Available online at www.sciencedirect.com
ScienceDirect
J. Differential Equations 232 (2007) 134–175
Journal of Differential Equations
www.elsevier.com/locate/jde

Commutators of flow maps of nonsmooth vector fields

Franco Rampazzo [a,*,1], Héctor J. Sussmann [b,2]

[a] *Dipartimento di Matematica Pura ed Applicata, Università di Padova, Via Belzoni 7, 35131 Padova, Italy*
[b] *Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA*

Received 12 April 2006

Available online 26 September 2006

Abstract

Relying on the notion of set-valued Lie bracket introduced in an earlier paper, we extend some classical results valid for smooth vector fields to the case when the vector fields are just Lipschitz. In particular, we prove that the flows of two Lipschitz vector fields commute for small times if and only if their Lie bracket vanishes everywhere (i.e., equivalently, if their classical Lie bracket vanishes almost everywhere). We also extend the asymptotic formula that gives an estimate of the lack of commutativity of two vector fields in terms of their Lie bracket, and prove a simultaneous flow box theorem for commuting families of Lipschitz vector fields.

Keywords: Lie bracket; Lipschitz vector field; Commutativity; Asymptotic formula; Simultaneous flow-box; Higher order bracket

1. Introduction

The purpose of this paper is to extend to nonsmooth vector fields the following three facts, known to be true if $f_1, \ldots, f_d$ are vector fields of class C^1 on a manifold M of class C^2:

* Corresponding author.
E-mail addresses: rampazzo@math.unipd.it (F. Rampazzo), sussmann@math.rutgers.edu (H.J. Sussmann).

1 Research supported in part by the Research Project *Alcuni aspetti analitici e geometrico-differenziali nei problemi di minimo a crescita lenta*, University of Padova, 2003.

2 Research supported in part by NSF Grants DMS01-03901 and DMS-05-09930.

0022-0396/$ – see front matter
doi:10.1016/j.jde.2006.04.016

(I) (*Asymptotics*) If $d = 2$, $f = f_1$, and $g = f_2$, then the asymptotic formula

$$\lim_{(t,s)\to(0,0),\, t\neq 0,\, s\neq 0} \frac{1}{st}\big(\big(\Phi^g_{-s}\Phi^f_{-t}\Phi^g_s\Phi^f_t\big)(q) - q\big) = [f,g](q) \tag{1}$$

holds for every $q \in M$.

(II) (*Commutativity*) If $d = 2$, $f = f_1$, and $g = f_2$, then the flow maps of f and g commute for small times if and only if the Lie bracket $[f,g]$ vanishes identically. Precisely,

$$\begin{aligned} &\big((\forall q \in M)(\exists \varepsilon > 0)\big(\forall t, s \in [-\varepsilon, \varepsilon]\big)\big(\Phi^g_{-s}\Phi^f_{-t}\Phi^g_s\Phi^f_t\big)(q) = q\big) \\ &\quad\Longleftrightarrow\quad \big((\forall q \in M)[f,g](q) = 0\big). \end{aligned} \tag{2}$$

(III) (*Simultaneous flow-box*) If $[f_i, f_j](q) = 0$ for all $q \in M$ and all $i, j = 1, \ldots, d$, and $\bar{q} \in M$ is such that the vectors $f_1(\bar{q}), \ldots, f_d(\bar{q})$ are linearly independent, then there exists a coordinate chart of class C^1 near $\bar{q}$ with respect to which all the f_i are represented by constant vectors.

Here, (i) if X is a vector field on M that has uniqueness of trajectories, and $r \in \mathbb{R}$, then Φ^X_r is the *time r flow map* corresponding to X; therefore, if $q \in M$, then $\mathbb{R} \ni r \mapsto \Phi^X_r(q)$ is the integral curve of X that goes through q at time $r = 0$, (ii) if X and Y are vector fields of class C^1 on M, then $[X, Y]$ is the Lie bracket of X and Y.

In view of these facts, it is natural to ask whether the asymptotic formula (1), the characterization of commutativity given by (2), and the "simultaneous flow-box" theorem (III), are valid for flows of locally Lipschitz vector fields, rather than for vector fields of class C^1. All three results involve Lie brackets, whose meaning for locally Lipschitz vector fields is not immediately clear, so the desired extension of (1), (2), and (III) would require that we first propose an adequate generalized notion of Lie bracket.

We will offer affirmative answers to these questions, using the notion of set-valued Lie bracket of locally Lipschitz vector fields introduced in [7]. If we write $[f,g]_{\mathrm{set}}(q)$, for each point q, to denote the value of this bracket at q, then $[f,g]_{\mathrm{set}}(q)$ is a nonempty compact convex subset of the tangent space T_qM, and the map $M \ni q \mapsto [f,g]_{\mathrm{set}}(q) \subseteq T_qM$ is upper semicontinuous. Furthermore, the set $[f,g]_{\mathrm{set}}(q)$ coincides with the singleton $\{[f,g](q)\}$ when f and g are of class C^1. (The precise definition is given in Definition 3.1 below.)

Using the set-valued bracket, our generalization of (I) will consist of the formula

$$\lim_{(t,s)\to(0,0),\, t\neq 0,\, s\neq 0} \frac{1}{st}\operatorname{dist}\big(\big(\Phi^g_{-s}\Phi^f_{-t}\Phi^g_s\Phi^f_t\big)(q) - q, [f,g]_{\mathrm{set}}(q)\big) = 0, \tag{3}$$

valid for locally Lipschitz vector fields f, g, as well as the formula

$$\lim_{t\to 0,\, t\neq 0} \frac{1}{t^2}\Delta\big(\big(\Phi^g_{-t}\Phi^f_{-t}\Phi^g_t\Phi^f_t\big)(q) - q, [F,G]_{\mathrm{set}}(q)\big) = 0, \tag{4}$$

valid for a pair of vector fields f, g that are semidifferentiable at a point q. (The "quasidistance" Δ is defined in (24) below. A vector field is *semidifferentiable* at a point q if it is continuous near q and can be approximated near q to first order by a Lipschitz vector field, cf. Section 4.5. In (4), F and G are Lipschitz vector fields that approximate f and g near q

to first order. Furthermore, the map $\Phi^g_{-t}\Phi^f_{-t}\Phi^g_t\Phi^f_t$ is possibly set-valued, since f and g need not have unique trajectories.) We will also show, by giving a counterexample, that (4) cannot be extended to a limiting statement for $(\Phi^g_{-s}\Phi^f_{-t}\Phi^g_s\Phi^f_t)(q)$ as $(t,s)\to(0,0)$.

Remark 1.1. Formula (4) is applicable, in particular, when f and g are continuous near q and classically differentiable at q. In that case, taking F and G to be first-order linear approximations of f and g near q, (4) implies

$$\lim_{t\to 0,\, t\neq 0}\frac{1}{t^2}\operatorname{dist}\big([f,g](q),\big(\Phi^g_{-t}\Phi^f_{-t}\Phi^g_t\Phi^f_t\big)(q)-q\big)=0. \tag{5}$$

In the special case when f and g are both Lipschitz near q and classically differentiable at q, formula (3) applies, and formula (5) is also applicable. The set $\{[f,g](q)\}$ is, in general, smaller than $[f,g]_{\mathrm{set}}(q)$, so the approximation result of (5) is better than the one obtained from (3) by taking $s=t$.

Our generalization of (II) will be the formula

$$\begin{aligned}&\big((\forall q\in M)(\exists\varepsilon>0)\big(\forall t,s\in[-\varepsilon,\varepsilon]\big)\big(\Phi^g_{-s}\Phi^f_{-t}\Phi^g_s\Phi^f_t\big)(q)=q\big)\\ &\quad\Longleftrightarrow\quad\big((\forall q\in M)[f,g]_{\mathrm{set}}(q)=0\big)\\ &\quad\Longleftrightarrow\quad\big([f,g](q)=0\text{ for a.e. }q\big),\end{aligned} \tag{6}$$

respectively. The statement generalizing (III) will be identical to (III), except only for the fact that "of class C^1" will be replaced by "Lipschitz."

The paper is organized as follows. In Section 2 we introduce some basic definitions and notations. In particular, in Sections 2.2–2.4, we present a self-contained introduction (with an example) to the "Agrachev–Gamkrelidze formalism," which will be used in many parts of the present paper.[3] In Section 3 we review the notion of set-valued bracket introduced in [7]. In Section 4 we derive asymptotic formulae similar to (1) for vector fields which are not C^1, and in particular (a) we prove (3) for locally Lipschitz vector fields, and (b) we obtain (4) for "semidifferentiable" vector fields. In Section 5 we prove a commutativity result (Theorem 5.3) for locally Lipschitz vector fields, which, in particular, yields the characterization (6). In Section 6, using the result on commutativity, we will prove the Lipschitz analogue of the simultaneous flow-box result (III) (cf. Theorem 6.1). Finally, in Section 7 we discuss the difficulties that arise when one tries to define higher-order brackets such as $[f,[g,h]]$ under minimal regularity assumptions, and show that the most obvious approach (in which, for example, one uses $[f,[g,h]]_{\mathrm{set}}$ as the nonsmooth analogue of $[f,[g,h]]$ if f is locally Lipschitz and g,h are of class C^1 with locally Lipschitz derivatives) does not lead to a good theory. We do this by constructing an example in which the asymptotic formula

$$\big(\Phi^f_{-t}\big(\Phi^h_{-t}\Phi^g_{-t}\Phi^h_t\Phi^g_t\big)^{-1}\Phi^f_t\big(\Phi^h_{-t}\Phi^g_{-t}\Phi^h_t\Phi^g_t\big)\big)(q)=q+t^3\big[f,[g,h]\big]_{\mathrm{set}}(q)+o\big(t^3\big),$$

[3] This formalism, introduced in a series of papers by A. Agrachev and R. Gamkrelidze, will be very convenient in computations involving compositions of several flow maps. Following [6], we include here a brief outline of the formalism and its rigorous justification, together with an example of a computation. The readers who wish to move on quickly to the results of the paper should just read Sections 2.2 and 2.3, skipping the justification provided in Section 2.4.

is not true. We conclude from this that a different definition of higher-order brackets is needed, but leave the full discussion of that definition and its properties to a subsequent paper.

2. Preliminary definitions and notational conventions

As usual, $\mathbb{Z}$ denotes the set of all integers. We write $\mathbb{Z}_+ = \{n \in \mathbb{Z}\colon n \geqslant 0\}$, $\mathbb{N} = \{n \in \mathbb{Z}\colon n \geqslant 1\}$, $\overline{\mathbb{Z}}_+ = \mathbb{Z}_+ \cup \{\infty\}$, $\overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$.

For any $n \in \mathbb{N}$, we use $\mathbb{R}^n$, $\mathbb{B}^n$, $\overline{\mathbb{B}}^n$ to denote, respectively, the space of all real n-dimensional column vectors, and the open and closed Euclidean unit balls $\{x \in \mathbb{R}^n\colon \|x\| < 1\}$, $\{x \in \mathbb{R}^n\colon \|x\| \leqslant 1\}$. For $x \in \mathbb{R}^n$ and $\rho > 0$, $x + \rho\mathbb{B}^n$, $x + \rho\overline{\mathbb{B}}^n$, will denote the open and closed balls of radius ρ and center x. We write $\rho\mathbb{B}^n$, $\rho\overline{\mathbb{B}}^n$, instead of $0 + \rho\mathbb{B}^n$, $0 + \rho\overline{\mathbb{B}}^n$. We use $e_1^n, \ldots, e_n^n$ to denote the members of the canonical basis of $\mathbb{R}^n$, so that $e_j^n = (\delta_j^1, \ldots, \delta_j^n)^\dagger$, where δ_j^i is Kronecker's delta, and † stands for "transpose."

If $\ell \in \overline{\mathbb{Z}}_+$, a *manifold of class* C^ℓ is a finite-dimensional, second countable, Hausdorff, differentiable (if $\ell > 0$) manifold of class C^ℓ. If M is an m-dimensional manifold of class C^1, and $\kappa : U \mapsto \mathbb{R}^m$ is a coordinate chart on M, then for each $j \in \{1, \ldots, m\}$ we use ∂_j^κ to denote the jth element of the canonical basis of vector fields on U corresponding to κ, so that, for example, if $f \in C^1(U, \mathbb{R})$ then $\partial_j^\kappa f$ is the function $\frac{\partial(f \circ \kappa^{-1})}{\partial x^j} \circ \kappa$, from $\kappa(U)$ to $\mathbb{R}$.

If A, B are real linear spaces, then $\mathcal{L}(A, B)$ denotes the space of all $\mathbb{R}$-linear maps from A to B.

2.1. Lipschitz maps

If E, F are metric spaces, with distance functions d_E, d_F, then a map: $m : E \mapsto F$ is *Lipschitz* if there exists $L \in \mathbb{R}$ such that $d_F(m(e_1), m(e_2)) \leqslant L d_E(e_1, e_2)$ for all $e_1, e_2 \in U$. (In that case the number L is a *Lipschitz constant* for m.) We say that m is *locally Lipschitz* if every $e \in E$ has a neighborhood U such that the restriction of m to U is Lipschitz. We say that m is a *lipeomorphism* if it is a bijection and both m and the inverse map $m^{-1} : F \mapsto E$ are locally Lipschitz.

Assume that $\ell \in \overline{\mathbb{N}}$, and N, M are manifolds of class C^ℓ and dimensions n, m. A map $f : N \mapsto M$ is *locally Lipschitz* if it is continuous and such that for every pair (ξ, η) of coordinate charts $\xi : U \mapsto \mathbb{R}^n$, $\eta : V \mapsto \mathbb{R}^m$ defined on open subsets U, V of N, M, the map $f^{\xi,\eta} = \eta \circ f \circ \xi^{-1} : \xi(U \cap f^{-1}(V)) \mapsto \mathbb{R}^m$ is locally Lipschitz. (It is easily shown that f is locally Lipschitz if and only if for every $\bar{q} \in N$ there exist charts ξ, η, defined on open neighborhoods U, V of $\bar{q}$, $f(\bar{q})$, such that $f(U) \subseteq V$ and $f^{\xi,\eta}$ is Lipschitz.) The well-known Rademacher theorem implies that if f is a locally Lipschitz map then it is differentiable almost everywhere, that is, $DIFF(f)$ is a full subset of N, where $DIFF(f)$ is the set of points $q \in N$ such that f is differentiable at q. (A *full subset* of N a subset F of N such that $N \setminus F$ is a null subset of N. A *null subset* of N is a subset S of N such that $\xi(U \cap S)$ is a subset of $\mathbb{R}^n$ of zero Lebesgue measure whenever $\xi : U \mapsto \mathbb{R}^n$ is a chart of N.)

Remark 2.1. Since all Riemannian metrics are locally equivalent on a manifold of class C^1, it is clear that a map $F : N \mapsto M$ is locally Lipschitz if and only if it is locally Lipschitz as a map between the metric spaces (N, d_{g_N}) and (M, d_{g_M}), where g_N, g_M are arbitrary Riemannian metrics on N and M, and d_{g_N}, d_{g_M} are the corresponding distance functions.

2.2. The Agrachev–Gamkrelidze formalism

In a series of papers (cf., e.g., [1,2]), A. Agrachev and R. Gamkrelidze proposed a very convenient formalism, henceforth referred to as the *Agrachev–Gamkrelidze formalism* (and abbreviated as AGF), for computations involving flow maps arising from various time-varying vector fields, based on "chronological exponentials." We now present an outline of this formalism, following [6].

The crucial point of the AGF is to write the pairing of a contravariant object q and a covariant object p consistently as qp. For example, points of a manifold M and tangent vectors to M are contravariant objects, while functions and differential forms are covariant objects, so in the AGF the value of a function φ at a point q is written $q\varphi$ rather than $\varphi(q)$. Similarly, the result of applying a tangent vector v at a point q to a function φ (i.e., the directional derivative at q of φ in the direction of v) is written $v\varphi$. Vector fields are first-order differential operators, acting on functions on the left. Hence they should act on points on the right, so we write qf rather than $f(q)$ for the value at q of a vector field f, and then $qf \in T_qM$. With this notation, $q(f\varphi)$ is the value at q of the function $f\varphi$, while the result of applying the tangent vector qf to the function φ is $(qf)\varphi$. It is clear that $(qf)\varphi = q(f\varphi)$, so we just write $qf\varphi$, omitting the parentheses.

A vector field f on a manifold M generates a one-parameter family $\{e^{tf}\}_{t\in\mathbb{R}}$ of possibly partially defined maps from M to M. Since f acts on points on the right, the maps e^{tf} should also be written as acting on the right, so we write qe^{tf} rather than $e^{tf}(q)$ or $e^{tf}q$. Then $t \mapsto qe^{tf}$ is the integral curve ξ of f that goes through q at time 0. The equation that would be written classically as $(d/dt)(\xi(t)) = f(\xi(t))$ now becomes $(d/dt)(qe^{tf}) = qe^{tf}f$.

More generally, a map Φ from M to another manifold N is written as acting on points on the right, so we write $q\Phi$ rather than $\Phi(q)$. (Notice that the notation qf for a vector field f is consistent with this more general convention, since f is a map from M to TM.) Maps also act on tangent vectors. If $q \in M$, $v \in T_qM$, and $\Phi : M \mapsto N$, then $v\Phi$ is the tangent vector at $q\Phi$ known as the *push-forward* of v, and often represented in the literature by expressions such as $D\Phi \cdot v$, or $D\Phi(q) \cdot v$, or $D\Phi(q)(v)$, or $\Phi_*(v)$, or $\Phi_* v$.

The dual action of maps on functions is written as a left action. Thus, if $\Phi : M \mapsto N$, and φ is a function on N, then $\Phi\varphi$ is the *pullback* of φ by Φ, i.e., the function $\varphi \circ \Phi$, sometimes written as $\Phi^*(\varphi)$. Then the identity $(\varphi \circ \Phi)(q) = \varphi(\Phi(q))$ simply says that $q(\Phi\varphi) = (q\Phi)\varphi$, so we simply write $q\Phi\varphi$, omitting the parentheses. Furthermore, the usual definition of the push-forward $\Phi_*(v)$ of a tangent vector says that $\Phi_*(v)\varphi = v(\varphi \circ \Phi)$. In the AGF, this just becomes $(v\Phi)\varphi = v(\Phi\varphi)$, so we can simply write $v\Phi\varphi$, omitting the parentheses.[4]

In particular, if f is a vector field on M, φ is a function on M, and $t \in \mathbb{R}$, then the action of the flow map e^{tf} on φ is written on the left, as $e^{tf}\varphi$, so $(qe^{tf})\varphi = q(e^{tf}\varphi)$, and we may just write $qe^{tf}\varphi$, omitting the parentheses.

