widehat{u}^*_{\widetilde{s}i}(t\widehat{\theta}^*_{\widetilde{s}})\right| dt \to 0. \tag{7.28}$$

Therefore,

$$\int_0^1 \left|u^*_{\widetilde{s}_{(q_r)}i}(t\theta^*_{\widetilde{s}_{(q_r)}}) - \widehat{u}^*_{\widetilde{s}_{(q_r)}i}(t\widehat{\theta}^*_{\widetilde{s}_{(q_r)}})\right| dt = \int_0^1 \left|u^*_{s_{(q_r)}i}(t\theta^*_{s_{(q_r)}}) - \widehat{u}^*_{s_{(q_r)}i}(t\widehat{\theta}^*_{s_{(q_r)}})\right| dt \to 0.$$

Since any subsequence of $s_{(q)}$ has a subsequence satisfying this relation, we get

$$\int_0^1 \left|u^*_{s_{(q)}i}(t\theta^*_{s_{(q)}}) - \widehat{u}^*_{s_{(q)}i}(t\widehat{\theta}^*_{s_{(q)}})\right| dt \to 0, \quad i = 1, \ldots, m, \quad \text{as } q \to \infty,$$

which can be rewritten as

$$\frac{1}{\theta^*_{s_{(q)}}}\int_0^{\theta^*_{s_{(q)}}}\big|u^*_{s_{(q)}i}(t)-\widehat{u}^*_{s_{(q)}i}(t\mu_q)\big|\,dt\to 0,\qquad i=1,\dots,m,\qquad \text{as } q\to\infty, \tag{7.29}$$

where $\mu_q=\widehat{\theta}^*_{s_{(q)}}/\theta^*_{s_{(q)}}\to 1$.

It remains to prove that

$$\frac{1}{\theta_q}\int_0^{\theta_q}\big|\widehat{u}^*_{s_{(q)}i}(t\mu_q)-\widehat{u}^*_{s_{(q)}i}(t)\big|\,dt\to 0,\quad \text{where}\quad \theta_q=\min\{\theta^*_{s_{(q)}},\widehat{\theta}^*_{s_{(q)}}\}.$$

Write $\widetilde{\theta}_q=\theta_q/\varepsilon_q$. Then $\mu_q\widetilde{\theta}_q\le\widehat{\theta}^*_{s_{(q)}}/\varepsilon_q=\widehat{\theta}^*_{\widetilde{s}_{(q)}}$. Introduce the sequences

$$\widetilde{s}'_{(q)}=P(\widetilde{\theta}_q,\widehat{u}^*_{\widetilde{s}_{(q)}}),\quad \widetilde{s}''_{(q)}=P(\mu_q\widetilde{\theta}_q,\widehat{u}^*_{\widetilde{s}_{(q)}}).$$

Then

$$|(\widetilde{s}'_{(q)}-\widetilde{s}_{(q)})_j|\le C'_1\big|(\widetilde{\theta}_q)^{w_j}-(\widehat{\theta}^*_{\widetilde{s}_{(q)}})^{w_j}\big|,\quad |(\widetilde{s}''_{(q)}-\widetilde{s}_{(q)})_j|\le C''_1\big|(\mu_q\widetilde{\theta}_q)^{w_j}-(\widehat{\theta}^*_{\widetilde{s}_{(q)}})^{w_j}\big|.$$

Due to (7.27), $\widetilde{\theta}_{q_r}\to\widehat{\theta}^*_{\widetilde{s}}$ and $\mu_{q_r}\widetilde{\theta}_{q_r}\to\widehat{\theta}^*_{\widetilde{s}}$. Since $\widetilde{s}_{(q_r)}\to\widetilde{s}$ and $\widehat{\theta}^*_{\widetilde{s}_{(q_r)}}\to\widehat{\theta}^*_{\widetilde{s}}$, we have

$$\widetilde{s}'_{(q_r)}\to\widetilde{s},\quad \widetilde{s}''_{(q_r)}\to\widetilde{s},\qquad \text{as } r\to\infty.$$