The product $f_1 f_2 \cdots f_k$ of several vector fields is a differential operator, which acts on functions on the left and on points on the right. For example, if f, g are vector fields of class C^1, then fg is a second-order differential operator with continuous coefficients (given in a coordinate chart $\kappa : U \to \mathbb{R}^m$, if $f = \sum_i f^i \partial_i^\kappa$, $g = \sum_j g^j \partial_j^\kappa$, by $fg = \sum_{i,j}(f^i(\partial_i^\kappa g^j)\partial_j^\kappa + f^i g^j \partial_i^\kappa \partial_j^\kappa)$), and qfg is the operator fg at the point q, i.e., the map that sends every function φ to the value of $fg\varphi$ at q, i.e., to $qfg\varphi$. The difference $[f, g] = fg - gf$—the *Lie bracket* of f and g—is also, in principle, a second-order differential operator, but $[f, g]$ happens, in fact, to be first-order, i.e.,

[4] Notice that in the AGF the notations $\Phi\varphi$, $v\Phi$, for pullback and push-forward correctly place the symbol Φ in the "back" and "forward" positions.

a continuous vector field, given in coordinates by $[f,g]=\sum_{i,j}(f^i(\partial_i^\kappa g^j)\partial_j^\kappa - g^i(\partial_i^\kappa f^j)\partial_j^\kappa)$, that is, by $[f,g]=\sum_j h_j\partial_j^\kappa$, where $h_j=\sum_i(f^i(\partial_i^\kappa g^j)-g^i(\partial_i^\kappa f^j))$.

If follows that a complicated expression such as $q\Phi f e^{tg}h\Psi e^{tk}\ell$ makes perfect sense, if M,N,P are manifolds, $\Phi:M\mapsto N$, f,g,h are vector fields on N, $\Psi:N\mapsto P$, and k,ℓ are vector fields on P. The precise meaning of this expression is as the map that takes a function φ on P, applies to it the first-order differential operator ℓ, pulls back the function $\ell\varphi$ by the map e^{tk}, then pulls back the function $e^{tk}\ell\varphi$ by Ψ, then applies to the resulting function the differential operator h, pulls back the function $h\Psi e^{tk}\ell\varphi$ by the map e^{tg}, applies to the function $e^{tg}\Psi h e^{tk}\ell\varphi$ the differential operator f, then pulls back $f e^{tg}\Psi h e^{tk}\ell\varphi$ by Φ and, finally, evaluates the resulting function $\Phi f e^{tg}h e^{tk}\ell\varphi$ at q.

Remark 2.2. Once it is understood that to a manifold M are associated two dual kinds of entities, namely, "test-function-like," or "covariant" objects, and "contravariant" ones, it becomes clear that the formalisms often used in textbooks are somewhat inconsistent, because the result of pairing a point q and a test function φ is usually written as $\varphi(q)$, whereas that of pairing a tangent vector v and a test function φ is usually written as $v\varphi$. The AGF is truly consistent, in that it always uses the notation qp for the result of pairing a contravariant object q and a covariant object p.

From now on, we will use the AGF whenever doing so is more convenient for calculations. But we will revert to the classical notation in many cases when using the AGF is unnecessary and the classical notation is preferable. (*For example, if $\gamma:\mathbb{R}\mapsto M$ is a curve, we will use $\gamma(t)$ rather than the AGF expression $t\gamma$.*) *We will even mix the formalisms, by writing, for example, formulae such as $\dot\gamma(t)=\gamma(t)X$* (*rather than the fully AGF equality $t\partial_t\gamma=t\gamma X$, or the fully classical identity $\dot\gamma(t)=X(\gamma(t))$*) *if γ is an integral curve of a vector field X. In all cases, the resulting formulae will be completely unambiguous.*

2.3. An example

With the AGF, many important formulae involving vector fields, their exponentials, and their Lie brackets, become completely trivial formally, and the formal calculations can be rigorously justified using the distributional interpretation, as will be explained in Section 2.4 below. We illustrate this with an example.

Let M be a manifold of class C^2, let $f_1,\dots,f_d$ be vector fields of class C^1 on M, and let $q\in M$. We will compute the first and second derivatives $\dot\gamma(0)$, $\ddot\gamma(0)$, at $t=0$ of the curve γ given by $\gamma(t)=q\Pi(t)$, where $\Pi(t)$ is the product $\Pi(t)=e^{tf_1}e^{tf_2}\cdots e^{tf_d}$.

We have

$$(d/dt)\Pi(t)=\sum_{i=1}^{d}e^{tf_1}\cdots e^{tf_i}f_ie^{tf_{i+1}}\cdots e^{tf_d},$$

$$\begin{aligned}\left(d^2/dt^2\right)\Pi(t)&=\sum_{i=1}^{d}\sum_{j=1}^{i}e^{tf_1}\cdots e^{tf_j}f_je^{tf_{j+1}}\cdots e^{tf_i}f_ie^{tf_{i+1}}\cdots e^{tf_d}\\&\quad+\sum_{i=1}^{d}\sum_{j=i+1}^{d}e^{tf_1}\cdots e^{tf_i}f_ie^{tf_{i+1}}\cdots e^{tf_j}f_je^{tf_{j+1}}\cdots e^{tf_d}\end{aligned}$$

and then

$$\begin{aligned}
\dot\gamma(0) &= q\sum_{i=1}^{d} f_i,\\
\ddot\gamma(0) &= q\left(\sum_{i=1}^{d}\left(\sum_{j=1}^{i} f_j\right) f_i + \sum_{i=1}^{d} f_i\left(\sum_{j=i+1}^{d} f_j\right)\right)\\
&= q\left(\sum_{i=1}^{d}\left(\sum_{j=1}^{d} f_j\right) f_i + \sum_{i=1}^{d} f_i\left(\sum_{j=i+1}^{d} f_j\right) - \sum_{i=1}^{d}\left(\sum_{j=i+1}^{d} f_j\right) f_i\right)\\
&= q\left(\sum_{i=1}^{d} f_i\right)^2 + \sum_{i<j} q[f_i, f_j].
\end{aligned}$$

In particular, this shows that $\ddot\gamma(0) = \sum_{i<j} q[f_i, f_j]$ if $\dot\gamma(0) = 0$, which is a special case of the general principle that "when the tangent vector to a curve γ at time 0 vanishes, then the second derivative $\ddot\gamma(0)$ is a tangent vector."

If we let $d = 4$, $f_1 = f$, $f_2 = g$, $f_3 = -f$, $f_4 = -g$, then $\sum_{i=1}^{d} f_i = 0$, and $\sum_{i<j} q[f_i, f_j] = [f, g] + [f, -f] + [f, -g] + [g, -f] + [g, -g] + [-f, -g] = 2[f, g]$, so

$$\left.\frac{d}{dt}\right|_{t=0}\left(qe^{tf}e^{tg}e^{-tf}e^{-tg}\right) = 0, \qquad \left.\frac{d^2}{dt^2}\right|_{t=0}\left(qe^{tf}e^{tg}e^{-tf}e^{-tg}\right) = 2q[f, g],$$

from which we get the asymptotic formula

$$qe^{tf}e^{tg}e^{-tf}e^{-tg} = q + t^2 q[f, g] + o\left(t^2\right),$$

so that

$$\lim_{t\to 0}\frac{qe^{tf}e^{tg}e^{-tf}e^{-tg} - q}{t^2} = q[f, g]. \tag{7}$$

2.4. *Justification of the AGF*

The rigorous justification of the formalism discussed above is obtained by regarding all "contravariant" objects such as points, tangent vectors, and differential operators evaluated at a point, as *distributions*, i.e., as members of the dual of a suitable space of test functions.

We now make this precise. Assume that $\ell \in \overline{\mathbb{Z}}_+$, $m \in \mathbb{Z}_+$, and M is an m-dimensional manifold of class C^ℓ. We use $\mathcal{E}^\ell(M)$ to denote the commutative $\mathbb{R}$-algebra of real-valued functions of class C^ℓ on M, topologized in the usual way. (A sequence $\{\varphi_j\}_{j\in\mathbb{N}}$ converges to a limit φ in $\mathcal{E}^\ell(M)$ if $\varphi_j \to \varphi$ uniformly on compact sets, and for every $k \in \mathbb{N}$ such that $k \leqslant \ell$ and every k-tuple $(X_1, \ldots, X_k)$ of vector fields of class C^ℓ on M the functions $X_1X_2\ldots X_k\varphi_j$ converge to $X_1X_2\ldots X_k\varphi$ uniformly on compact sets.) We let $\mathcal{E}'^\ell(M)$ denote the dual space of $\mathcal{E}^\ell(M)$, i.e., the space of compactly supported Schwarz distributions on M or order ℓ. We remark that, in particular, $\mathcal{E}^k(M)$ and $\mathcal{E}'^k(M)$ are well defined for all $k \in \mathbb{Z}_+$ such that $k \leqslant \ell$, because a manifold of class C^ℓ has a canonical structure of class C^k whenever $k \leqslant \ell$. If $j \leqslant k \leqslant \ell$, then $\mathcal{E}^k(M)$ is a dense subspace of $\mathcal{E}^j(M)$ whenever $j \leqslant k \leqslant \ell$, and the inclusion from $\mathcal{E}^k(M)$ to $\mathcal{E}^j(M)$ is

continuous; it follows that $\mathcal{E}'^{j}(M)$ is canonically embedded in $\mathcal{E}'^{k}(M)$. It is clear that $\mathcal{E}'^{0}(M)$ is the space of signed Borel measures on M that have compact support.

Every point q of M gives rise to a linear functional $\delta_q \in \mathcal{E}'^{0}(M)$—the Dirac delta function at q—defined by letting $\delta_q(\varphi) = \varphi(q)$ for $\varphi \in \mathcal{E}^0(M)$. The map $M \ni q \mapsto \delta_q \in \mathcal{E}'^{0}(M)$ is clearly injective, so we can use this map to regard M as embedded in $\mathcal{E}'^{0}(M)$, and then M is embedded in $\mathcal{E}'^{k}(M)$ whenever $k \leqslant \ell$.

We endow each space $\mathcal{E}'^{k}(M)$ with the weak* topology arising from the duality with $\mathcal{E}^k(M)$, so a net $\{\nu_\alpha\}_{\alpha\in A}$ of members of $\mathcal{E}'^{k}(M)$ converges to a $\nu \in \mathcal{E}'^{k}(M)$ if and only if the net $\{\nu_\alpha(\varphi)\}_{\alpha\in A}$ converges to $\nu(\varphi)$ for every $\varphi \in \mathcal{E}^k(M)$. Then many linear operations and limiting processes that in principle appear not to make intrinsic sense on M become completely meaningful in the spaces $\mathcal{E}'^{k}(M)$. It follows that, in addition to the points of M, many other objects related to M can also be naturally regarded as members of $\mathcal{E}'^{\ell}(M)$. For example:

(1) If $\ell > 0$, $\gamma : [0, \varepsilon] \mapsto M$ is a curve of class C^1, and $\gamma(0) = q$, then the limit

$$\dot\gamma(0) = \lim_{h\downarrow 0} \frac{\gamma(h) - q}{h}, \tag{8}$$

makes perfect sense as a limit in $\mathcal{E}'^{1}(M)$, where $\gamma(h)$, q mean, naturally, the Dirac delta functions of the points $\gamma(h)$, q. So $\dot\gamma(0)$ (that is, the functional $\mathcal{E}^1(M) \ni \varphi \mapsto \lim_{h\downarrow 0} h^{-1}(\varphi(\gamma(h)) - \varphi(q))$) is a well defined member of $\mathcal{E}'^{1}(M)$.
Thus formula (8), which is the natural way to define $\dot\gamma(0)$ when $M = \mathbb{R}^m$, remains perfectly meaningful as written—and gives the right answer—for a general manifold M, provided only that it is properly reinterpreted. (In particular, there is no need to define $\dot\gamma(0)$ in a more roundabout way by, for example, writing (8) with respect to some fixed coordinate chart, and then proving that the resulting tangent vector does not depend on the chart.)

(2) The tangent bundle TM is embedded in $\mathcal{E}'^{1}(M)$ as follows. The tangent space T_qM of M at a point $q \in M$ is, by definition, the set of all linear functionals $v : \mathcal{E}^1(M) \mapsto \mathbb{R}$ such that $v = \dot\gamma(0)$ for some curve $\gamma : [0, \varepsilon] \mapsto M$ of class C^1 such that $\gamma(0) = q$. Hence T_qM is already a linear subspace of $\mathcal{E}'^{1}(M)$.

(3) Similarly, if we use PDO^k_qM, for $k \leqslant \ell$, to denote the set of all partial differential operators of order $\leqslant k$ at q (so that $V \in PDO^k_qM$ if and only if V is a map $\mathcal{E}^k(M) \ni \varphi \mapsto V\varphi \in \mathbb{R}$ given, for some coordinate chart $\kappa : U \mapsto \mathbb{R}^m$ such that $q \in U$, by

$$V\varphi = a_\emptyset q\varphi + \sum_{\nu=1}^{k}\sum_{i_1=1}^{m}\sum_{i_2=1}^{m}\cdots\sum_{i_\nu=1}^{m} a_{i_1,i_2,\ldots,i_\nu}\left(q\,\partial^\kappa_{i_1}\partial^\kappa_{i_2}\cdots\partial^\kappa_{i_\nu}\varphi\right),$$

where the coefficients $a_{i_1,i_2,\ldots,i_\nu}$ are real numbers), then each PDO^k_qM is automatically a linear subspace of $\mathcal{E}'^{k}(M)$, and then it follows that the set $PDO^kM \stackrel{\text{def}}{=} \bigcup_{q\in M} PDO^k_pM$ is a subset of $\mathcal{E}'^{k}(M)$.

To justify rigorously the use of the AGF, it suffices to regard a manifold M as embedded in $\mathcal{E}'^{k}(M)$ as explained before. Then

- If $q \in M$ and $\varphi \in \mathcal{E}^0(M)$, then $q\varphi$ is simply an alternative way of writing $\varphi(q)$, or $\delta_q(\varphi)$, or $\delta_q\varphi$.

- If $\ell \geqslant 1$, $q \in M$, $v \in T_q M$, and $\varphi \in \mathcal{E}^1(M)$, then the notation $v\varphi$ for the directional derivative of φ at q in the direction of v (which, in this case, is the one commonly used in textbooks) reflects the fact that $v \in \mathcal{E}'^1(M)$.
- If f is a vector field on M (i.e., a section of the tangent bundle TM) and $\varphi \in \mathcal{E}^1(M)$, then $f\varphi$ is a well defined function on M, which belongs to $\mathcal{E}^{k-1}(M)$ if $0 < k \leqslant \ell$, f is a vector field of class C^{k-1}, and $\varphi \in \mathcal{E}^k(M)$.
- If M, N are manifolds of class C^ℓ, and Φ is a map from M to N, we have already explained that Φ is written as acting on points of M on the right, so the AGF notation for $\Phi(q)$, if $q \in M$, is $q\Phi$. If Φ is continuous then the dual action of Φ on test functions is the "pulling back" map $\mathcal{E}^0(N) \ni \varphi \mapsto \varphi \circ \Phi \in \mathcal{E}^0(M)$. In the AGF, we write $\Phi\varphi$ rather than $\varphi \circ \Phi$. If Φ is of class C^k, then $\Phi\varphi \in \mathcal{E}^k(M)$ whenever $\varphi \in \mathcal{E}^k(N)$, and the map $\mathcal{E}^k(N) \ni \varphi \mapsto \Phi\varphi \in \mathcal{E}^k(M)$ is linear and continuous, so its adjoint (i.e., the map $\mathcal{E}'^k(M) \ni \mu \mapsto \mu\Phi \in \mathcal{E}'^k(N)$, where $\mu\Phi$ is the map $\mathcal{E}^k(N) \ni \varphi \mapsto \mu(\Phi\varphi) \in \mathbb{R}$) is linear and continuous as well. If μ belongs to $\mathcal{E}'^k(M)$, then μ is a compactly supported distribution on M of order k, and $\mu\Phi$ is the "push-forward" of μ, which is a compactly supported distribution of order k on M (with support contained in the set $\Phi(\operatorname{supp}\mu)$, i.e., $(\operatorname{supp}\mu)\Phi$ in AGF notation).
  It follows that *the "pushing forward" map $\mathcal{E}'^k(M) \ni \mu \mapsto \mu\Phi \in \mathcal{E}'^k(N)$ is the unique linear continuous extension to $\mathcal{E}'^k(M)$ of the original map $\Phi : M \mapsto N \subseteq \mathcal{E}'^k(N)$*. This justifies using the same name Φ for the pushing forward map,
  In particular, if $v \in T_q M$ for some $q \in M$, and $k > 0$, then $v\Phi$ makes sense. Since the map $\mathcal{E}'^k(M) \ni \mu \mapsto \mu\Phi \in \mathcal{E}'^k(N)$ is linear and continuous, formula (8) implies that, if $\gamma : [0, \varepsilon] \mapsto M$ is a curve of class C^1, and $\gamma(0) = q$, then

$$\dot\gamma(0)\Phi = \lim_{h \downarrow 0} \frac{\gamma(h)\Phi - q\Phi}{h}, \tag{9}$$

  so $\dot\gamma(0)\Phi = \dot\eta(0)$, where η is the curve $t \mapsto \gamma(t)\Phi$, i.e., $\eta = \Phi \circ \gamma$.

3. Lie brackets of locally Lipschitz vector fields

Let M be a manifold of class C^2, and let f, g be vector fields of class C^1 on M. We write $[f, g]$ to denote the difference $fg - gf$ which, as we have already pointed out, is a continuous vector field, called the *Lie bracket* of f and g.

In [7] we proposed the following extension of the notion of Lie bracket to the case when the vector fields f and g are only locally Lipschitz.[5] First of all, we point out that $q[f, g]$ is a well defined tangent vector at q for each point q belonging to $DIFF(f) \cap DIFF(g)$. (Recall that the sets $DIFF(X)$ were defined in Section 2.1.)

Definition 3.1. Let f, g be locally Lipschitz vector fields on a manifold M of class C^2. The *Lie bracket* of f and g is the set-valued section $[f, g]_{\text{set}}$ of the tangent bundle TM constructed as follows. For every $q \in M$ we let $q[f, g]_{\text{set}}$—*the Lie bracket of f and g at q*—be the convex hull of the set of all vectors

$$v = \lim_{j \to \infty} q_j[f, g], \tag{10}$$

for all sequences $\{q_j\}_{j \in \mathbb{N}}$ such that

[5] See [8] for a different kind of Lie bracket, which happens to be defined almost everywhere.

1. $q_j \in DIFF(f) \cap DIFF(g)$ for all j,
2. $\lim_{j\to\infty} q_j = q$,
3. the limit v of (10) exists.

Remark 3.2 *(An equivalent definition).* For every full subset $\mathcal{F} \subseteq M$, one can equivalently define the set $q[f,g]_{\text{set}}$ by replacing condition 1 in Definition 3.1 with the weaker condition

$1^{\mathcal{F}}$. $q_j \in DIFF(f) \cap DIFF(g) \cap \mathcal{F}$.

A proof of this fact can be obtained by (a) writing $q[f,g]_{\text{set}}$ in terms of the Clarke generalized Jacobian $\partial h(q)$ (at q) of the map $h = (f,g)$—see Remark 3.6 below—and (b) recalling that, for every full subset $\mathcal{F} \subseteq M$, one can obtain $\partial h(q)$ by only considering the sequences $\{q_j\}_{j\in\mathbb{N}}$ ranging in $DIFF(h) \cap \mathcal{F}$.[6]

Proposition 3.3. *Let f, g be locally Lipschitz vector fields on a manifold M of class C^2. Then $q \mapsto q[f,g]_{\text{set}}$ is an upper semicontinuous set-valued map such that, for every $q \in M$, $q[f,g]_{\text{set}}$ is a convex, compact, nonempty subset of T_qM.*[7] *Moreover, the* skew-symmetry identity

$$q[f,g]_{\text{set}} = -q[g,f]_{\text{set}} \tag{11}$$

holds for all $q \in M$.[8] *In addition, each locally Lipschitz vector field g satisfies the identity*

$$q[g,g]_{\text{set}} = \{0\} \quad \textit{for every } q \in M. \tag{12}$$

Proof. Identities (11) and (12) are straightforward consequences of Definition 3.1 and the skew-symmetry of the ordinary Lie bracket. (But notice that (12) is *not* a direct consequence of (11), because if a set S is such that $-S = S$ it does not follow that $S = \{0\}$.)

The convexity of the sets $q[f,g]$ follows directly from the definition.

If $S = \bigcup_{q\in DIFF(f)\cap DIFF(g)} q[f,g]$, and $\overline{S}$ is the closure of S in TM, then each set $\overline{S}(q) = T_qM \cap \overline{S}$ is compact. By definition, $q[f,g]_{\text{set}}$ is the convex hull of $\overline{S}(q)$, so $q[f,g]_{\text{set}}$ is compact.

The fact that $q[f,g]_{\text{set}} \neq \emptyset$ follows from (i) Rademacher's theorem, which implies that $DIFF(f) \cap DIFF(g)$ is a full subset of M, from which it follows in particular that $DIFF(f) \cap DIFF(g)$ is dense in M, together with (ii) the local Lipschitz property of f and g, which implies that any sequence $\{(q_j, v_j)\}_{j\in\mathbb{N}}$ such that $q_j \to q$, $q_j \in DIFF(f) \cap DIFF(g)$, and $v_j = q_j[f,g]$, has a convergent subsequence.

Finally, it is easy to show that the graph[9] $\bigcup_{q\in M} q[f,g]_{\text{set}}$ is a closed subset of TM, so the set-valued map $q \mapsto q[f,g]_{\text{set}}$ is upper semicontinuous. □

---

6 See, e.g., [4, Theorem 2.5.1].

7 We recall that, if E and F are topological spaces, then a set-valued map $\mu : E \mapsto F$ is *upper semicontinuous* if the set $\mu^{-1}(C) = \{x \in E\colon \mu(x) \cap C \neq \emptyset\}$ is closed whenever C is a closed subset of F. If μ has compact values, then μ is upper semicontinuous if and only if the graph $\bigcup_{e\in E}\{e\} \times \mu(e)$ is a closed subset of $E \times F$.

8 This means that $q[f,g]_{\text{set}} = \{w\colon -w \in q[g,f]_{\text{set}}\}$.

9 When $\sigma : E \to F$ is a section of a bundle F over a topological space E, we define the *graph* of σ to be the set $\{\sigma(e)\colon e \in E\}$, rather than the set $\{(e,\sigma(e))\colon e \in E\}$, because the fibers F_e of F are pairwise disjoint, so $\sigma(e)$ already determines e.

Remark 3.4. If $q \in DIFF(f) \cap DIFF(g)$, then the set $q[f,g]_{\text{set}}$ *does not coincide*, in general, with the singleton $\{q[f,g]\}$ even though the latter is obviously a subset of the former. For example, let $M = \mathbb{R}$ and let us consider the locally Lipschitz vector fields f, g defined by $f = \partial_x$, $g = \alpha(x)\partial_x$, where

$$\alpha(x) = \begin{cases} x^2 \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then, if we take $q = 0$, it is clear that $q[f,g] = 0$, while on the other hand, $q[f,g]_{\text{set}} = [-1,1]$.

Remark 3.5. If f and g are of class C^1 near q, then $q[f,g]_{\text{set}} = \{q[f,g]\}$.

Remark 3.6. There is a simple relationship between the set-valued Lie bracket and the notion of Clarke generalized Jacobian of a map. Let us recall that, if M and N are manifolds of class C^1, $h : M \mapsto N$ is locally Lipschitz, and $q \in M$, then the *Clarke generalized Jacobian* of h at q is the subset $\partial h(q)$ of $\mathcal{L}(T_qM, T_{h(q)}N)$ defined as follows. First, we let $\tilde{\partial} h(q)$ be the set of all linear maps $L \in \mathcal{L}(T_qM, T_{h(q)}N)$ such that $L = \lim_{j\to\infty} Dh(q_j)$ for some sequence $\{q_j\}_{j\in\mathbb{N}}$ such that

1. $q_j \in DIFF(h)$ for all j,
2. $\lim_{j\to\infty} q_j = q$,
3. the limit[10] $\lim_{j\to\infty} Dh(q_j)$ exists.

Then $\partial h(q)$ is the convex hull of $\tilde{\partial} h(q)$.

In the special case when f is a locally Lipschitz vector field on an m-dimensional manifold M, we can take $N = TM$, so $\partial f(q)$ is a subset of the $2m^2$-dimensional linear space $\mathcal{L}(T_qM, T_{f(q)}TM)$. If $\pi_M : TM \mapsto M$ is the canonical projection, then $\pi_M \circ f = \mathrm{id}_M$, the identity map of M. So the equality $D\pi_M(f(q)) \circ Df(q) = \mathrm{id}_{T_qM}$ holds whenever q belongs to $DIFF(f)$. It follows that, for each $q \in DIFF(f)$, $Df(q)$ belongs to the m^2-dimensional affine subspace $\mathcal{L}^0(T_qM, T_{f(q)}TM)$ of the space $\mathcal{L}(T_qM, T_{f(q)}TM)$ whose members are the linear maps $L : T_qM \mapsto T_{f(q)}TM$ such that $D\pi_M \circ L = \mathrm{id}_{T_qM}$. By taking limits, it follows that $\tilde{\partial} f(q)$ is a subset of $\mathcal{L}^0(T_qM, T_{f(q)}TM)$ for every $q \in M$, and then the convex hull $\partial f(q)$ (which makes sense because $\mathcal{L}^0(T_qM, T_{f(q)}TM)$ is an affine space, though not a linear one) is a subset of $\mathcal{L}^0(T_qM, T_{f(q)}TM)$.

When M is an open subset of $\mathbb{R}^m$, then a vector field f on M is just a map from M to $\mathbb{R}^m$, so the sets $\partial f(q)$, for $q \in M$, can be regarded as subsets of $\mathbb{R}^{m\times m}$, the space of m by m real matrices. In this situation, it might appear natural to define a "Lie bracket" $[f,g]_C$ of two locally Lipschitz vector fields, by analogy with the formula $[f,g](q) = Dg(q)\cdot f(q) - Df(q)\cdot g(q)$, by letting

$$[f,g]_C(q) = \partial g(q)\cdot f(q) - \partial f(q)\cdot g(q) \tag{13}$$

(that is, $[f,g]_C(q) = \{(B\cdot f(q) - A\cdot g(q)) \colon (A,B) \in \partial f(q) \times \partial g(q)\}$).

[10] The limit is taken in $\Lambda(M,N)$, the bundle over $M \times N$ whose fiber at $(q,r) \in M \times N$ is $\mathcal{L}(T_qM, T_rN)$. Clearly, $\Lambda(M,N)$ is a manifold of class $C^{\ell-1}$ if M, N are of class C^ℓ.

This does *not* yield our set-valued bracket $[f,g]_{\mathrm{set}}$. The correct formula for $[f,g]_{\mathrm{set}}$ in terms of Clarke Jacobians is

$$[f,g]_{\mathrm{set}}(q) = \bigl\{\bigl(B\cdot f(q) - A\cdot g(q)\bigr)\colon\ (A,B)\in\partial(f,g)(q)\bigr\},$$

where (f,g) is the map $M\ni q\mapsto (f(q),g(q))\in\mathbb{R}^m\times\mathbb{R}^m\sim\mathbb{R}^{2m}$.

It is clear that $[f,g]_{\mathrm{set}}(p)\subseteq[f,g]_C(p)$, but it is easy to see that the inequality can be strict since, for example, if $M=\mathbb{R}$ and $f(x)=g(x)=1+|x|$, then $[f,f]_C(0)=[-1,1]$, while $[f,f]_{\mathrm{set}}(0)=\{0\}$. This example also gives us a good reason for not using $[\cdot,\cdot]_C$ as the set-valued bracket, since it is obviously desirable for a bracket to satisfy the identity $[f,f]=\{0\}$, but we have shown that this identity is not true for $[\cdot,\cdot]_C$.

In the general case when M is a manifold, then (f,g) is a section of the bundle $TM^{(2)}$ whose fiber $T_qM^{(2)}$ is the product $T_qM\times T_qM$. The Clarke Jacobian $\partial(f,g)(q)$ is a compact convex subset of $\mathcal{L}^0(T_qM, T_{(f(q),g(q))}TM^{(2)})$ where, if v,w belong to T_qM, we use $\mathcal{L}^0(T_qM, T_{(v,w)}TM^{(2)})$ to denote the set of all linear maps $L\in\mathcal{L}(T_qM, T_{(v,w)}TM^{(2)})$ such that $d\pi_M^{(2)}\circ L=\mathrm{id}_{T_qM}$, and $\pi_M^{(2)}$ is the canonical projection from $TM^{(2)}$ to M. Then $q[f,g]_{\mathrm{set}}$ is the set $\{L(g(q),-f(q))\colon\ L\in\partial(f,g)(q)\}$.

4. Asymptotic formulae for $qe^{tf}e^{sg}e^{-tf}e^{-sg}$

It is well known—and proved above, cf. (7)—that

$$qe^{tf}e^{tg}e^{-tf}e^{-tg} = q + t^2q[f,g] + o\bigl(t^2\bigr)\quad\text{as } t\to 0, \tag{14}$$

if f and g are vector fields of class C^1 on a manifold M of class C^2. (The precise meaning of this is that

$$qe^{tf}e^{tg}e^{-tf}e^{-tg}\varphi = q\varphi + t^2\bigl(q[f,g]\varphi\bigr) + o\bigl(t^2\bigr)\quad\text{as } t\to 0$$

whenever $\varphi\in\mathcal{E}^1(M)$.)

The goal of this section is to prove more general asymptotic formulae, valid for Lipschitz vector fields, or for continuous vector fields that are "semidifferentiable" at one point q. The result for Lipschitz vector fields is similar to (14), except that the classical Lie bracket in the right-hand side is replaced by $q[f,g]_{\mathrm{set}}$, and the resulting equation has to be properly reinterpreted. If f and g are both classically differentiable at q and Lipschitz near q then the result for semidifferentiable vector fields yields stronger information than the Lipschitz result, as shown in Remark 4.8 below.

4.1. An exact formula for $qe^{tf}e^{sg}e^{-tf}e^{-sg}$ when f and g are of class C^1

We first obtain an exact formula for the commutator $qe^{tf}e^{sg}e^{-tf}e^{-sg}$ when f and g are vector fields of class C^1 on a manifold M of class C^2. Formally, both the statement and the proof of the formula are identical to the ones in [7], where the case when M is a Euclidean space is treated. We give the proof for completeness, and because the argument is quite short and constitutes a good example on how the AGF facilitates computations.

For each $r\in\mathbb{R}$, we use I_r to denote the compact interval $[\min(0,r),\max(0,r)]$. For each ordered pair (t,s) of real numbers, the rectangle $R(t,s)$ is defined by $R(t,s)=I_t\times I_s$.

Lemma 4.1. *Let f and g be vector fields of class C^1 on a manifold M of class C^2. Then, for all $q \in M$, $t, s \in \mathbb{R}$ such that $qe^{\tau f}e^{sg}e^{-\tau f}e^{-sg}$ is defined whenever $\tau \in I_t$, the identity*[11]

$$qe^{tf}e^{sg}e^{-tf}e^{-sg} - q = \int_0^t \int_0^s \left(qe^{\tau f}e^{(s-\sigma)g}[f,g]e^{\sigma g}e^{-\tau f}e^{-sg}\right) d\tau\, d\sigma \tag{15}$$

holds.

Remark 4.2. Under the regularity hypotheses of Lemma 4.1, the vector field $[f, g]$ is continuous, so the integrand function

$$R(t,s) \ni (\tau,\sigma) \mapsto qe^{\tau f}e^{(s-\sigma)g}[f,g]e^{\sigma g}e^{-\tau f}e^{-sg} \in TM \subseteq \mathcal{E}'^1(M)$$

is continuous. Furthermore, this function is equal to $q[f, g] + o(1)$ as $(s, t) \to (0, 0)$. Therefore (15) implies the usual second-order estimate

$$qe^{tf}e^{sg}e^{-tf}e^{-sg} - q = st\big(q[f,g]\big) + o\big(|st|\big). \tag{16}$$

Proof of Lemma 4.1. As in [7], the proof of (15) reduces to the following chain of equalities:

$$\begin{aligned}
qe^{tf}e^{sg}e^{-tf}e^{-sg} - q &= \int_0^t \left(qe^{\tau f}fe^{sg}e^{-\tau f}e^{-sg} - qe^{\tau f}e^{sg}fe^{-\tau f}e^{-sg}\right) d\tau \\
&= \int_0^t qe^{\tau f}e^{sg}\left(e^{-sg}fe^{sg} - f\right)e^{-\tau f}e^{-sg}\, d\tau \\
&= \int_0^t \int_0^s qe^{\tau f}e^{sg}\left(e^{-\sigma g}[f,g]e^{\sigma g}\right)e^{-\tau f}e^{-sg}\, d\sigma\, d\tau \\
&= \int_0^t \int_0^s qe^{\tau f}e^{(s-\sigma)g}[f,g]e^{\sigma g}e^{-\tau f}e^{-sg}\, d\sigma\, d\tau,
\end{aligned}$$

where we have used the identities:

$$\frac{d}{d\tau}\left(qe^{\tau f}e^{sg}e^{-\tau f}e^{-sg}\right) = qe^{\tau f}fe^{sg}e^{-\tau f}e^{-sg} - qe^{\tau f}e^{sg}fe^{-\tau f}e^{-sg} \quad \text{and}$$

$$\frac{d}{d\sigma}\left(y\left(e^{-\sigma g}fe^{\sigma g} - f\right)\right) = ye^{-\sigma g}[f,g]e^{\sigma g}. \qquad \square$$

[11] The meaning of this identity is clear if $M = \mathbb{R}^n$, but the formula is also valid on a more general manifold, if regarded as an equality of members of $\mathcal{E}'^1(M)$.

4.2. Regularizations

Regularizations of vector fields on $\mathbb{R}^n$ are obtained by means of a standard mollification procedure.

We fix, once and for all, a nonnegative real-valued function φ on $\mathbb{R}^n$, such that $\varphi \in C^\infty$, $\int_{\mathbb{R}^n} \varphi(x)\,dx = 1$ and $\varphi(x) = 0$ whenever $\|x\| > 1$,

If Ω is an open subset of $\mathbb{R}^n$ and $\rho > 0$, then Ω_ρ will denote the open set $\{x \in \mathbb{R}^n\colon\ x + \rho\overline{\mathbb{B}}^n \subseteq \Omega\}$.

Definition 4.3. For any continuous vector field k on an open subset Ω of $\mathbb{R}^n$ and any $\rho > 0$, the *ρ-regularization* of k is the vector field k^ρ on Ω_ρ obtained by setting, for every $x \in \Omega_\rho$,

$$k^\rho(x) = \int_{\mathbb{R}^n} \varphi(h) k(x + \rho h)\,dh. \tag{17}$$

It is the clear that k^ρ is a vector field of class C^∞ on Ω_ρ. It will be important for us to have an explicit expression for the differential Dk^ρ of k^ρ when k is locally Lipschitz. The formula we need is given by the following well-known result.

Proposition 4.4. *If n, Ω, k, ρ are as above, and k is locally Lipschitz, then*

$$Dk^\rho(x) = \int_{\overline{\mathbb{B}}^n} \varphi(h) Dk(x + \rho h)\,dh \quad \textit{for all } x \in \Omega_\rho.$$

4.3. A technical lemma

We are going to assume that

(A1) Ω is an open subset of $\mathbb{R}^n$;
(A2) f and g are bounded Lipschitz vector fields on Ω;
(A3) $\mathcal{F}$ is a full subset of $DIFF(f) \cap DIFF(g)$.

We choose a positive constant C such that

$$\max\big(\|f(x)\|, \|g(x)\|, \|Df(x)\|, \|Dg(x)\|\big) \leqslant C \quad \text{for a.e. } x \in \Omega.$$

For each subset S of Ω, we let $[f, g]_{\mathrm{set},S}$ denote the closed convex hull of the set of all vectors $[f, g](x)$, for all $x \in S \cap DIFF(f) \cap DIFF(g)$. Then $[f, g]_{\mathrm{set},S}$ is a convex compact subset of $\mathbb{R}^n$. Furthermore, $[f, g]_{\mathrm{set},S}$ is clearly nonempty if the set $S \cap DIFF(f) \cap DIFF(g)$ is nonempty.

Given a compact convex subset V of $\mathbb{R}^n$, and a nonnegative real number λ, we write

$$V^{(\lambda)} = \big\{v \in \mathbb{R}^n\colon\ \mathrm{dist}(v, V) \leqslant \lambda\big\},$$

$$\check{V}^{((\lambda))} = \big\{v \in \mathbb{R}^n\colon\ \|v - w\| \leqslant \lambda\|w\| \text{ for some } w \in V\big\},$$

so $V^{(\lambda)}$ is compact convex and $\check{V}^{((\lambda))}$ is compact. We let $V^{((\lambda))}$ be the convex hull of $\check{V}^{((\lambda))}$, so $V^{((\lambda))}$ is compact and convex.

Lemma 4.5. *Assume that* (A1)–(A3) *hold. Let $q \in \Omega$, and let t, s be nonzero real numbers having the property that $qe^{\tau f}e^{sg}e^{-\tau f}e^{-sg}$ is defined whenever $\tau \in I_t$. Let $\nu = 2C(|s| + |t|)e^{2C(|s|+|t|)}$. Then*

$$\frac{qe^{tf}e^{sg}e^{-tf}e^{-sg} - q}{ts} \in \big([f, g]_{\mathrm{set},\mathcal{F}}\big)^{((\nu))}. \tag{18}$$

In particular, if $x[f, g] = 0$ for every $x \in \mathcal{F}$, then $qe^{tf}e^{sg}e^{-tf}e^{-sg} = q$.

Proof. Let K be a compact subset of Ω whose interior U contains (i) all the points $qe^{\tau f}e^{\sigma g}$, for $\tau \in I_t$, and $\sigma \in I_s$, as well as (ii) all the points $qe^{\tau f}e^{sg}e^{-\tilde{\tau} f}$, for $\tau \in I_t$, $\tilde{\tau} \in I_\tau$, and (iii) all the $qe^{\tau f}e^{sg}e^{-\tau f}e^{-\tilde{\sigma} g}$, for $\tau \in I_t$, $\tilde{\sigma} \in I_s$. (Such a set exists because the set of points of the three types listed above is a compact subset of Ω.) Choose a positive $\bar{\rho}$ such that $K \subseteq \Omega_{\bar{\rho}}$. Let $K^{\bar{\rho}} = \bigcup_{x \in K}(x + \bar{\rho}\overline{\mathbb{B}}^n)$, so $K^{\bar{\rho}}$ is a compact subset of Ω.