Hence, Corollary 7.15 implies

$$\int_0^1\big|\widehat{u}^*_{\widetilde{s}_{(q_r)}i}(t\widetilde{\theta}_{q_r})-\widehat{u}^*_{\widetilde{s}i}(t\widehat{\theta}^*_{\widetilde{s}})\big|\,dt\to 0,\quad \int_0^1\big|\widehat{u}^*_{\widetilde{s}_{(q_r)}i}(t\mu_{q_r}\widetilde{\theta}_{q_r})-\widehat{u}^*_{\widetilde{s}i}(t\widehat{\theta}^*_{\widetilde{s}})\big|\,dt\to 0,$$

which gives

$$\int_0^1\big|\widehat{u}^*_{\widetilde{s}_{(q_r)}i}(t\widetilde{\theta}_{q_r})-\widehat{u}^*_{\widetilde{s}_{(q_r)}i}(t\mu_{q_r}\widetilde{\theta}_{q_r})\big|\,dt=\int_0^1\big|\widehat{u}^*_{s_{(q_r)}i}(t\theta_{q_r})-\widehat{u}^*_{s_{(q_r)}i}(t\mu_{q_r}\theta_{q_r})\big|\,dt\to 0.$$

Since any subsequence of $s_{(q)}$ has a subsequence satisfying this relation, we get

$$\int_0^1\big|\widehat{u}^*_{s_{(q)}i}(t\mu_q\theta_q)-\widehat{u}^*_{s_{(q)}i}(t\theta_q)\big|\,dt\to 0,\qquad i=1,\dots,m,\qquad \text{as } q\to\infty.$$

Rewriting, we obtain

$$\frac{1}{\theta_q}\int_0^{\theta_q}\big|\widehat{u}^*_{s_{(q)}i}(t\mu_q)-\widehat{u}^*_{s_{(q)}i}(t)\big|\,dt\to 0,\qquad i=1,\dots,m,\qquad \text{as } q\to\infty. \tag{7.30}$$

Combining (7.29) and (7.30), we finally get

$$\frac{1}{\theta_q}\int_0^{\theta_q}\big|u^*_{s_{(q)}i}(t)-\widehat{u}^*_{s_{(q)}i}(t)\big|\,dt\to 0,\qquad i=1,\dots,m,\qquad \text{as } s_{(q)}\to 0,\qquad s_{(q)}\in\Omega(\delta),$$

which coincides with (7.15) (for $\Phi(s)=s$). ∎

Remark 7.18. The asymptotic relation (7.14) for a system and its homogeneous approximation was obtained in [6]; this relation means that the sub-Riemannian distances to the origin defined by a system and by its homogeneous approximation are asymptotically equivalent. However, our definition of the approximation in the sense of time optimality also includes the asymptotic relation concerning optimal controls (7.15), which was not considered in [6].

Remark 7.19. Also notice that Theorem 7.1 allows us to give a partial answer to the question analogous to the open problem of [52]. Namely, in the case of the time-optimal control problem for a control-affine system of the form (1.2), the following condition is important for the approximation theorem analogous to Theorem 7.17: *For the set* $K = \{\widehat{u}_s^*(t\widehat{\theta}_s^*) : s \in \Omega,\ t \in [0,1]\}$ *considered as a set in* $L_2[0,1]$, *the weak convergence of a sequence of elements from* K *implies the strong convergence.* The open question is whether this condition follows from the other conditions of the theorem [52]. For the case of control-linear systems, Theorem 7.1 implies that the set K is contained in the unit sphere of the Hilbert space $L_2([0,1];\mathbb{R}^m)$, therefore it satisfies the above-mentioned condition.

8. Conclusion

In this paper we give a self-contained survey of the main ideas and techniques of the approach that is based on applying free algebras to the study of nonlinear control systems. Namely, a class of control-linear systems with m controls satisfying the Rashevsky–Chow condition is considered. The appropriate algebraic object is an m-generated free Lie algebra $\mathcal{L}$ with a natural grading and the corresponding m-generated free associative algebra $\mathcal{F}$. A control system can be replaced by a formal power series of elements of $\mathcal{F}$ with constant coefficients from $\mathbb{R}^n$, which corresponds to a series representation of the endpoint map $\mathcal{E}_{X_1,\dots,X_m}$ for the initial system.