Then, if $x \in K$ and ρ is such that $0 < \rho \leqslant \bar{\rho}$, we have

$$\big[f^\rho, g^\rho\big](x) = Dg^\rho(x) \cdot f^\rho(x) - Df^\rho(x) \cdot g^\rho(x).$$

Furthermore, if $k = f$ or $k = g$, then Proposition 4.4 tells us that

$$Dk^\rho(x) = \int\limits_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h) Dk(x + \rho h)\, dh,$$

where $\mathcal{F}^{\rho,x} = \{h \in \mathbb{R}^n\colon x + \rho h \in \mathcal{F}\}$. Therefore

$$\begin{aligned} Dg^\rho(x) \cdot f^\rho(x) &= \int\limits_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h) Dg(x + \rho h) \cdot f^\rho(x)\, dh \\ &= \int\limits_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h) Dg(x + \rho h) \cdot f(x + \rho h)\, dh + E_1(\rho, x), \end{aligned}$$

where

$$E_1(\rho, x) = \int\limits_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h) Dg(x + \rho h) \cdot \big(f^\rho(x) - f(x + \rho h)\big)\, dh.$$

Similarly,

$$Df^\rho(x) \cdot g^\rho(x) = \int\limits_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h) Df(x + \rho h) \cdot g(x + \rho h)\, dh + E_2(\rho, x),$$

where

$$E_2(\rho, x) = \int\limits_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h) Dg(x + \rho h) \cdot \big(g^\rho(x) - g(x + \rho h)\big)\, dh.$$

Then

$$\left[f^\rho, g^\rho\right](x) = \int\limits_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h)[f,g](x+\rho h)\,dh + E_1(\rho,x) + E_2(\rho,x). \tag{19}$$

Now, if $h \in \overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}$ then $x + \rho h \in \mathcal{F}$, so the vector $[f,g](x+\rho h)$ belongs to $[f,g]_{\mathrm{set},\mathcal{F}}$. It then follows (since $\varphi(h) \geqslant 0$ for all h and $\int_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h)\,dh = 1$) that

$$\int\limits_{\overline{\mathbb{B}}^n \cap \mathcal{F}^{\rho,x}} \varphi(h)[f,g](x+\rho h)\,dh \in [f,g]_{\mathrm{set},\mathcal{F}}.$$

Now

$$f^\rho(x) = \int\limits_{\overline{\mathbb{B}}^n} \varphi(u) f(x+\rho u)\,du \quad \text{and} \quad f(x+\rho h) = \int\limits_{\overline{\mathbb{B}}^n} \varphi(u) f(x+\rho h)\,du,$$

so

$$f^\rho(x) - f(x+\rho h) = \int\limits_{\overline{\mathbb{B}}^n} \varphi(u)\big(f(x+\rho u) - f(x+\rho h)\big)\,du,$$

so $\|E_1(\rho,x)\| \leqslant 2C\rho$. A similar argument shows that $\|E_2(\rho,x)\| \leqslant 2C\rho$. Therefore $[f^\rho, g^\rho](x)$ belongs to $([f,g]_{\mathrm{set},\mathcal{F}})^{(4C\rho)}$.

We now apply (15), with the open set U in the role of the manifold M of Lemma 4.1. The fact that U contains the points $qe^{\tau f}e^{\sigma g}$, for $\tau \in I_t$, and $\sigma \in I_s$, as well as the $qe^{\tau f}e^{sg}e^{-\tilde{\tau} f}$, for $\tau \in I_t$, $\tilde{\tau} \in I_\tau$, and the $qe^{\tau f}e^{sg}e^{-\tau f}e^{-\tilde{\sigma} g}$, for $\tau \in I_t$, $\tilde{\sigma} \in I_s$, implies that there exists a ρ^* such that $0 < \rho^* \leqslant \bar{\rho}$ having the property that U also contains the points of a similar form with f^ρ, g^ρ instead of f, g, for all $\rho \in [0,\rho^*]$. This implies, if $f^{\rho,U}$ $g^{\rho,U}$ denote the restrictions to U of the vector fields f^ρ, g^ρ, that $qe^{\tau f^{\rho,U}}e^{sg^{\rho,U}}e^{-\tau f^{\rho,U}}e^{-sg^{\rho,U}}$ is defined whenever $\tau \in I_t$. Then, if $0 < \rho \leqslant \rho^*$, we have

$$qe^{tf^\rho}e^{sg^\rho}e^{-tf^\rho}e^{-sg^\rho} - q = \int\limits_0^t\int\limits_0^s \left(qe^{\tau f^\rho}e^{(s-\sigma)g^\rho}\left[f^\rho, g^\rho\right]e^{\sigma g^\rho}e^{-\tau f^\rho}e^{-sg^\rho}\right)d\tau\,d\sigma. \tag{20}$$

For any fixed $(\tau,\sigma) \in R(t,s)$, let $x = qe^{\tau f^\rho}e^{(s-\sigma)g^\rho}$, $y = qe^{\tau f^\rho}e^{sg^\rho}e^{-\tau f^\rho}e^{-sg^\rho}$, $v_0 = qe^{\tau f^\rho}e^{(s-\sigma)g^\rho}[f^\rho, g^\rho]$, $v_1 = v_0 e^{\sigma g^\rho}$, $\mu_1 = |\sigma|$, $v_2 = v_1 e^{-\tau f^\rho}$, $\mu_2 = |\tau|$, $v_3 = v_2 e^{-sg^\rho}$, $\mu_3 = |t|$, $\mu = \mu_1 + \mu_2 + \mu_3$. Then v_3 is computed by solving a differential equation $\dot{V}(u) = M(u)\cdot V(u)$ with initial condition $V(0) = v_0$ on the interval $[0,\mu]$, where M is a matrix-valued function such that $\|M(u)\| \leqslant C$ for all u. Gronwall's inequality then implies that $\|V(u)\| \leqslant e^{C\mu}\|v_0\|$ for all u, so that $\|v_3 - v_0\| \leqslant C\mu e^{C\mu}\|v_0\| \leqslant 2C(|s|+|t|)e^{2C(|s|+|t|)}\|v_0\|$. Since v_0 belongs to the set $([f,g]_{\mathrm{set},\mathcal{F}\cap K})^{(4C\rho)}$, we conclude that $v_3 \in W(\rho)$, where

$$W(\rho) = \left(\left([f,g]_{\mathrm{set},\mathcal{F}}\right)^{(4C\rho)}\right)^{((2C(|s|+|t|)e^{2C(|s|+|t|)})}.$$

So the integrand of (20) belongs to $W(\rho)$ for each τ, σ. Since $W(\rho)$ is compact and convex, we conclude that

$$\frac{qe^{tf^\rho}e^{sg^\rho}e^{-tf^\rho}e^{-sg^\rho}-q}{ts} \in W(\rho). \tag{21}$$

If we now let $\rho \downarrow 0$, we find that (18) holds.

The last assertion is trivial, since the hypothesis that $x[f,g]=0$ for every $x \in \mathcal{F}$ implies that $([f,g]_{\mathrm{set},\mathcal{F}})^{((\nu))} = \{0\}$. □

4.4. An asymptotic formula for Lipschitz vector fields

We now prove a result stating that, asymptotically as $(t,s) \to (0,0)$, $t \neq 0$, $s \neq 0$, the difference $qe^{tf}e^{sg}e^{-tf}e^{-sg}-q$ "behaves like $ts(q[f,g]_{\mathrm{set}})+o(|ts|)$." The precise meaning of this, if $q[f,g]_{\mathrm{set}}$ is the singleton of a vector v, is that

$$\lim_{(t,s)\to(0,0),\, t\neq 0,\, s\neq 0} \frac{qe^{tf}e^{sg}e^{-tf}e^{-sg}-q}{ts} = v.$$

In the more general case when $q[f,g]_{\mathrm{set}}$ is a set, the conclusion is as follows.

Proposition 4.6. *Assume that M is an m-dimensional manifold of class C^2, f and g are locally Lipschitz vector fields on M, and $q \in M$. Then, if $\kappa : \Omega \mapsto \mathbb{R}^m$ is any coordinate chart of M defined on a neighborhood Ω of q, the identity*

$$\lim_{(t,s)\to(0,0),\, t\neq 0,\, s\neq 0} \operatorname{dist}\left(\frac{\kappa(qe^{tf}e^{sg}e^{-tf}e^{-sg})-\kappa(q)}{ts}, D\kappa\cdot\big(q[f,g]_{\mathrm{set}}\big)\right) = 0 \tag{22}$$

holds, where $D\kappa \cdot (q[f,g]_{\mathrm{set}})$ is the subset of $\mathbb{R}^m$ which is the image under the differential of κ of the subset $q[f,g]_{\mathrm{set}}$ of T_qM.

Proof. It suffices to apply Lemma 4.5. First, observe that since our conclusion is local we may assume that M is an open subset of $\mathbb{R}^n$ and κ is the identity map. Fix a positive number $\bar{\alpha}$ such that $q+\bar{\alpha}\overline{\mathbb{B}}^n \subseteq M$, and then let N be a number which is both an upper bound for $\|f(x)\|$ and $\|g(x)\|$ for all $x \in q+\bar{\alpha}\overline{\mathbb{B}}^n$ and a Lipschitz constant for f and g on $q+\bar{\alpha}\overline{\mathbb{B}}^n$. Then, if $0<\alpha\leqslant\bar{\alpha}$, $f_1,\dots,f_k$ is an arbitrary finite sequence such that each f_j is either f or g, and $t_1,\dots,t_k$ are real numbers such that $|t_1|+\cdots+|t_k| \leqslant \frac{\alpha}{N}$, if follows that $qe^{t_1f_1}e^{t_2f_2}\cdots e^{t_kf_k}$ is defined and belongs to $q+\alpha\overline{\mathbb{B}}^n$. Lemma 4.5 then implies that $(ts)^{-1}(qe^{tf}e^{sg}e^{-tf}e^{-sg}-q)$ belongs to the set $([f,g]_{\mathrm{set},q+\alpha\overline{\mathbb{B}}^n})^{((\nu(s,t)))}$ whenever

$$t \neq 0, \qquad s \neq 0 \quad \text{and} \quad 2N\big(|t|+|s|\big) \leqslant \alpha, \tag{23}$$

where $\nu(s,t) = 2N(|s|+|t|)e^{2N(|s|+|t|)}$. It follows that

$$\frac{qe^{tf}e^{sg}e^{-tf}e^{-sg}-q}{ts} \in \big([f,g]_{\mathrm{set},q+\alpha\overline{\mathbb{B}}^n}\big)^{((\alpha e^\alpha))} \quad \text{whenever (23) holds.}$$

It is clear that

$$\bigcap_{\alpha>0}\left([f,g]_{\mathrm{set},q+\alpha\bar{\mathbb{B}}^n}\right)^{((\alpha e^{\alpha}))}=q[f,g]_{\mathrm{set}}.$$

Therefore, given any positive ε we can find α such that $([f,g]_{\mathrm{set},q+\alpha\bar{\mathbb{B}}^n})^{((\alpha e^{\alpha}))}$ is a subset of the ε-neighborhood of $q[f,g]_{\mathrm{set}}$. Then

$$\operatorname{dist}\left(\frac{qe^{tf}e^{sg}e^{-tf}e^{-sg}-q}{ts},q[f,g]_{\mathrm{set}}\right)\leqslant\varepsilon\quad\text{whenever (23) holds.}$$

Therefore (22) holds, and our proof is complete. □

4.5. An asymptotic formula for semidifferentiable vector fields

A continuous vector field f on a manifold M of class C^2 is said to be *semidifferentiable* at a point $q\in M$ if there exists a locally Lipschitz vector field F on M such that

$$\lim_{x\to q}\frac{f(x)-F(x)}{\|x-q\|}=0.$$

(This formula has a clear meaning relative to a particular coordinate chart, and it is easily proved that if it is valid in some chart then it is valid in every chart. Alternatively, it is not hard to give an intrinsic interpretation.)

It is clear that any vector field f which is Lipschitz on some neighborhood of q is semidifferentiable at q, since we can take $F=f$. Also, a continuous vector field f which is classically differentiable at q is semidifferentiable at q. For an example of a vector field which is neither Lipschitz nor classically differentiable at a point q but is semidifferentiable at q, take $M=\mathbb{R}$, $q=0$, $f(x)=\varphi(x)\partial_x$, where $\varphi(x)=|x|+|x|^{3/2}\sin 1/x$ if $x\neq 0$, and $\varphi(0)=0$.

If we are given two nonempty subsets A, B of a metric space X, we define the *quasidistance* $\Delta(A,B)$ from A to B by the formula

$$\Delta(A,B)=\sup\bigl\{\operatorname{dist}(a,B)\colon a\in A\bigr\}. \tag{24}$$

(This function is closely related to, but not the same as, the Hausdorff distance $\Delta_{\mathrm{Ha}}(A,B)$ between A and B. The precise relation between the two functions is that $\Delta_{\mathrm{Ha}}(A,B)=\max\{\Delta(A,B),\Delta(B,A)\}$.)

In the proposition stated below, if f and g are continuous vector fields then $qe^{tf}e^{tg}e^{-tf}e^{-tg}$ is a set, since the vector fields may fail to have unique trajectories. Precisely, $qe^{tf}e^{tg}e^{-tf}e^{-tg}$ is the set of all z such that there exist u,v,w for which $u\in qe^{tf}$, $v\in ue^{tg}$, $w\in ve^{-tf}$, and $w\in ve^{-tg}$. (If $k=f$ or $k=g$, then qe^{tk} is the set of all points of the form $\xi(t)$, where ξ is an integral curve of k such that $\xi(0)=q$.)

Proposition 4.7. *Let f, g be continuous vector fields on a manifold M of class C^2, and let $q\in M$ be such that f and g are semidifferentiable at q. Let F,G be locally Lipschitz vector*

fields on M such that $\lim_{x\to q}\frac{f(x)-F(x)}{\|x-q\|}=0$ *and* $\lim_{x\to q}\frac{g(x)-G(x)}{\|x-q\|}=0$. *Then, if* $\kappa:\Omega\mapsto\mathbb{R}^m$ *is any coordinate chart of M defined on a neighborhood Ω of q, the identity*

$$\lim_{t\to 0}\Delta\left(\frac{\kappa(q\,e^{tf}e^{tg}e^{-tf}e^{-tg})-\kappa(q)}{t^2},\,D\kappa\cdot\big(q[F,G]_{\mathrm{set}}\big)\right)=0 \tag{25}$$

holds, where $D\kappa\cdot(q[F,G]_{\mathrm{set}})$ is the subset of $\mathbb{R}^m$ which is the image under the differential of κ of the subset $q[F,G]_{\mathrm{set}}$ of T_qM.

Proof. The conclusion is clearly local, so we assume, without loss of generality, that $M=\mathbb{R}^n$, κ is the identity map, $q=0$, f and g are continuous globally bounded maps from $\mathbb{R}^n$ to $\mathbb{R}^n$, F and G are globally Lipschitz globally bounded maps from $\mathbb{R}^n$ to $\mathbb{R}^n$, and $\theta:]0,+\infty[\mapsto[0,+\infty]$ is an increasing function such that $\lim_{r\to 0}\theta(r)=0$ and the inequalities $\|f(x)-F(x)\|\leqslant\theta(\|x\|)\|x\|$, $\|g(x)-G(x)\|\leqslant\theta(\|x\|)\|x\|$, hold for all $x\in\mathbb{R}^n$.

We fix a positive number C which is a global Lipschitz constant for F and G and a global upper bound for f, g, F, G.

If $x,\tilde{x}\in\mathbb{R}^n$ and $t\in\mathbb{R}$, then the expressions $\|xe^{tf}-\tilde{x}e^{tF}\|$, $\|xe^{tg}-\tilde{x}e^{tG}\|$ will denote, respectively, the supremum of the set $\{\|y-\tilde{x}e^{tF}\|:\ y\in xe^{tf}\}$, and the supremum of the set $\{\|y-\tilde{x}e^{tG}\|:\ y\in xe^{tg}\}$.

Next, fix $x,\tilde{x},y\in\mathbb{R}^n$ and $t\in\mathbb{R}\setminus\{0\}$, and assume that $y\in xe^{tf}$. Pick absolutely continuous maps $\xi,\Xi:I_t\mapsto\mathbb{R}^n$ such that $\dot{\xi}(s)=f(\xi(s))$ and $\dot{\Xi}(s)=F(\Xi(s))$ for almost all $s\in I_t$, $\xi(0)=x$, $\xi(t)=y$, and $\Xi(0)=\tilde{x}$. Then, if $s\in I_t$, we have

$$\begin{aligned}\xi(s)-\Xi(s)&=x-\tilde{x}+\sigma_t\int_{I_s}\big(f\big(\xi(r)\big)-F\big(\Xi(r)\big)\big)\,dr\\&=x-\tilde{x}+\sigma_t\int_{I_s}\big(f\big(\xi(r)\big)-F\big(\xi(r)\big)\big)\,dr+\sigma_t\int_{I_t}\big(F\big(\xi(r)\big)-F\big(\Xi(r)\big)\big)\,dr,\end{aligned}$$

where $\sigma_t=1$ if $t>0$ and $\sigma_t=-1$ if $t<0$. Then

$$\big\|\xi(s)-\Xi(s)\big\|\leqslant\|x-\tilde{x}\|+|s|\Theta_{f,F,\xi}(s)+C\int_{I_s}\big\|\xi(r)-\Xi(r)\big\|\,dr,$$

where

$$\Theta_{f,F,\xi}(s)=\sup\big\{\big\|f\big(\xi(r)\big)-F\big(\xi(r)\big)\big\|:\ r\in I_s\big\}.$$

Then Gronwall's inequality implies that

$$\big\|\xi(s)-\Xi(s)\big\|\leqslant e^{C|s|}\big(\|x-\tilde{x}\|+|s|\Theta_{f,F,\xi}(s)\big).$$

On the other hand, if we let

$$\|\xi\|_{\sup,s}=\sup\big\{\big\|\xi(r)\big\|:\ r\in I_s\big\},$$

then

$$\left\| f\big(\xi(r)\big) - F\big(\xi(r)\big) \right\| \leqslant \theta\big(\|\xi\|_{\sup,s}\big)\|\xi\|_{\sup,s} \quad \text{for all } r \in I_s,$$

so

$$\Theta_{f,F,\xi}(s) \leqslant \theta\big(\|\xi\|_{\sup,s}\big)\|\xi\|_{\sup,s}$$

and then

$$\left\|\xi(s) - \Xi(s)\right\| \leqslant e^{C|s|}\big(\|x-\tilde{x}\| + |s|\theta\big(\|\xi\|_{\sup,s}\big)\|\xi\|_{\sup,s}\big).$$

Therefore

$$\left\| F\big(\xi(s)\big) - F\big(\Xi(s)\big) \right\| \leqslant Ce^{C|s|}\big(\|x-\tilde{x}\| + |s|\theta\big(\|\xi\|_{\sup,s}\big)\|\xi\|_{\sup,s}\big),$$

so that, if we let $K = 1 + Ce^C$, we have

$$\left\| f\big(\xi(s)\big) - F\big(\Xi(s)\big) \right\| \leqslant K\big(\|x-\tilde{x}\| + \theta\big(\|\xi\|_{\sup,s}\big)\|\xi\|_{\sup,s}\big) \quad \text{if } |s| \leqslant 1.$$

Then, if $|s| \leqslant 1$, we find that

$$\begin{aligned}\left\|\xi(s) - \Xi(s) - (x-\tilde{x})\right\| &= \left\|\sigma_t \int_{I_s} \big(f\big(\xi(r)\big) - F\big(\Xi(r)\big)\big)\,dr\right\| \\ &\leqslant K|s|\big(\|x-\tilde{x}\| + \theta\big(\|\xi\|_{\sup,s}\big)\|\xi\|_{\sup,s}\big).\end{aligned}$$

On the other hand, $\|\xi(r)\| \leqslant \|x\| + C|r|$ for every r, so $\|\xi\|_{\sup,s} \leqslant \|x\| + C|s|$. It follows that

$$\left\|\xi(s) - \Xi(s) - (x-\tilde{x})\right\| \leqslant K|s|\big(\|x-\tilde{x}\| + \theta\big(\|x\| + C|s|\big)\big(\|x\| + C|s|\big)\big).$$

If $|t| \leqslant 1$, then we can let $s = t$, in which case $\xi(s) = y$, and we get

$$\left\| y - \tilde{x}e^{tF} - (x-\tilde{x}) \right\| \leqslant K|t|\big(\|x-\tilde{x}\| + \theta\big(\|x\| + C|t|\big)\big(\|x\| + C|t|\big)\big).$$

If we let $\sigma(t,\alpha) = (1+C)K\theta(\alpha + C|t|)$, we see that

- *if $|t| \leqslant 1$ and $y \in xe^{tf}$, then*

$$\left\| y - \tilde{x}e^{tF} - (x-\tilde{x}) \right\| \leqslant K|t|\|x-\tilde{x}\| + \sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big); \tag{26}$$

- *$\sigma(t,\alpha)$ goes to zero as $(t,\alpha) \to (0,0)$.*

Naturally, the same conclusion is true for g and G, so that

- *if $|t| \leqslant 1$ and $y \in xe^{tg}$, then*

$$\left\| y - \tilde{x}e^{tG} - (x-\tilde{x}) \right\| \leqslant K|t|\|x-\tilde{x}\| + \sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big). \tag{27}$$

We now estimate the quantity

$$\begin{aligned} Q(t,x,\tilde{x}) &\stackrel{\text{def}}{=} \left\| x e^{tf} e^{tg} e^{-tf} e^{-tg} - \tilde{x} e^{tF} e^{tG} e^{-tF} e^{-tG} - (x-\tilde{x}) \right\| \\ &\stackrel{\text{def}}{=} \sup\left\{ \left\| y - \tilde{x} e^{tF} e^{tG} e^{-tF} e^{-tG} - (x-\tilde{x}) \right\| \colon\ y \in x e^{tf} e^{tg} e^{-tf} e^{-tg} \right\}, \end{aligned}$$

assuming that $|t| \leqslant 1$. For this purpose, pick a member y of $x e^{tf} e^{tg} e^{-tf} e^{-tg}$, and let $\tilde{y} = \tilde{x} e^{tF} e^{tG} e^{-tF} e^{-tG}$. Let y_1, y_2, y_3 be such that $y_1 \in x e^{tf}$, $y_2 \in y_1 e^{tg}$, $y_3 \in y_2 e^{-tf}$, and $y \in y_3 e^{-tg}$. Let $\tilde{y}_1 = \tilde{x} e^{tF}$, $\tilde{y}_2 = \tilde{y}_1 e^{tG}$, $\tilde{y}_3 = \tilde{y}_2 e^{-tf}$, $\tilde{y} = \tilde{y}_3 e^{-tG}$.

Then (26) and (27) imply the estimates

$$\left\| y_1 - \tilde{y}_1 - (x - \tilde{x}) \right\| \leqslant K|t| \|x - \tilde{x}\| + \sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big), \tag{28}$$

$$\left\| y_2 - \tilde{y}_2 - (y_1 - \tilde{y}_1) \right\| \leqslant K|t| \|y_1 - \tilde{y}_1\| + \sigma\big(t, \|y_1\|\big)|t|\big(\|y_1\| + |t|\big), \tag{29}$$

$$\left\| y_3 - \tilde{y}_3 - (y_2 - \tilde{y}_2) \right\| \leqslant K|t| \|y_2 - \tilde{y}_2\| + \sigma\big(t, \|y_2\|\big)|t|\big(\|y_2\| + |t|\big), \tag{30}$$

$$\left\| y - \tilde{y} - (y_3 - \tilde{y}_3) \right\| \leqslant K|t| \|y_3 - \tilde{y}_3\| + \sigma\big(t, \|y_3\|\big)|t|\big(\|y_3\| + |t|\big). \tag{31}$$

If follows from (28) that

$$\|y_1 - \tilde{y}_1\| \leqslant \big(1 + K|t|\big)\|x - \tilde{x}\| + \sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big)$$

and, in addition, it is clear that $\|y_1\| \leqslant \|x\| + C|t|$. Then (28) and (29) imply

$$\begin{aligned} \left\| y_2 - \tilde{y}_2 - (x - \tilde{x}) \right\| &\leqslant \left\| y_2 - \tilde{y}_2 - (y_1 - \tilde{y}_1) \right\| + \left\| y_1 - \tilde{y}_1 - (x - \tilde{x}) \right\| \\ &\leqslant K|t| \|y_1 - \tilde{y}_1\| + \sigma\big(t, \|y_1\|\big)|t|\big(\|y_1\| + |t|\big) \\ &\quad + K|t| \|x - \tilde{x}\| + \sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big) \\ &\leqslant K|t|\Big(\big(1 + K|t|\big)\|x - \tilde{x}\| + \sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big)\Big) \\ &\quad + \sigma\big(t, \|x\| + C|t|\big)|t|\big(\|x\| + C|t| + |t|\big) \\ &\quad + K|t| \|x - \tilde{x}\| + \sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big) \\ &\leqslant \mathcal{K}(t)\|x - \tilde{x}\| + K|t|\sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big) \\ &\quad + \sigma\big(t, \|x\| + C|t|\big)|t|\big(\|x\| + C|t| + |t|\big) + \sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big), \quad \text{so} \end{aligned}$$

$$\left\| y_2 - \tilde{y}_2 - (x - \tilde{x}) \right\| \leqslant \mathcal{K}(t)\|x - \tilde{x}\| + \Sigma\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big), \tag{32}$$

where

$$\mathcal{K}(t) = 2K|t| + K^2 t^2, \qquad \Sigma(t,\alpha) = \big(1 + K|t|\big)\sigma(t,\alpha) + (1 + C)\sigma\big(t, \alpha + C|t|\big).$$

A similar calculation yields

$$\left\| y - \tilde{y} - (y_2 - \tilde{y}_2) \right\| \leqslant \mathcal{K}(t)\|y_2 - \tilde{y}_2\| + \Sigma\big(t, \|y_2\|\big)|t|\big(\|y_2\| + |t|\big). \tag{33}$$

If we combine (32) and (33), we find, by means of an argument similar to the one used to derive (32) from (28) and (29), that

$$\left\| y - \tilde{y} - (x - \tilde{x}) \right\| \leqslant \mathcal{K}^*(t)\|x - \tilde{x}\| + \Sigma^*\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big), \tag{34}$$

where

$$\mathcal{K}^*(t) = 2\mathcal{K}(t) + \mathcal{K}(t)^2, \qquad \Sigma^*(t,\alpha) = \big(1 + \mathcal{K}(t)\big)\Sigma(t,\alpha) + (1 + 2C)\Sigma\big(t, \alpha + 2C|t|\big).$$

(The factor $2C$ appears because, instead of inequality $\|y_1\| \leqslant \|x\| + C|t|$, we now have $\|y_2\| \leqslant \|x\| + 2C|t|$.)

If we now take $x = \tilde{x}$, we find

$$\|y - \tilde{y}\| \leqslant \Sigma^*\big(t, \|x\|\big)|t|\big(\|x\| + |t|\big). \tag{35}$$

If we specialize further to $x = 0$, we get

$$\|y - \tilde{y}\| \leqslant \Sigma^*(t, 0)t^2. \tag{36}$$

This shows that

$$\lim_{t\to 0} \Delta\left(\frac{q e^{tf} e^{tg} e^{-tf} e^{-tg}}{t^2}, \frac{q e^{tF} e^{tG} e^{-tF} e^{-tG}}{t^2} \right) = 0.$$

On the other hand, we know from Proposition 4.6 that

$$\lim_{t\to 0} \operatorname{dist}\left(\frac{q e^{tF} e^{tG} e^{-tF} e^{-tG} - q}{t^2}, q[F, G]_{\text{set}} \right) = 0.$$

These formulae clearly imply (25), completing our proof. □

4.6. An asymptotic formula for continuous classically differentiable vector fields

An important special case of Proposition 4.7 arises when f and g are continuous vector fields that are classically differentiable at a point. In that case, we can take F and G to be linear vector fields, relative to some chart κ defined near q, and conclude that

$$\lim_{t\to 0} \Delta\left(\frac{\kappa(q e^{tf} e^{tg} e^{-tf} e^{-tg}) - \kappa(q)}{t^2}, D\kappa \cdot \big(q[f, g]\big) \right) = 0, \tag{37}$$

because $D\kappa \cdot (q[F, G]_{\text{set}}) = \{D\kappa \cdot (q[f, g])\}$.

Remark 4.8. The result for semidifferentiable vector fields yields new information even when the vector fields are also Lipschitz. For instance, if $M = \mathbb{R}$ and f, g are defined by letting $f(q) = q^2 \sin(\frac{1}{q})$ $\forall q \in \mathbb{R} \setminus \{0\}$, $f(0) = 0$, and $g(q) = 1$ $\forall q \in \mathbb{R}$, then both f and g are differentiable everywhere and Lipschitz continuous. Notice that for $q = 0$ one has

$$q[f, g] = 0, \qquad q[f, g]_{\text{set}} = [-1, 1].$$

So, on one hand, by applying formula (22) with κ equal to the identity map and $t = s$, we obtain

$$\lim_{t\to 0}\left|\frac{qe^{tf}e^{tg}e^{-tf}e^{-tg}}{t^2}\right| \in [-1,1]. \tag{38}$$

On the other hand, by applying formula (37)—still with κ equal to the identity map—we get the much stronger relation

$$\lim_{t\to 0}\left|\frac{qe^{tf}e^{tg}e^{-tf}e^{-tg}}{t^2}\right| = 0.$$

Remark 4.9. A "natural" generalization of (37) would be the formula

$$\lim_{t\to 0,\, s\to 0,\, t\neq 0,\, s\neq 0} \Delta\left(\frac{\kappa(qe^{tf}e^{sg}e^{-tf}e^{-sg}) - \kappa(q)}{st}, D\kappa\cdot\big(q[f,g]\big)\right) = 0, \tag{39}$$

which, presumably, might be true when f and g are continuous vector fields that are classically differentiable at q.

It turns out, however, that *formula* (39) *is not true in general under these conditions.* One trivial reason for this is that, if (39) was true, it would follow that

$$qe^{tf}e^{sg}e^{-tf}e^{-sg} = q + st\big(q[f,g]\big) = o\big(|st|\big),$$

so in particular we could plug in $s = 0$ and conclude that $qe^{tf}e^{-tf} = \{q\}$. But, if f is just continuous, then f need not have uniqueness of trajectories, and this clearly implies that the set $qe^{tf}e^{-tf}$ need not coincide with $\{q\}$.

Furthermore, *formula* (39) *can fail to be true even when f and g have unique trajectories*, as shown by the following example.

Example 4.10. Let us define the map $\psi:\mathbb{R}\to\mathbb{R}$ by setting $\psi(\rho) = 0$ for all $\rho \leqslant -1$, $\psi(\rho) = e^{\rho^2/(\rho^2-1)}$ for $\rho \in]-1,0[$ and $\psi(\rho) = 1$ for $\rho \geqslant 0$. The map ψ is of class C^∞, and, for every $\rho \in]-1,0[$, one has

$$\frac{d\psi}{d\rho}(\rho) = -\frac{2\rho e^{\frac{\rho^2}{\rho^2-1}}}{(\rho^2-1)^2}.$$

Let us consider the vector fields on $\mathbb{R}^2$

$$f(x,y) = \begin{pmatrix} 1 \\ \varphi(x,y) \end{pmatrix}, \qquad g(x,y) = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

where the function $\varphi:\mathbb{R}^2\to\mathbb{R}$ is defined as follows:

$$\varphi(x,y) = x^2 \quad \text{if } y \geqslant 0, \qquad \varphi(x,y) = 0 \quad \text{if } y \leqslant -x^4,$$

and

$$\varphi(x,y) = x^2\psi\left(\frac{y}{x^4}\right) \quad \text{if } -x^4 < y < 0.$$

The map φ is continuous on $\mathbb{R}^2$ and, since

$$\frac{|\varphi(x,y) - \varphi(0,0)|}{|(x,y)|} \leqslant \frac{x^2}{|(x,y)|},$$

it is differentiable at the origin, with $D\varphi(0,0) = (0,0)$. Actually, it can be easily checked that φ is differentiable at any point $(x,y) \in \{(x,-x^4)\colon x \in \mathbb{R}\}$, and, at any such point one has $D\varphi(0,0) = (0,0)$. Hence, the map φ is differentiable everywhere in $\mathbb{R}^2$. However *it is not of class* C^1. Indeed, for every (x,y) such that $x \neq 0$ and $-x^4 < y < 0$, one has

$$\frac{\partial\varphi}{\partial y}(x,y) = -2x^{-2}\frac{\frac{y}{x^4}e^{\frac{y^2}{y^2-x^8}}}{(\frac{y^2}{x^8}-1)^2} = -\frac{2x^{10}ye^{\frac{y^2}{y^2-x^8}}}{(y^2-x^8)^2}.$$

Hence, since $-x^5 > -x^4$ for every $x \in]0,1[$, one has

$$\lim_{x\to 0+}\frac{\partial\varphi}{\partial y}(x,y)\bigg|_{y=-x^5} = -\infty.$$

Therefore the vector field f is everywhere differentiable and is not of class C^1. Let us observe that, though it is not locally Lipschitz in a neighborhood of the origin, the corresponding Cauchy problem has a unique local solution for every initial point. This is trivial when the initial point is not the origin, for in a small compact neighborhood of such a point the vector field f is C^1, hence Lipschitz. As for the origin, let us observe that the half-plane $\Lambda = \{(x,y) \mid y \geqslant 0\}$ is invariant for the vector field f, that is, every integral curve of f starting in Λ remains in Λ during its interval of existence. Hence such an integral curve is unique, for f is locally Lipschitz on Λ. Actually, for every $(x,y) \in \Lambda$ one has

$$\begin{pmatrix} x \\ y \end{pmatrix} e^{tf} = \begin{pmatrix} t+x \\ y + \frac{(t+x)^3}{3} \end{pmatrix}.$$

Let $M_1 > 0$ be an upper bound for both $|f|$ and $|g|$ on the square $[-1,1]^2$. In particular, setting $\epsilon = \min\{1, 1/M_1\}$ and $q_0 = (0,0)^\dagger$, we can define the (single-valued) map

$$(t,s) \mapsto q_0 e^{tf} e^{sg} e^{-tf} e^{-sg}$$

on the set $]-\epsilon,\epsilon[^2$. Observe that

$$q_0[f,g] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hence, if formula (39) were true, we would have that

$$q_0 e^{tf} e^{sg} e^{-tf} e^{-sg} = o\big(|st|\big),$$

which, in particular, would imply

$$q_0 e^{tf} e^{t^3 g} e^{-tf} e^{-t^3 g} = o\big(|t|^4\big) \tag{40}$$

for every $t \in]-\epsilon^3, \epsilon^3[$. Let us set

$$q_1(t) = q_0 e^{tf}, \qquad q_2(t) = q_0 e^{tf} e^{t^3 g}, \qquad q_3(t) = q_0 e^{tf} e^{t^3 g} e^{-tf}.$$

Then

$$q_1(t) = \begin{pmatrix} t \\ \frac{t^3}{3} \end{pmatrix}, \qquad q_2(t) = q_1(t) e^{t^3 g} = \begin{pmatrix} x_2(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} t \\ -\frac{2t^3}{3} \end{pmatrix}.$$

In particular, if t is sufficiently small, one has

$$y_2(t) = -\frac{2t^3}{3} < -t^4,$$

so that $q_2(t)$ belongs to the set $C = \{(x, y) \mid y < -x^4\}$. Since $f = (1, 0)^\dagger$ on C one has

$$q_3(t) = q_2(t) e^{-tf} = \begin{pmatrix} 0 \\ -\frac{2t^3}{3} \end{pmatrix}$$

which yields

$$q_0 e^{tf} e^{t^3 g} e^{-tf} e^{-t^3 g} = q_3(t) e^{-t^3 g} = \begin{pmatrix} 0 \\ \frac{t^3}{3} \end{pmatrix},$$

so that, in particular, (40) turns out to be false.

5. Commutativity of flows of locally Lipschitz vector fields

For a pair (f, g) of vector fields of class C^1, it is well known that local commutativity of the flows of f and g is equivalent to the vanishing of the Lie bracket $[f, g]$.[12] We now prove the extension of this result to the locally Lipschitz case.

To begin with, we have to be precise about the various ways in which the flows of two vector fields may be said to "commute." (We recall that $R(t, s)$ is defined in Section 4.1.)

Definition 5.1. Let M be a manifold of class C^2, and let f, g be locally Lipschitz vector fields on M. We say that

[12] Incidentally, notice that this is also a trivial consequence of formula (15).

- *the flows of f and g commute* if, for every pair of real numbers s, t, (i) $qe^{sg}e^{tf}$ is defined if and only if $qe^{tf}e^{sg}$ is defined, and (ii) if $qe^{sg}e^{tf}$ is defined then $qe^{sg}e^{tf} = qe^{tf}e^{sg}$;
- *the flows of f and g commute on rectangles* if, whenever $q \in M$, $t, s \in \mathbb{R}$ are such that $qe^{\tau f}e^{\sigma g}$ is defined for all $(\tau, \sigma) \in R(t, s)$, it follows that $qe^{sg}e^{tf}$ is defined and $qe^{sg}e^{tf} = qe^{tf}e^{sg}$;
- *the flows of f and g commute for small times near a point $q_* \in M$* if there exist a neighborhood U of q_* and a positive number ε such that

$$qe^{tf}e^{sg} \overset{\text{bsd}}{=} qe^{sg}e^{tf} \tag{41}$$

for all $q' \in U$, $t, s \in \mathbb{R}$ such that $|t| \leqslant \varepsilon$ and $|s| \leqslant \varepsilon$, where "$A \overset{\text{bsd}}{=} B$" means "$A$ and B are both defined, and in addition they are equal."

Remark 5.2. In the definition of "commuting on rectangles," it is not immediately obvious that the roles of f and g can be interchanged, i.e., that if the flows of f and g commute on rectangles then the flows of g and f commute on rectangles. This is true, however, because Theorem 5.3 below says that both conditions are equivalent to commutativity for small times, which is symmetric with respect to the interchange of f and g.

Theorem 5.3. *Let M be a manifold of class C^2, and let f, g be locally Lipschitz vector fields on M. Then the following conditions are equivalent*:

(i) $q[f, g] = 0$ *for almost every* $q \in M$;[13]
(ii) $q[f, g] = 0$ *for every member* $q \in DIFF(f) \cap DIFF(g)$;
(iii) $q[f, g]_{\text{set}} = \{0\}$ *for every* $q \in M$;
(iv) *the flows of f and g commute for small times near q for every $q \in M$*;
(v) *the flows of f and g commute on rectangles.*

Proof. It is clear that (iii) $\Leftrightarrow$ (ii) $\Rightarrow$ (i) and (v) $\Rightarrow$ (iv). The implication (i) $\Rightarrow$ (iv) is a trivial consequence of Lemma 4.5.