In this way the analysis of properties of a control system is reduced to the study of corresponding properties of some structures in the free algebra. More specifically, the coefficients of the series define the so-called core Lie subalgebra $\mathcal{L}_{X_1,\dots,X_m}$, which is responsible for the homogeneous approximation of the system; an equivalent role is played by the left ideal $\mathcal{J}_{X_1,\dots,X_m}$.

This leads to an algebraic definition of a homogeneous approximation; in particular, this shows that the homogeneous approximation is unique (up to a polynomial change of coordinates). Moreover, any graded Lie subalgebra of $\mathcal{L}$ of codimension n is a core Lie subalgebra for some system, which gives a complete classification of all possible homogeneous approximations.

The algebraic technique allows us to find the homogeneous approximation and an approximating system explicitly, by use of an "elementary" operation of orthogonal projection in $\mathcal{F}$, without finding any special (privileged) coordinates; on the other hand, all privileged coordinates are effectively described.

We also give an algebraic characteristic of systems that are regular and homogeneous at the origin. For such systems, we give an explicit formula that expresses a series representation of the endpoint map $\mathcal{E}^z_{X_1,\dots,X_m}$ from an arbitrary point z via a series representation of the endpoint map $\mathcal{E}_{X_1,\dots,X_m}$ from the origin.

Finally, we show that the homogeneous approximation of a system of a given class is closely related to the approximation in the sense of time optimality.

References

[1] A. A. Agrachev and R. V. Gamkrelidze, *Exponential representation of flows and a chronological enumeration*, Mat. Sb. (N.S.) 107 (1978), no. 4, 467–532 (in Russian); English transl.: Math. USSR-Sb. 35 (1979), 727–785.

[2] A. A. Agrachev and R. V. Gamkrelidze, *The shuffle product and symmetric groups*, in: Differential Equations, Dynamical Systems, and Control Science, Lecture Notes Pure Appl. Math. 152, Dekker, New York, 1998, 365–382.

[3] A. A. Agrachev, R. V. Gamkrelidze, and A. V. Sarychev, *Local invariants of smooth control systems*, Acta Appl. Math. 14 (1989), 191–237.

[4] A. Agrachev and A. Marigo, *Nonholonomic tangent spaces: intrinsic construction and rigid dimensions*, Electron. Res. Announc. Amer. Math. Soc. 9 (2003), 111–120.

[5] N. I. Akhiezer and M. G. Krein, *Some Questions in the Theory of Moments*, GONTI, Kharkiv, 1938 (in Russian); English transl.: Transl. Math. Monogr. 2, Amer. Math. Soc., Providence, RI, 1962.

[6] A. Bellaïche, *The tangent space in sub-Riemannian geometry*, in: Sub-Riemannian Geometry, Progr. Math. 144, Birkhäuser, Basel, 1996, 4–78.

[7] A. Bellaïche, F. Jean, and J.-J. Risler, *Geometry of nonholonomic systems*, in: J.-P. Laumond (ed.), Robot Motion Planning and Control, Lecture Notes in Control and Inform. Sci. 229, Springer, 1998, 55–92.

[8] R. M. Bianchini and G. Stefani, *Graded approximation and controllability along a trajectory*, SIAM J. Control Optim. 28 (1990), 903–924.

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

[10] K. T. Chen, *Integration of paths—a faithful representation of parths by noncommutative formal power series*, Trans. Amer. Math. Soc. 89 (1958), 395–407.

[11] W. L. Chow, *Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung*, Math. Ann. 117 (1939), 98–105.

[12] P. E. Crouch, *Solvable approximations to control systems*, SIAM J. Control Optim. 22 (1984), 40–54.

[13] P. E. Crouch and F. Lamnabhi-Lagarrigue, *Algebraic and multiple integral identities*, Acta Appl. Math. 15 (1989), 235–274.

[14] S. Eilenberg and S. Mac Lane, *On the groups* $H(\Pi, n)$*, I*, Ann. of Math. 58 (1953), 55–106.

[15] A. F. Filippov, *On some questions in the theory of optimal regulation: existence of a solution of the problem of optimal regulation in the class of bounded measurable functions*, Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him. 1959, no. 2, 25–32 (in Russian).

[16] A. F. Filippov, *On certain questions in the theory of optimal control*, J. SIAM Control Ser. A 1 (1962), 76–84.

[17] M. Fliess, *Développements fonctionnels en indéterminées non commutatives des solutions d'équations différentielles non linéaires forcées*, C. R. Acad. Sci. Paris Sér. A–B 287 (1978), 1133–1135.

[18] M. Fliess, *Realizations of nonlinear systems and abstract transitive Lie algebras*, Bull. Amer. Math. Soc. 2 (1980), 444–446.

[19] M. Fliess, *Fonctionnelles causales non linéaires et indéterminées non commutatives*, Bull. Soc. Math. France 109 (1981), 3–40.

[20] M. Fliess, M. Lamnabhi, and F. Lamnabhi-Lagarrigue, *An algebraic approach to nonlinear functional expansions*, IEEE Trans. Circuits and Systems 30 (1983), 554–570.

[21] E. Gehrig and M. Kawski, *A Hopf-algebraic formula for compositions of noncommuting flows*, in: Proc. 47th IEEE Conference on Decision and Control, 2008, 1569–1574.

[22] E. G. Gilbert, *Functional expansions for the response of nonlinear differential systems*, IEEE Trans. Automatic Control AC-22 (1977), 909–921.

[23] H. Hermes, *Nilpotent approximations of control systems and distributions*, SIAM J. Control Optim. 24 (1986), 731–736.

[24] H. Hermes, *Nilpotent and high-order approximations of vector field systems*, SIAM Rev. 33 (1991), 238–264.

[25] S. Yu. Ignatovich, *Realizable growth vectors of affine control systems*, J. Dynam. Control Systems 15 (2009), 557–585.

[26] S. Yu. Ignatovich, *Normalization of homogeneous approximations of symmetric affine control systems with two controls*, J. Dynam. Control Systems 17 (2011), 1–48.

[27] A. Isidori, *Nonlinear control systems: an introduction*, in: Lecture Notes in Control and Inform. Sci. 72, Springer, Berlin, 1985.

[28] B. Jakubczyk, *Local realizations of nonlinear causal operators*, SIAM J. Control Optim. 24 (1986), 230–242.

[29] B. Jakubczyk, *Realization theory for nonlinear systems; three approaches*, in: M. Fliess and M. Hazewinkel (eds.), Algebraic and Geometric Methods in Nonlinear Control Theory, Reidel, Dordrecht, 1986, 3–31.

[30] B. Jakubczyk, *Convergence of power series along vector fields and their commutators; a Cartan–Kähler type theorem*, Ann. Polon. Math. 74 (2000), 117–132.

[31] F. Jean, *Uniform estimation of sub-Riemannian balls*, J. Dynam. Control Systems 7 (2001), 473–500.

[32] M. Kawski, *Nonlinear control and combinatorics of words*, in: Geometry of Feedback and Optimal Control, Monogr. Textbooks Pure Appl. Math. 207, Dekker, New York, 1998, 305–346.

[33] M. Kawski, *The combinatorics of nonlinear controllability and noncommuting flows*, in: ICTP Lect. Notes 8, Trieste, 2002, 223–312.

[34] M. Kawski, *Control interpretations of products in the Hopf algebra*, in: Proc. 48th IEEE Conference on Decision and Control, 2009, 7503–7508.

[35] M. Kawski and H. Sussmann, *Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory*, in: Operators, Systems, and Linear Algebra, Teubner, 1997, 111–128.

[36] V. I. Korobov, *The continuous dependence of a solution of an optimal-control problem with a free time for initial data*, Differentsial'nye Uravneniya 7 (1971), 1120–1123 (in Russian); English transl.: Differ. Equations 7 (1971), 850–852 (1973).

[37] V. I. Korobov and G. M. Sklyar, *Time-optimality and the power moment problem*, Mat. Sb. (N.S.) 134 (176) (1987), no. 2, 186–206, 287 (in Russian); English transl.: Math. USSR-Sb. 62 (1989), no. 1, 185–206.