The implication (v) $\Rightarrow$ (ii) is a straightforward consequence of the classical asymptotic formula

$$qe^{tf}e^{tg}e^{-tf}e^{-tg} - q = t^2\big(q[f, g]\big) + o\big(t^2\big) \quad \text{as } t \downarrow 0 \tag{42}$$

which, as we have seen in Section 4.6, holds true at each point $q \in DIFF(f) \cap DIFF(g)$.

To conclude, we have to prove that (iv) $\Rightarrow$ (v).

Lemma 5.4. *Let f and g be locally Lipschitz vector fields on a manifold M of class C^2. Assume that the flows of f and g commute for small times near q for every $q \in M$. Then the flows of f and g commute on rectangles.*

Proof. Let $q \in M$, $t, s \in \mathbb{R}$ be such that $qe^{\tau f}e^{\sigma g}$ is defined whenever $(\tau, \sigma) \in R(t, s)$. We want to prove that $qe^{tf}e^{sg} = qe^{sg}e^{tf}$. It clearly suffices to assume that $t > 0$ and $s > 0$. (If $t < 0$, we

[13] That is, there is a full subset $\mathcal{F}$ of M such that $q[f, g] = 0$ for every member q of the set $DIFF(f) \cap DIFF(g) \cap \mathcal{F}$.

may substitute $|t|$ for t and $-f$ for f; if $s < 0$, we may substitute $|s|$ for s and $-g$ for g.) For $r \geqslant 0$, let $I_r = [0, r]$.

Let $K = \{qe^{\tau f}e^{\sigma g}\colon\ \tau \in I_t,\ \sigma \in I_s\}$. Then K is a compact subset of M, because $qe^{\tau f}e^{\sigma g}$ exists whenever $\tau \in I_t$, $\sigma \in I_s$, and the map $I_t \times I_s \ni (\tau, \sigma) \mapsto qe^{\tau f}e^{\sigma g}$ is continuous. So there exists a positive ε such that $q'e^{\tau f}e^{\sigma g}$ is defined and equal to $q'e^{\sigma g}e^{\tau f}$ whenever $q' \in K$, $|\tau| \leqslant \varepsilon$ and $|\sigma| \leqslant \varepsilon$.

Let N be a positive integer such that $\frac{t}{N} < \varepsilon$ and $\frac{s}{N} < \varepsilon$. Let $t_j = \frac{jt}{N}$, $s_k = \frac{ks}{N}$, for $j = 0, \ldots, N$, $k = 0, \ldots, N$. We claim that

(∗) *If $j, k \in \{0, \ldots, N\}$, then $qe^{s_k g}e^{t_j f}$ is defined and equal to $qe^{t_j f}e^{s_k g}$.*

To prove (∗), we first show that

(#) *If $j \in \{0, \ldots, N\}$, $q' \in K$, and $q'e^{t_i f} \in K$ for all $i \in \{0, \ldots, j\}$, then $q'e^{t_j f}e^{s_1 g}$ is defined and equal to $q'e^{s_1 g}e^{t_j f}$.*

We prove (#) by induction on j. The case when $j = 0$ is trivial. Assume that our conclusion is known to be true for a j such that $0 \leqslant j < N$, and let $q' \in K$ be such that $q'e^{t_i f} \in K$ for all $i \in \{0, \ldots, j+1\}$. Then in particular $q'e^{t_i f} \in K$ for all $i \in \{0, \ldots, j\}$, so the inductive assumption implies that $q'e^{t_j f}e^{s_1 g}$ is defined and equal to $q'e^{s_1 g}e^{t_j f}$. Since $q'e^{t_j f} \in K$, $t_1 \leqslant \varepsilon$, and $s_1 \leqslant \varepsilon$, we can conclude that $q'e^{t_j f}e^{t_1 f}e^{s_1 g}$ (which is equal to $q'e^{t_{j+1} f}e^{s_1 g}$) is defined and equal to $q'e^{t_j f}e^{s_1 g}e^{t_1 f}$. But $q'e^{t_j f}e^{s_1 g} = q'e^{s_1 g}e^{t_j f}$. Therefore $q'e^{t_j f}e^{s_1 g}e^{t_1 f} = q'e^{s_1 g}e^{t_j f}e^{t_1 f} = q'e^{s_1 g}e^{t_{j+1} f}$. It follows that $q'e^{t_{j+1} f}e^{s_1 f}$ is defined and equal to $q'e^{s_1 g}e^{t_{j+1} f}$, completing the proof of (#).

To prove (∗), we let $S(j, k)$ be the statement "$qe^{s_k g}e^{t_j f}$ is defined and equal to $qe^{t_j f}e^{s_k g}$," and then let $\Sigma(k)$ be the statement "$S(j, k)$ is true for every index $j \in \{0, \ldots, N\}$." We prove $\Sigma(k)$ by induction on k.

It is clear that $\Sigma(0)$ is true. Assume that $\Sigma(k)$ is true for a particular k such that $0 \leqslant k < N$. Let $j \in \{0, \ldots, N\}$. We want to prove that $qe^{s_{k+1} g}e^{t_j f}$ is defined and equal to $qe^{t_j f}e^{s_{k+1} g}$. Since $\Sigma(k)$ is true, $qe^{s_k g}e^{t_i f}$ is defined and equal to $qe^{t_i f}e^{s_k g}$ for $i = 0, \ldots, j$. Hence $qe^{s_k g}e^{t_i f}$ is defined and belongs to K for $i = 0, \ldots, j$, because $qe^{t_i f}e^{s_k g} \in K$. It follows from (#), with $q' = qe^{s_k g}$, that $qe^{s_k g}e^{t_j f}e^{s_1 g}$ is defined and equal to $qe^{s_k g}e^{s_1 g}e^{t_j f}$, i.e., to $qe^{s_{k+1} g}e^{t_j f}$. Hence $qe^{t_j f}e^{s_{k+1} g} = qe^{t_j f}e^{s_k g}e^{s_1 g} = qe^{s_k g}e^{t_j f}e^{s_1 g} = qe^{s_{k+1} g}e^{t_j f}$, completing the proof of (∗).

Now that we have proved (∗), we take $k = N$, $j = N$, and conclude that $qe^{tf}e^{sg} = qe^{sg}e^{tf}$, completing our proof. □

6. A simultaneous flow-box theorem for a family of vector fields

Roughly speaking, the so-called "flow-box theorem" states that if $\ell \in \overline{\mathbb{N}}$, f is a vector field of class C^ℓ on an m-dimensional manifold M of class $C^{\ell+1}$, and $q \in M$ is such that $qf \neq 0$, then there exists a coordinate chart $\kappa : U \mapsto \Omega \subseteq \mathbb{R}^m$ of class C^ℓ of M near q such that the coordinate representation f^κ of f on U is a constant vector field on Ω.

As is well known, if two vector fields f and g of class C^ℓ are given, then in general there does not exist a chart κ near q such that *both* vector fields f^κ, g^κ are constant near q. In fact, if qf and qg are linearly independent, then the chart κ exists if and only if the Lie bracket $[f, g]$ vanishes identically on a neighborhood of q. A similar result holds for more than two vector fields: if $f_1, \ldots, f_d$ are vector fields of class C^ℓ such that $qf_1, \ldots, qf_d$ are linearly independent,

then there exists a chart κ of class C^ℓ such that f_j^κ is a constant vector field for $j = 1, \ldots, d$ if and only if all the Lie brackets $[f_i, f_j]$ vanish identically on a neighborhood of q. (This is sometimes referred as the "simultaneous flow box theorem.")

The commutativity result stated in Theorem 5.3 makes it reasonable to expect that a simultaneous flow-box theorem will hold for locally Lipschitz vector fields on a manifold of class C^2, yielding a Lipschitz chart, i.e., a lipeomorphism κ from a neighborhood of a given point q onto an open subset W of $\mathbb{R}^m$. (See [3] for a generalization in the case of a single Lipschitz vector field on a Banach space.) We will now show that this is indeed true, and that the resulting leaves are submanifolds of class $C^{1,1}$.

6.1. The simultaneous flow box theorem

We recall that e_i^m is the ith vector of the canonical basis of $\mathbb{R}^m$, so that $e_i^m = (\delta_i^1, \ldots, \delta_i^m)$, where δ_i^j is the usual Kronecker delta.

Theorem 6.1. *Let M be an m-dimensional manifold of class C^2. Let d be a positive integer, and let $f_1, \ldots, f_d$ be locally Lipschitz vector fields on M such that $q[f_i, f_j] = 0$ for almost all $q \in M$ and all $i, j \in \{1, \ldots, d\}$. (In view of Theorem 5.3, this is equivalent to assuming that $q[f_i, f_j]_{\mathrm{set}} = 0$ for all $q \in M$ and all $i, j \in \{1, \ldots, d\}$.) Let $\bar{q} \in M$ be such that the vectors $\bar{q} f_1, \ldots, \bar{q} f_d$, are linearly independent.*

Then there exists an open neighborhood U of $\bar{q}$, an open cube $W =]-\alpha, \alpha[^m$ in $\mathbb{R}^m$, and a homeomorphism κ from U onto W, such that,

- $\kappa(\bar{q}) = 0$;
- *κ and κ^{-1} are locally Lipschitz;*
- *if $q \in U$, $\kappa(q) = (x^1, \ldots, x^m)^\dagger$, and $i \in \{1, \ldots, d\}$, then qe^{tf_i} is defined for every $t \in \mathbb{R}$ such that $-\alpha < x^i + t < \alpha$, and satisfies*

$$\kappa\left(qe^{tf_i}\right) = \kappa(q) + te_i^m. \tag{43}$$

Proof. Without loss of generality, we may assume that M is an open subset of $\mathbb{R}^m$ and $\bar{q} = 0$, since we can always (i) choose a coordinate chart $\tilde{\kappa} : \widetilde{U} \mapsto \widetilde{W}$ near $\bar{q}$, of class C^2 and such that $\tilde{\kappa}(\bar{q}) = 0$, (ii) replace M with $\widetilde{U}$, and (iii) identify $\widetilde{U}$ with $\widetilde{W}$ via $\tilde{\kappa}$.

Then the f_i are just locally Lipschitz maps from M to $\mathbb{R}^m$. Since the d vectors $f_1(0), \ldots, f_d(0)$ are linearly independent, there exists an invertible linear map $L : \mathbb{R}^m \mapsto \mathbb{R}^m$ such that $L \cdot f_i(0) = e_i^m$ for $i \in \{1, \ldots, d\}$. We can then identify M with $L(M)$, and assume that $f_i(0) = e_i^m$ for $i \in \{1, \ldots, d\}$.

Let r, R be real numbers such that $0 < r < R$ and $R\overline{\mathbb{B}}^m \subseteq M$.

Since the vector fields f_i are locally Lipschitz, by standard results on ordinary differential equations there exists a positive real number $\overline{T}$ such that for every positive integer μ and every μ-tuple $(i(1), \ldots, i(\mu))$ of indices belonging to $\{1, \ldots, d\}$ the point $xe^{t_1 f_{i(1)}} \cdots e^{t_\mu f_{i(\mu)}}$ is well defined and belongs to $R\overline{\mathbb{B}}^m$ whenever $x \in r\overline{\mathbb{B}}^m$ and $t_1, \ldots, t_\mu$ are real numbers such that $|t_1| + \cdots + |t_\mu| \leqslant \overline{T}$.

It then follows from Theorem 5.3 that, if x belongs to $r\overline{\mathbb{B}}^m$, $(i(1), \ldots, i(d))$ is a permutation of the set $\{1, \ldots, d\}$, and $|t_1| + \cdots + |t_d| < \overline{T}$, then

$$xe^{t_1 f_1} \cdots e^{t_d f_d} = xe^{t_{i(1)} f_{i(1)}} \cdots e^{t_{i(d)} f_{i(d)}}.$$

Let us identify $\mathbb{R}^m$ with the product $\mathbb{R}^d \times \mathbb{R}^{m-d}$ and write, for $z = (z^1, \dots, z^m) \in \mathbb{R}^m$, $z = (z^I, z^{II})$, where $z^I = (z^1, \dots, z^d)^\dagger$ and $z^{II} = (z^{d+1}, \dots, z^m)^\dagger$,

Let $\widetilde{W}$ be the cube $]-\tilde{\alpha}, \tilde{\alpha}[^m$, where $\tilde{\alpha}$ is a positive number such that $\sqrt{d}\tilde{\alpha} < r$ and $(m-d)\tilde{\alpha} < \overline{T}$. Write $\widetilde{W} = \widetilde{W}^I \times \widetilde{W}^{II}$, where $\widetilde{W}^I =]-\tilde{\alpha}, \tilde{\alpha}[^d$ and $\widetilde{W}^{II} =]-\tilde{\alpha}, \tilde{\alpha}[^{m-d}$. Then $\widetilde{W}^I \subseteq r\overline{\mathbb{B}}^d$ and

$$\widetilde{W}^{II} \subseteq \left\{\left(\xi^1, \dots, \xi^{m-d}\right)\colon \left|\xi^1\right| + \dots + \left|\xi^{m-d}\right| < \overline{T}\right\}.$$

Define a map $F : \widetilde{W} \mapsto R\overline{\mathbb{B}}^m$ by letting

$$F(x) = F\left(x^I, x^{II}\right) = \left(0, x^{II}\right)e^{x^1 f_1} \cdots e^{x^d f_d} \quad \text{for } x = \left(x^1, \dots, x^n\right)^\dagger \in \widetilde{W}.$$

By standard methods, involving Gronwall's inequality, it is easy to verify that this map is locally Lipschitz. More precisely, it verifies the inequality

$$\left\| F(\tilde{x}) - F(x) \right\| \leqslant (1 + C)e^{dN\overline{T}} \|\tilde{x} - x\|,$$

where N denotes a Lipschitz constant for all the vector fields f_i on $R\overline{\mathbb{B}}^m$ and C is an upper bound for the numbers $\|f_i(x)\|$, for all $x \in R\overline{\mathbb{B}}^m$ and all i. (*Proof.* Given points $x, \tilde{x} \in R\overline{\mathbb{B}}^m$, $i \in \{1, \dots, d\}$, and $t \in \mathbb{R}$ such that xe^{sf_i} and $\tilde{x}e^{sf_i}$ belong to $R\overline{\mathbb{B}}^m$ for all $s \in [\min(t, 0), \max(t, 0)]$, Gronwall's inequality implies that $\|xe^{tf_i} - \tilde{x}e^{tf_i}\| \leqslant e^{N|t|}\|x - \tilde{x}\|$. Hence, if $t, \tilde{t}$ are such that xe^{sf_i} and $\tilde{x}e^{sf_i}$ belong to $R\overline{\mathbb{B}}^m$ for all $s \in [\min(t, 0), \max(t, 0)]$ and all $s \in [\min(\tilde{t}, 0), \max(\tilde{t}, 0)]$, we have

$$\begin{aligned}\left\|xe^{tf_i} - \tilde{x}e^{\tilde{t}f_i}\right\| &\leqslant \left\|xe^{tf_i} - \tilde{x}e^{tf_i}\right\| + \left\|\tilde{x}e^{tf_i} - \tilde{x}e^{\tilde{t}f_i}\right\| \leqslant e^{N|t|}\|x - \tilde{x}\| + C|t - \tilde{t}| \\ &\leqslant e^{N|t|}\left(\|x - \tilde{x}\| + C|t - \tilde{t}|\right),\end{aligned}$$

using the fact that $\tilde{x}e^{tf_i} - \tilde{x}e^{\tilde{t}f_i} = \int_{\tilde{t}}^t \tilde{x}e^{sf_i} f_i\, ds$. In then follows by induction on k that

$$\left\|xe^{t^1 f_1} \cdots e^{t^k f_k} - \tilde{x}e^{\tilde{t}^1 f_1} \cdots e^{\tilde{t}^k f_k}\right\| \leqslant e^{N\sum_{j=1}^k |t^j|}\left(\|x - \tilde{x}\| + C\left|t^1 - \tilde{t}^1\right| + \dots + C\left|t^k - \tilde{t}^k\right|\right)$$

if $x \in r\overline{\mathbb{B}}^m$, $\tilde{x} \in r\overline{\mathbb{B}}^m$, $|t^1| + \dots + |t^k| \leqslant \overline{T}$, and $|\tilde{t}^1| + \dots + |\tilde{t}^k| \leqslant \overline{T}$. If we then take $k = d$, $x = (0, x^{d+1}, \dots, x^m)$, $\tilde{x} = (0, \tilde{x}^{d+1}, \dots, \tilde{x}^m)$, and let $t^i = x^i$ and $\tilde{t}^i = \tilde{x}^i$ for $i = 1, \dots, m$, we get the desired inequality.)

We claim that F *defines a local lipeomorphism near* $x = 0$. More precisely, we will prove the following result.

Lemma 6.2. *There exists a positive real number α such that, if $W =]-\alpha, \alpha[^m$, then $\Omega = F(W)$ is an open neighborhood of* 0 *and the restriction of F to W is a lipeomorphism onto Ω.*

Proof. The crucial fact is the following local invertibility result for locally Lipschitz maps (cf. [5, Theorem 3.12]).

Proposition 6.3. *Let ν be a positive integer, let y, z be points of $\mathbb{R}^\nu$, let N_y and N_z be neighborhoods of y and z, respectively, and let $\Phi : N_y \mapsto N_z$ be a locally Lipschitz map. Assume that $\Phi(y) = z$ and that all the members of the Clarke generalized Jacobian $\partial\Phi(y)$ of Φ at y*

are invertible linear maps. Then there exist a positive real number η and an open neighborhood $A_z \subseteq N_z$ such that the restriction of G to the open ball $y+\eta\mathbb{B}^m$ is a lipeomorphism from $y+\eta\mathbb{B}^m$ onto A_z.

In view of Proposition 6.3 it is sufficient to prove that $\partial F(0)$ does not contain any noninvertible matrix. We will, in fact, prove the much stronger conclusion that

$$\partial F(0) = \{\mathrm{id}_m\}, \tag{44}$$

where id_m denotes the $m \times m$ identity matrix.

By definition, $\partial F(0)$ is the convex hull of the set of $m \times m$ matrices J such that

$$J = \lim_{k\to\infty} DF(x_k),$$

where $(x_k)_{k\in\mathbb{N}}$ is a sequence in $DIFF(F)$ which converges to $x = 0$ and is such that the above limit exists.

For every $x \in R\mathbb{B}^m \cap DIFF(F)$, let us set

$$\frac{\partial F}{\partial x^j}(x) = \begin{pmatrix} \frac{\partial F^1}{\partial x^j}(x) \\ \vdots \\ \frac{\partial F^m}{\partial x^j}(x) \end{pmatrix},$$

so that we can write

$$DF(x) = \left(\frac{\partial F}{\partial x^1}(x), \dots, \frac{\partial F}{\partial x^m}(x)\right).$$

The first d columns of $DF(x)$ can be easily calculated. In fact, let us choose $j \in \{1, \dots, d\}$, and let $(\gamma_1, \dots, \gamma_{d-1})$ be the string obtained by deleting j from $(1, \dots, d)$. Then, thanks to the commutativity of the flows, one has

$$\begin{aligned} \frac{\partial F}{\partial x^j}(x) &= \frac{\partial}{\partial x^j}\big((0, x^{II})e^{x^1 f_1} \cdots e^{x^d f_d}\big) = \frac{\partial}{\partial x^j}\big((0, x^{II})e^{x^{\gamma_1} f_{\gamma_1}} \cdots e^{x^{\gamma_{d-1}} f_{\gamma_{d-1}}} e^{x^j f_j}\big) \\ &= \big(0, x^{II}\big)e^{x^{\gamma_1} f_{\gamma_1}} \cdots e^{x^{\gamma_{d-1}} f_{\gamma_{d-1}}} e^{x^j f_j} f_j = \big(0, x^{II}\big)e^{x^1 f_1} \cdots e^{x^d f_d} f_j = xFf_j, \end{aligned}$$

which, with a more conventional notation, can also be written as

$$\frac{\partial F}{\partial x^j}(x) = f_j\big(F(x)\big). \tag{45}$$

As for the derivatives of F with respect to the last $m-d$ variables, we shall prove that, *for every $j \in \{d+1, \dots, m\}$, every $\delta < R$, and every $x \in \delta\mathbb{B}^m \cap DIFF(F)$ one has*

$$\frac{\partial F}{\partial x^j}(x) \in e_j^m + \varepsilon_\delta \mathbb{B}^m, \tag{46}$$

where $\varepsilon_\delta = d\delta N e^{Nd\delta}$.

For this purpose, for any $z = (z^I, z^{II}) \in \mathbb{R}^m$, let us set $S_z = \sum_{j=1}^d |z^j|$, and, for $j \in \{1, \dots, d\}$, define $\lambda_j = 1$ if $z^j \geqslant 0$ and $\lambda_j = -1$ if $z^j < 0$. Let

$$\tau_z : [0, S_z] \mapsto \mathbb{R}^d$$

be the unique continuous path such that $\tau(0) = (0, \dots, 0)$ and

$$\frac{d\tau_x}{ds}(s) = \lambda_j e_j^d \quad \text{whenever } s \in \left]\sum_{i=1}^{j-1}|z^i|, \sum_{i=1}^{j}|z^i|\right[, \ j \in \{1, \dots, d\}.$$

Notice, in particular, that for every $z \in \mathbb{R}^n$ the number S_z and the path τ_z depend on z^I only. Moreover, one has

$$\tau_z(S_z) = z^I.$$

Write $\tau_z(s) = (\tau_z^1(s), \dots, \tau_z^d(s))$. Then, for each $j \in \{1, \dots, d\}$, the function $\tau_z^j : [0, S_z] \mapsto \mathbb{R}$ is continuous, vanishes identically on the interval $[0, \sum_{i=1}^{j-1} |z^i|]$, has the constant value z^j on $[\sum_{i=1}^{j} |z^i|, S_z]$, and is linear on $[\sum_{i=1}^{j-1} |z^i|, \sum_{i=0}^{j} |z^i|]$.

Let $\delta \in [0, R[$, and, for every $y \in \delta\mathbb{B}^m$, let us consider the curve $\xi_y : [0, S_y] \mapsto \mathbb{R}^m$ defined by

$$\xi_y(s) = F\big(\tau_y(s), y^{II}\big).$$

Notice that, in particular,

$$\xi_y(S_y) = F(y). \tag{47}$$

It is easy to check that the curve ξ_y is the unique solution of the Cauchy problem

$$\begin{cases} \frac{d\xi}{ds}(s) = \sum_{j=1}^d f_j(\xi(s)) \cdot \frac{d\tau_y^j}{ds}(s), \\ \xi(0) = (0, y^{II}). \end{cases}$$

For every $y \in \delta\mathbb{B}^m$ we define a time-varying vector field $G_y : r\mathbb{B}^m \times [0, S_y] \mapsto \mathbb{R}^m$ by setting

$$G_y(z, s) = \sum_{j=1}^d f_j(z) \cdot \frac{d\tau_y^j}{ds}(s),$$

so that $\xi_y(\cdot)$ is the solution on $[0, S_y]$ of the Cauchy problem

$$\begin{cases} \frac{d\xi}{ds}(s) = G_y(\xi(s), s), \\ \xi(0) = (0, y^{II}). \end{cases}$$

Fix $x \in \delta\mathbb{B}^m \cap DIFF(F)$ and $j \in \{d+1, \dots, n\}$, and observe that $S_{x+he_j^m} = S_x$ and $G_{x+he_j^m} = G_x$. Therefore, for a sufficiently small $h > 0$, (47) implies

$$\frac{F(x+he_j^m) - F(x)}{h} = \frac{1}{h}\left(he_j^m + \int_0^{S_x} \left[G_x\big(\xi_{x+he_j^m}(s), s\big) - G_x\big(\xi_x(s), s\big)\right] ds\right).$$

Once again using Gronwall's inequality, one obtains

$$\left\|\xi_{x+he_j^m}(s) - \xi_x(s)\right\| \leqslant he^{Ns}$$

for every $s \in [0, S_x]$. Since the map $G_x(\cdot, s)$ is N-Lipschitz on rB for every $s \in [0, S_x]$, it follows that

$$\begin{aligned}\left\|\frac{F(x+he_j^m) - F(x)}{h} - e_j^m\right\| &\leqslant \frac{1}{h}\int_0^{S_x} \left\|G_x\big(\xi_{x+he_j^m}(s), s\big) - G_{x+}\big(\xi_x(s), s\big)\right\| ds \\ &\leqslant \frac{1}{h}\int_0^{S_x} Nhe^{Ns}\, ds \leqslant S_x N e^{NS_x} \leqslant d\delta\, N e^{Nd\delta}.\end{aligned}$$

If we let h go to 0, we get estimate (46). Then (45) and (46) imply (44), and the lemma is proved. □

Conclusion of the proof of Theorem 6.1. To conclude the proof, let us set $\kappa = F^{-1} : U \to W$, which, by Lemma 6.2 is a lipeomorphism such that $\kappa(\bar{q}) = 0$. We now need to verify that κ satisfies (43). If $t \in \mathbb{R}$ is such that $-\alpha < x^i + t < \alpha$, setting $(y^1, \dots, y^d)^\dagger = y = x^I + te_d^i$, one has $(y, x^{II}) \in W$. Hence, from

$$q = F(x) = \big(0, x^{II}\big)e^{x^1 f^1} \cdots e^{x^d f^d}$$

and Theorem 5.3, one obtains

$$\begin{aligned}\kappa(q) + te_m^i = x + te_m^i = \kappa \circ \kappa^{-1}\big(y, x^{II}\big) &= \kappa\big(\big(0, x^{II}\big)e^{y^1 f^1} \cdots e^{y^d f^d}\big) \\ &= \kappa\big(\big(0, x^{II}\big)e^{x^1 f^1} \cdots e^{x^d f^d} e^{tf_i}\big) = \kappa\big(qe^{tf_i}\big). \qquad \square\end{aligned}$$

6.2. Regularity of the leaves

Let M, $f_1, \dots, f_d$, $q \in M$, α, U, W, and κ be as in the statement of Theorem 6.1. The *leaves* of the foliation defined by $f_1, \dots, f_d$ on U are the subsets $\mathcal{L}_\gamma = \kappa^{-1}(]-\alpha, \alpha[^d \times \{\gamma\})$, for all $\gamma \in]-\alpha, \alpha[^{m-d}$.

We recall that a map $\mu : P \mapsto Q$ between manifolds of class C^2 is *of class* $C^{1,1}$ if it is of class C^1 and its first-order partial derivatives are locally Lipschitz (with respect to arbitrary coordinate charts of class C^2 on P, Q).

Theorem 6.4. *For every $\gamma \in]-\alpha, \alpha[^{m-d}$, the leaf $\mathcal{L}_\gamma$ is a submanifold of M of class $C^{1,1}$.*

Proof. It is clear that L_γ is the image of the cube $]-\alpha, \alpha[^d$ under the map μ_γ given by $\mu_\gamma(t_1, \dots, t_d) = \kappa^{-1}(0, \gamma)e^{t_1 f_1} \cdots e^{t_d f_d}$. This map is of class $C^{1,1}$, because in view of Theorem 5.3, its first-order partial derivatives $\frac{\partial \mu_\gamma}{\partial t_j}$ are given by

$$\frac{\partial \mu_\gamma}{\partial t_j}(t_1, \dots, t_d) = \kappa^{-1}(0, \gamma)e^{t_1 f_1} \cdots e^{t_d f_d} f_j.$$

Furthermore, μ_γ is injective, because $\kappa(\mu_\gamma(t_1, \dots, t_d)) = (t_1, \dots, t_d, \gamma)$. Finally, the differential $d\mu_\gamma(t_1, \dots, t_d)$ is also injective, because the vectors qf_i, $i = 1, \dots, d$, are linearly independent for each $q \in \kappa^{-1}(W)$. It then follows by standard arguments using the inverse function theorem that L_γ is a submanifold of class $C^{1,1}$. □

7. A counterexample about higher-order brackets

7.1. The need for a definition of higher degree brackets

In this paper, we have shown that the notion of set-valued Lie bracket $[f, g]_{\text{set}}$ of two locally Lipschitz vector fields f and g is a reasonable generalization of the classical Lie bracket, which has enabled us to extend to Lipschitz vector fields facts (I)–(III) of the introduction. In addition, this notion has also been used in our nonsmooth version of Chow's local controllability theorem, proved in [7].

In view of this, it is natural to wonder whether our approach can be used to define higher-order brackets and prove higher-order asymptotic formulae and a more general Chow theorem. In fact, under suitable regularity conditions, our definition of the Lie bracket of two locally Lipschitz vector fields leads directly to a notion of set-valued high-order bracket. For example, if g and h are of class $C^{1,1}$, and f is locally Lipschitz, then $[g, h]$ is a locally Lipschitz vector field, so $[f, [g, h]]_{\text{set}}$ is well defined, according to Definition 3.1.

We are going to show that *this definition of* $[f, [g, h]]_{\text{set}}$ *does not lead to the correct asymptotic formula*. Precisely, if we define

$$\begin{aligned} S(t, q) &= qe^{tf}\left(e^{tg}e^{th}e^{-tg}e^{-th}\right)e^{-tf}\left(e^{tg}e^{th}e^{-tg}e^{-th}\right)^{-1} \\ &= qe^{tf}e^{tg}e^{th}e^{-tg}e^{-th}e^{-tf}e^{th}e^{tg}e^{-th}e^{-tg} \end{aligned} \tag{48}$$

then it is well known that

$$S(t, q) \sim q + t^3 q\big[f, [g, h]\big] + o\big(t^3\big),$$

if g and h are of class C^2 and f is of class C^1. So, when $g \in C^{1,1}$, $h \in C^{1,1}$, and f is locally Lipschitz, the correct asymptotic formula would have to say that

$$\lim_{t \to 0} t^{-3} \operatorname{dist}\big(S(t, q) - q, q\big[f, [g, h]\big]_{\text{set}}\big) = 0.$$

If, in addition, $f = [g, h]$, then $[f, [g, h]]_{\text{set}} = 0$, so the formula would imply

$$S(t, q) - q = o\big(t^3\big). \tag{49}$$

7.2. The counterexample

We now show, by means of an example, that *the asymptotic formula* (49) *can fail to be true, if* $g \in C^{1,1}$, $h \in C^{1,1}$, $f = [g, h]$, *and* $S(t, q)$ *is defined as in* (48). For this purpose, we define three vector fields f, g, h on $\mathbb{R}^2$, by taking g and h to be two vector fields of class $C^{1,1}$, to be chosen below, and letting $f = [g, h]$.

In what follows, for any nonnegative integer r, we shall say that a function is *of class* $C^{r,1}$ if it is of class C^r and its derivatives of order r are Lipschitz continuous.

In order to define g and h, we first let $\Phi : \mathbb{R} \mapsto \mathbb{R}$ be a function of class $C^{2,1}$ to be chosen later. We let $\varphi(x) = (x + 1)\Phi'(x)$ and $\psi = \varphi'$, so φ and ψ are of class $C^{1,1}$ and $C^{0,1}$, respectively. We let $\sigma = \psi'$, so σ is an L^∞ function. Let us define three vector fields f, g, h, by

$$g = e_1, \qquad h = (1 + x)e_1 + \varphi(x)e_2, \qquad f = [g, h] = e_1 + \psi(x)e_2.$$

Notice that, if we impose the extra conditions

$$\psi(0) = \varphi(0) = \Phi(0) = 0, \tag{50}$$

and restrict ourselves to values of x such that $x > -1$, then the choice of σ completely determines our functions, since

$$\psi(x) = \int_0^x \sigma(r)\,dr, \qquad \varphi(x) = \int_0^x \psi(r)\,dr, \qquad \Phi(x) = \int_0^x \frac{\varphi(r)}{1 + r}\,dr.$$

We shall prove the following two facts:

Fact 1. *If* ψ *is differentiable at* 0, *then* $\mathcal{S}(t, (0, 0)^\dagger) = o(t^3)$ *as* $t \to 0$, *as it is expected from the classical case.*

Fact 2. *There exists a function* σ, *necessarily discontinuous at* $x = 0$, *such that the formula* $\mathcal{S}(t, (0, 0)^\dagger) = o(t^3)$ *fails to be true.*

7.2.1. Some preliminary computations

We first compute $S(t, q)$ in terms of the functions Φ and φ. For this purpose, we need to compute the flows of f, g, and h. We claim that, if we let $q = \binom{x}{y}$, then

$$\begin{pmatrix} x \\ y \end{pmatrix} e^{tg} = \begin{pmatrix} x + t \\ y \end{pmatrix}, \tag{51}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} e^{tf} = \begin{pmatrix} x + t \\ y + \varphi(x + t) - \varphi(x) \end{pmatrix}, \tag{52}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} e^{th} = \begin{pmatrix} (1 + x)e^t - 1 \\ y + \Phi((1 + x)e^t - 1) - \Phi(x) \end{pmatrix}. \tag{53}$$

Indeed, formulae (51) and (52) are trivial. To verify (53), we let $\xi(t)$, $\eta(t)$ denote the components of the right-hand side of (53), and observe that $\dot{\xi}(t) = 1 + \xi(t)$, $\eta(t) = \Phi(\xi(t)) + y - \Phi(x)$, $\dot{\eta}(t) = \Phi'(\xi(t))\dot{\xi}(t) = \Phi'(\xi(t))(1+\xi(t)) = \varphi(\xi(t))$, and $\xi(0) = x$, $\eta(0) = y$, so $t \mapsto (\xi(t), \eta(t))$ is an integral curve of h that goes through q at time 0. Using these formulae, we find:

$$qe^{tf} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x+t \\ y+\varphi(x+t)-\varphi(x) \end{pmatrix} = \begin{pmatrix} x+t \\ y+A \end{pmatrix},$$

$$qe^{tf}e^{tg} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x+2t \\ y+A \end{pmatrix},$$

where

$$A = \varphi(x+t) - \varphi(x).$$

Then

$$qe^{tf}e^{tg}e^{th} = \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} (1+x+2t)e^t - 1 \\ y+A+B \end{pmatrix},$$

where

$$B = \Phi(x_3) - \Phi(x_2) = \Phi\big((1+x+2t)e^t - 1\big) - \Phi(x+2t).$$

In the next step, we get

$$qe^{tf}e^{tg}e^{th}e^{-tg} = \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} = \begin{pmatrix} (1+x+2t)e^t - 1 - t \\ y+A+B \end{pmatrix}$$

and then

$$qe^{tf}e^{tg}e^{th}e^{-tg}e^{-th} = \begin{pmatrix} x_5 \\ y_5 \end{pmatrix} = \begin{pmatrix} x+2t-te^{-t} \\ y+A+B+C \end{pmatrix},$$

where

$$C = \Phi(x_5) - \Phi(x_4) = \Phi\big(x+2t-te^{-t}\big) - \Phi\big((1+x+2t)e^t - 1 - t\big).$$

Then

$$qe^{tf}e^{tg}e^{th}e^{-tg}e^{-th}e^{-tf} = \begin{pmatrix} x_6 \\ y_6 \end{pmatrix} = \begin{pmatrix} x+t-te^{-t} \\ y+A+B+C+D \end{pmatrix},$$

where

$$D = \varphi(x_6) - \varphi(x_5) = \varphi\big(x+t-te^{-t}\big) - \varphi\big(x+2t-te^{-t}\big).$$

One more step yields

$$qe^{tf}e^{tg}e^{th}e^{-tg}e^{-th}e^{-tf}e^{th} = \begin{pmatrix} x_7 \\ y_7 \end{pmatrix} = \begin{pmatrix} (1+x+t)e^t - 1 - t \\ y + A + B + C + D + E \end{pmatrix},$$

where

$$E = \Phi(x_7) - \Phi(x_6) = \Phi\big((1+x+t)e^t - 1 - t\big) - \Phi\big(x + t - te^{-t}\big).$$

Next, we have

$$qe^{tf}e^{tg}e^{th}e^{-tg}e^{-th}e^{-tf}e^{th}e^{tg} = \begin{pmatrix} x_8 \\ y_8 \end{pmatrix} = \begin{pmatrix} (1+x+t)e^t - 1 \\ y + A + B + C + D + E \end{pmatrix},$$

$$qe^{tf}e^{tg}e^{th}e^{-tg}e^{-th}e^{-tf}e^{th}e^{tg}e^{-th} = \begin{pmatrix} x_9 \\ y_9 \end{pmatrix} = \begin{pmatrix} x + t \\ y + A + B + C + D + E + F \end{pmatrix},$$

where

$$F = \Phi(x_9) - \Phi(x_8) = \Phi(x+t) - \Phi\big((1+x+t)e^t - 1\big).$$

Finally, we get

$$\begin{aligned} S(t,q) &= qe^{tf}e^{tg}e^{th}e^{-tg}e^{-th}e^{-tf}e^{th}e^{tg}e^{-th}e^{-tg} \\ &= \begin{pmatrix} x_{10} \\ y_{10} \end{pmatrix} = \begin{pmatrix} x \\ y + A + B + C + D + E + F \end{pmatrix}. \end{aligned}$$

So what we need to know is whether the sum $\mathcal{S} = A + B + C + D + E + F$ (which, of course, depends on x, y, and t) is $o(t^3)$ for fixed (x, y), as $t \downarrow 0$. We will take $x = 0$, and use the fact that $\varphi(0) = 0$. Then

$$\begin{aligned} \mathcal{S} = {} & \varphi(t) + \Phi\big((1+2t)e^t - 1\big) - \Phi(2t) + \Phi\big(2t - te^{-t}\big) - \Phi\big((1+2t)e^t - 1 - t\big) \\ & + \varphi\big(t - te^{-t}\big) - \varphi\big(2t - te^{-t}\big) + \Phi\big((1+t)e^t - 1 - t\big) \\ & - \Phi\big(t - te^{-t}\big) + \Phi(t) - \Phi\big((1+t)e^t - 1\big), \end{aligned}$$

so

$$\begin{aligned} \mathcal{S} = {} & \varphi(t) + \Phi\big(e^t - 1 + 2te^t\big) - \Phi(2t) + \Phi\big(2t - te^{-t}\big) - \Phi\big(e^t - 1 + 2te^t - t\big) \\ & + \varphi\big(t - te^{-t}\big) - \varphi\big(2t - te^{-t}\big) + \Phi\big(e^t - 1 + te^t - t\big) - \Phi\big(t - te^{-t}\big) + \Phi(t) \\ & - \Phi\big(e^t - 1 + te^t\big). \end{aligned}$$

If $0 < t \leqslant 1$, then the bounds $e^t \leqslant e$, $e^{-t} \leqslant 1$, $e^t - 1 \leqslant et$, guarantee that all the arguments of the Φ and φ functions in the above expression lie in the interval $[0, 3et]$. If we make the further restriction that $t < \frac{1}{10}$, this guarantees that all these arguments belong to the interval $[0, \rho]$, where