[38] V. I. Korobov and G. M. Sklyar, *The Markov moment min-problem and time optimality*, Sibirsk. Mat. Zh. 32 (1991), no. 1, 60–71, 220 (in Russian); English transl.: Siberian Math. J. 32 (1991), no. 1, 46–55.

[39] M. G. Krein and A. A. Nudel'man, *The Markov moment problem and extremal problems. Ideas and problems of P. L. Chebyshev and A. A. Markov and their further development*, Nauka, Moscow, 1973 (in Russian); English transl.: Transl. Math. Monogr. 50, Amer. Math. Soc., Providence, RI, 1977.

[40] C. Lesiak and A. Krener, *The existence and uniqueness of Volterra series for nonlinear systems*, IEEE Trans. Automat. Control 23 (1978), 1090–1095.

[41] W. S. Liu and H. J. Sussmann, *Shortest paths for sub-Riemannian metrics of rank two distributions*, Mem. Amer. Math. Soc. 118 (1995), no. 564.

[42] M. Lothaire, *Combinatorics on Words*, Addison-Wesley, Reading, MA, 1983.

[43] A. A. Markov, *Nouvelles applications des fractions continues*, Math. Ann. 47 (1896), 579–597.

[44] G. Melançon and C. Reutenauer, *Lyndon words, free algebras and shuffles*, Canad. J. Math. 41 (1989), 577–591.

[45] P. K. Rashevski, *About connecting two points of a completely nonholonomic space by an admissible curve*, Uch. Zapiski Ped. Inst. Liebknechta 2 (1938), 83–94 (in Russian).

[46] R. Ree, *Lie elements and an algebra associated with shuffles*, Ann. of Math. 68 (1958), 210–220.

[47] C. Reutenauer, *Free Lie Algebras*, Clarendon Press, Oxford, 1993.

[48] G. M. Sklyar and S. Yu. Ignatovich, *A classification of linear time-optimal control problems in a neighborhood of the origin*, J. Math. Anal. Appl. 203 (1996), 791–811.

[49] G. M. Sklyar and S. Yu. Ignatovich, *Moment approach to nonlinear time optimality*, SIAM J. Control Optim. 38 (2000), 1707–1728.

[50] G. M. Sklyar and S. Yu. Ignatovich, *Representations of control systems in the Fliess algebra and in the algebra of nonlinear power moments*, Systems Control Lett. 47 (2002), 227–235.

[51] G. M. Sklyar and S. Yu. Ignatovich, *Approximation of time-optimal control problems via nonlinear power moment min-problems*, SIAM J. Control Optim. 42 (2003), 1325–1346.

[52] G. M. Sklyar and S. Yu. Ignatovich, *Determining of various asymptotics of solutions of nonlinear time optimal problems via right ideals in the moment algebra* (*Problem 3.8*), in: V. D. Blondel and A. Megretski (eds.), Unsolved Problems in Mathematical Systems and Control Theory, Princeton Univ. Press, Princeton, 2004, 117–121.

[53] G. M. Sklyar and S. Yu. Ignatovich, *Description of all privileged coordinates in the homogeneous approximation problem for nonlinear control systems*, C. R. Math. Acad. Sci. Paris 344 (2007), 109–114.

[54] G. M. Sklyar and S. Yu. Ignatovich, *Fliess series, a generalization of the Ree's theorem, and an algebraic approach to a homogeneous approximation problem*, Int. J. Control 81 (2008), 369–378.

[55] G. M. Sklyar, S. Yu. Ignatovich, and P. Yu. Barkhaev, *Algebraic classification of nonlinear steering problems with constraints on control*, in: G. Oyibo (ed.), Adv. Math. Res. 6, Nova Science, 2005, 37–96.

[56] H. J. Sussmann, *A product expansion for the Chen series*, in: Theory and Applications of Nonlinear Control Systems (Stockholm, 1985), North-Holland, Amsterdam, 1986, 323–335.

[57] Y. Wang and E. D. Sontag, *Generating series and nonlinear systems: analytic aspects, local realizability, and i/o representations*, Forum Math. 4 (1992), 299–322.