$\rho = \frac{3e}{10}$, so $\rho < 1$. Let c be an upper bound for $|\sigma(s)|$, for $s \in [0, \rho]$. Then (50) implies that the functions ψ, φ satisfy the bounds

$$\left|\psi(s)\right| \leqslant cs, \qquad \left|\varphi(s)\right| \leqslant \frac{cs^2}{2}, \qquad \left|\Phi(s)\right| \leqslant \frac{cs^3}{6} \quad \text{for } s \in [0, \rho].$$

Clearly, the function Φ is Lipschitz on $[0, 3et]$ with constant kt^2, where $k = \frac{c}{2}$. Therefore, if in an expression $\Phi(x)$ we replace x by $\tilde{x}$, where $|\tilde{x} - x|$ is bounded by a constant κ times t^2, the resulting error will be $O(t^4)$, and will not affect our asymptotic calculation to order t^3. Hence the argument $e^t - 1 + 2te^t$ of the first Φ term can be replaced by $3t$. Similarly, $2t - te^{-t}$ can be replaced by t, $e^t - 1 + 2te^t - t$ can be replaced by $2t$, $e^t - 1 + te^t - t$ can be replaced by t, $t - te^{-t}$ can be replaced by 0, and $e^t - 1 + te^t$ can be replaced by $2t$. The result is an expression $\widehat{\mathcal{S}}$ such that $\widehat{\mathcal{S}} - \mathcal{S} = o(t^3)$, given by

$$\begin{aligned}\widehat{\mathcal{S}} &= \varphi(t) + \Phi(3t) - \Phi(2t) + \Phi(t) - \Phi(2t) \\ &\quad + \varphi\bigl(t - te^{-t}\bigr) - \varphi\bigl(2t - te^{-t}\bigr) + \Phi(t) - \Phi(0) + \Phi(t) - \Phi(2t) \\ &= \varphi(t) + \varphi\bigl(t - te^{-t}\bigr) - \varphi\bigl(2t - te^{-t}\bigr) + \Phi(3t) - 3\Phi(2t) + 3\Phi(t).\end{aligned}$$

A similar replacement is possible for the arguments of φ, except that in this case we only know that φ is Lipschitz on $[0, 3et]$ with constant ct, so we can only replace x by $\tilde{x}$ if $|\tilde{x} - x| = O(t^3)$. Since $t - te^{-t} = t^2 + O(t^3)$, and $2t - te^{-t} = t + t^2 + O(t^3)$, we can replace $\widehat{\mathcal{S}}$ by $\mathcal{S}^{\#}$, given by

$$\mathcal{S}^{\#} = \varphi(t) + \varphi\bigl(t^2\bigr) - \varphi\bigl(t + t^2\bigr) + \Phi(3t) - 3\Phi(2t) + 3\Phi(t),$$

without changing the asymptotic behavior to order t^3 as $t \to 0$. Finally, if we let

$$\overline{\Phi}(x) = \int_0^x \varphi(r)\,dr,$$

we have

$$\Phi(x) - \overline{\Phi}(x) = \int_0^x \varphi(r)\left(\frac{1}{1+r} - 1\right)dr = -\int_0^x \frac{r\varphi(r)}{1+r}\,dr,$$

so

$$\left|\Phi(x) - \overline{\Phi}(x)\right| \leqslant x\int_0^x \left|\varphi(r)\right| dr \leqslant kx\int_0^x r^2\,dr = \frac{kx^4}{3}.$$

Hence we can replace Φ by $\overline{\Phi}$ without affecting the desired asymptotics. The result is

$$\overline{\mathcal{S}} = \varphi(t) + \varphi\bigl(t^2\bigr) - \varphi\bigl(t + t^2\bigr) + \overline{\Phi}(3t) - 3\overline{\Phi}(2t) + 3\overline{\Phi}(t).$$

7.2.2. Proof of Fact 1

We shall prove that *if ψ is differentiable at* 0, *then* $\overline{\mathcal{S}} = o(t^3)$ *as* $t \to 0$, which is equivalent to Fact 1. Let $a = \psi'(0)$. Then

$$\psi(x) = ax + o(x), \qquad \varphi(x) = \frac{ax^2}{2} + o\big(x^2\big) \quad \text{and} \quad \overline{\Phi}(x) = \frac{ax^3}{6} + o\big(x^3\big).$$

In particular, $\varphi(t^2) = \frac{at^4}{2} + o(t^4) = o(t^3)$. Also,

$$\begin{aligned}
&\varphi\big(t + t^2\big) - \varphi(t) \int_t^{t+t^2} \psi(s)\,ds \\
&\quad = \psi(t)t^2 + \int_t^{t+t^2} \big(\psi(s) - \psi(t)\big)\,ds = \big(at + o(t)\big)t^2 + \int_t^{t+t^2} \int_t^s \sigma(r)\,dr\,ds \\
&\quad = at^3 + o\big(t^3\big) + \int_t^{t+t^2} \int_t^s \sigma(r)\,dr\,ds = at^3 + o\big(t^3\big).
\end{aligned}$$

Finally,

$$\begin{aligned}
\overline{\Phi}(3t) &= \frac{a(3t)^3}{6} + o\big(t^3\big) = \frac{27at^3}{6} + o\big(t^3\big), \\
\overline{\Phi}(2t) &= \frac{a(2t)^3}{6} + o\big(t^3\big) = \frac{8at^3}{6} + o\big(t^3\big), \qquad \overline{\Phi}(t) = \frac{at^3}{6} + o\big(t^3\big),
\end{aligned}$$

so

$$\overline{\Phi}(3t) - 3\overline{\Phi}(2t) + 3\overline{\Phi}(t) = \frac{at^3}{6}(27 - 24 + 3) + o\big(t^3\big) = at^3 + o\big(t^3\big).$$

Hence

$$\begin{aligned}
\overline{\mathcal{S}} &= \varphi(t) + \varphi\big(t^2\big) - \varphi\big(t + t^2\big) + \overline{\Phi}(3t) - 3\overline{\Phi}(2t) + 3\overline{\Phi}(t) \\
&= -at^3 + o\big(t^3\big) + at^3 + o\big(t^3\big) = o\big(t^3\big),
\end{aligned}$$

and the proof is complete.

7.2.3. Proof of Fact 2

In order to prove Fact 2, let us evaluate $\overline{\mathcal{S}}$ for a particular choice of a bounded discontinuous σ. We define

$$\sigma(x) = (-1)^k \quad \text{if } x \in \mathbb{R},\ 2^{-k-1} < x \leqslant 2^{-k},\ k \in \mathbb{Z},$$

and supplement the definition by letting $\sigma(x) = 0$ for $x \leqslant 0$. Clearly, then,

$$\int_{2^{-k-1}}^{2^{-k}} \sigma(x)\,dx = (-1)^k 2^{-k-1}.$$

Therefore

$$\begin{aligned}\psi\big(2^{-k}\big) &= \sum_{j\in\mathbb{Z},\, j\geqslant k} \int_{2^{-j-1}}^{2^{-j}} \sigma(x)\,dx = \sum_{j\in\mathbb{Z},\, j\geqslant k} (-1)^j 2^{-j-1} \\ &= (-1)^k 2^{-k-1} \sum_{j=0}^{\infty} (-1)^j 2^{-j} = (-1)^k 2^{-k-1} \frac{1}{1+\frac{1}{2}} = \frac{(-1)^k 2^{-k}}{3}.\end{aligned}$$

Then, since ψ is linear on the interval $[2^{-k-1}, 2^{-k}]$, and has derivative equal to $(-1)^k$, we find

$$\psi(x) = \psi\big(2^{-k-1}\big) + (-1)^k\big(x - 2^{-k-1}\big) = \frac{(-1)^{k+1} 2^{-k-1}}{3} + (-1)^k\big(x - 2^{-k-1}\big)$$

for $x \in \mathbb{R}$, $2^{-k-1} < x \leqslant 2^{-k}$, $k \in \mathbb{Z}$. We can then integrate this and find

$$\varphi(x) - \varphi\big(2^{-k-1}\big) = \int_{2^{-k-1}}^{x} \psi(r)\,dr = \psi\big(2^{-k-1}\big)\big(x - 2^{-k-1}\big) + (-1)^k \frac{(x - 2^{-k-1})^2}{2},$$

so

$$\varphi(x) - \varphi\big(2^{-k-1}\big) = \frac{(-1)^{k+1} 2^{-k-1}}{3}\big(x - 2^{-k-1}\big) + (-1)^k \frac{(x - 2^{-k-1})^2}{2}. \tag{54}$$

In particular, if we take $x = 2^{-k}$, so that $x - 2^{-k-1} = 2^{-k-1}$, we get

$$\begin{aligned}\varphi\big(2^{-k}\big) - \varphi\big(2^{-k-1}\big) &= \frac{(-1)^{k+1} 2^{-k-1}}{3} 2^{-k-1} + (-1)^k \frac{(2^{-k-1})^2}{2} \\ &= (-1)^k 2^{-2k-2}\left(-\frac{1}{3} + \frac{1}{2}\right) = \frac{(-1)^k 2^{-2k}}{24}.\end{aligned}$$

Therefore

$$\begin{aligned}\varphi\big(2^{-k}\big) &= \sum_{j\in\mathbb{Z},\, j\geqslant k} \varphi\big(2^{-j}\big) - \varphi\big(2^{-j-1}\big) = \sum_{j\in\mathbb{Z},\, j\geqslant k} \frac{(-1)^j 2^{-2j}}{24} = \frac{(-1)^k 2^{-2k}}{24} \sum_{j=0}^{\infty} (-1)^j 2^{-2j} \\ &= \frac{(-1)^k 2^{-2k}}{24} \times \frac{1}{1+\frac{1}{4}} = \frac{(-1)^k 2^{-2k}}{24} \times \frac{4}{5} = \frac{(-1)^k 2^{-2k}}{30}.\end{aligned}$$

Then (54) implies that, for $2^{-k-1} < x \leqslant 2^{-k}$, $\varphi(x)$ is given by

$$\varphi(x) = \frac{(-1)^{k+1}2^{-2k-2}}{30} + \frac{(-1)^{k+1}2^{-k-1}}{3}\left(x - 2^{-k-1}\right) + (-1)^k \frac{(x - 2^{-k-1})^2}{2},$$

i.e., by

$$\varphi(x) = \frac{(-1)^k}{120}\left(-2^{-2k} - 20 \cdot 2^{-k}\left(x - 2^{-k-1}\right) + 60\left(x - 2^{-k-1}\right)^2\right). \tag{55}$$

We are going to use this expression to evaluate the φ part of $\widehat{\mathcal{S}}$, taking $t = t_k = 2^{-k}$. First, we observe that

$$\varphi\left(t_k^2\right) = \varphi\left(2^{-2k}\right) = \frac{2^{-4k}}{30} = O\left(t_k^4\right).$$

To compute $\varphi(t_k + t_k^2) - \varphi(t_k)$ we use (54) with $k-1$ instead of k, and $x = 2^{-k} + 2^{-2k}$. We get

$$\begin{aligned}
&\varphi\left(t_k + t_k^2\right) - \varphi(t_k) \\
&\quad = \frac{(-1)^k 2^{-k}}{3} \cdot 2^{-2k} + (-1)^{k+1}\frac{(2^{-2k})^2}{2} \\
&\quad = \frac{(-1)^k}{3} \cdot 2^{-3k} + \frac{(-1)^{k+1}}{2}2^{-4k} = \frac{(-1)^k}{3} \cdot t_k^3 + \frac{(-1)^{k+1}}{2}t_k^4 = \frac{(-1)^k}{3} \cdot t_k^3 + O\left(t_k^4\right).
\end{aligned}$$

Hence

$$\varphi(t_k) + \varphi\left(t_k^2\right) - \varphi\left(t_k + t_k^2\right) = \frac{(-1)^{k+1}}{3} \cdot t_k^3 + O\left(t_k^4\right). \tag{56}$$

In order to evaluate the $\overline{\Phi}$ part of $\widehat{\mathcal{S}}$, taking $t = t_k = 2^{-k}$, we need one more integration. We have

$$\begin{aligned}
&\overline{\Phi}(x) - \overline{\Phi}\left(2^{-k-1}\right) \\
&\quad = \varphi\left(2^{-k-1}\right)\left(x - 2^{-k-1}\right) + \frac{1}{2}\psi\left(2^{-k-1}\right)\left(x - 2^{-k-1}\right)^2 + \frac{(-1)^k}{6}\left(x - 2^{-k-1}\right)^3 \\
&\quad = \frac{(-1)^{k+1}2^{-2k-2}}{30}\left(x - 2^{-k-1}\right) + \frac{1}{2}\frac{(-1)^{k+1}2^{-k-1}}{3}\left(x - 2^{-k-1}\right)^2 + \frac{(-1)^k}{6}\left(x - 2^{-k-1}\right)^3.
\end{aligned} \tag{57}$$

If we substitute 2^{-k} for x, we get

$$\begin{aligned}
&\overline{\Phi}\left(2^{-k}\right) - \overline{\Phi}\left(2^{-k-1}\right) \\
&\quad = \frac{(-1)^{k+1}2^{-2k-2}}{30} \cdot 2^{-k-1} + \frac{1}{2}\frac{(-1)^{k+1}2^{-k-1}}{3}\left(2^{-k-1}\right)^2 + \frac{(-1)^k}{6}\left(2^{-k-1}\right)^3 \\
&\quad = (-1)^{k+1}2^{-3k-3} \cdot \frac{1}{30} = \frac{(-1)^{k+1}2^{-3k}}{240}.
\end{aligned}$$

Then

$$\begin{aligned}\overline{\Phi}\big(2^{-k}\big) &= \sum_{j\in\mathbb{Z},\, j\geqslant k}\big(\overline{\Phi}\big(2^{-j}\big)-\overline{\Phi}\big(2^{-j-1}\big)\big)\\ &= \sum_{j\in\mathbb{Z},\, j\geqslant k}\frac{(-1)^{j+1}2^{-3j}}{240} = \frac{(-1)^{k+1}2^{-3k}}{240}\sum_{j=0}^{\infty}(-1)^j 2^{-3j} = \frac{(-1)^{k+1}2^{-3k}}{240}\cdot\frac{1}{1+\frac{1}{8}}\\ &= \frac{(-1)^{k+1}2^{-3k}}{240}\cdot\frac{1}{1+\frac{1}{8}} = \frac{(-1)^{k+1}2^{-3k}}{270}.\end{aligned}$$

We now compute $\overline{\Phi}(3t_k)-3\overline{\Phi}(2t_k)+3\overline{\Phi}(t_k)$. For this purpose, we rewrite this expression as $(\overline{\Phi}(3t_k)-\overline{\Phi}(2t_k))-2\overline{\Phi}(2t_k)+3\overline{\Phi}(t_k)$. To compute $\overline{\Phi}(3t_k)-\overline{\Phi}(2t_k)$, we observe that (57), with k replaced by $k-2$, yields

$$\begin{aligned}&\overline{\Phi}(x)-\overline{\Phi}\big(2^{1-k}\big)\\ &\quad= \frac{(-1)^{k+1}2^{2-2k}}{30}\big(x-2^{1-k}\big)+\frac{1}{2}\frac{(-1)^{k+1}2^{1-k}}{3}\big(x-2^{1-k}\big)^2+\frac{(-1)^k}{6}\big(x-2^{1-k}\big)^3.\end{aligned}$$

Substitute $x=3t_k$, so $x=3\cdot 2^{-k}=2^{1-k}+2^{-k}$, and $x-2^{1-k}=2^{-k}$. Then (since $2^{1-k}=2t_k$), we get

$$\begin{aligned}\overline{\Phi}(3t_k)-\overline{\Phi}(2t_k) &= \frac{(-1)^{k+1}2^{2-2k}}{30}\cdot 2^{-k}+\frac{1}{2}\frac{(-1)^{k+1}2^{1-k}}{3}\cdot 2^{-2k}+\frac{(-1)^k}{6}\cdot 2^{-3k}\\ &= (-1)^k 2^{-3k}\left(-\frac{4}{30}-\frac{1}{3}+\frac{1}{6}\right) = (-1)^{k+1}2^{-3k}\cdot\frac{3}{10}.\end{aligned}$$

The formula $\overline{\Phi}(2^{-k})=\frac{(-1)^{k+1}2^{-3k}}{270}$ tells us that

$$\overline{\Phi}(t_k)=\frac{(-1)^{k+1}2^{-3k}}{270}.$$

If we apply the formula with $k-1$ instead of k, we get

$$\overline{\Phi}(2t_k)=\frac{(-1)^k 2^{3-3k}}{270}=(-1)^k 2^{-3k}\cdot\frac{4}{135}.$$

Hence

$$\begin{aligned}&\big(\overline{\Phi}(3t_k)-\overline{\Phi}(2t_k)\big)-2\overline{\Phi}(2t_k)+3\overline{\Phi}(t_k)\\ &\quad= (-1)^{k+1}2^{-3k}\cdot\frac{3}{10}-2\left((-1)90k2^{-3k}\cdot\frac{4}{135}\right)+3\left(\frac{(-1)^{k+1}2^{-3k}}{270}\right)\\ &\quad= (-1)^{k+1}2^{-3k}\left(\frac{3}{10}+\frac{8}{135}+\frac{1}{90}\right) = (-1)^{k+1}2^{-3k}\cdot\frac{3\cdot 27+8\cdot 2+3}{270}\\ &\quad= (-1)^{k+1}2^{-3k}\cdot\frac{81+16+3}{270} = (-1)^{k+1}2^{-3k}\cdot\frac{100}{270} = (-1)^{k+1}\frac{10}{27}t_k^3.\end{aligned}$$

If we combine this with (56), we find

$$\begin{aligned}\overline{\mathcal{S}} &= \varphi(t_k) + \varphi\big(t_k^2\big) - \varphi\big(t_k + t_k^2\big) + \big(\overline{\Phi}(3t_k) - \overline{\Phi}(2t_k)\big) - 2\overline{\Phi}(2t_k) + 3\overline{\Phi}(t_k) \\ &= \frac{(-1)^{k+1}}{3} \cdot t_k^3 + O\big(t_k^4\big) + (-1)^{k+1}\frac{10}{27}t_k^3 = (-1)^{k+1}t_k^3\left(\frac{1}{3} + \frac{10}{27}\right) + O\big(t_k^4\big) \\ &= (-1)^{k+1}\frac{19}{27}t_k^3 + O\big(t_k^4\big).\end{aligned}$$

Hence $\overline{\mathcal{S}}$ is *not* $o(t^3)$ as $t \downarrow 0$. As explained before, this shows that $S(t,0)$ is not $o(t^3)$ as $t \downarrow 0$. Hence Fact 2 is proved.

References

[1] A. Agrachev, R. Gamkrelidze, Exponential representation of flows and chronological calculus, Math. USSR-Sb. 107 (4) (1978) 487–532 (in Russian).

[2] A. Agrachev, R. Gamkrelidze, Chronological algebras and nonstationary vector fields, J. Soviet Math. 17 (1) (1979) 1650–1675.

[3] C. Calcaterra, A. Boldt, Flow-box theorem for Lipschitz continuous vector fields, J. Math. Anal. Appl., in press, math/0305207.

[4] F.H. Clarke, Optimization and Nonsmooth Analysis, Jonh Wiley & Sons, New York, 1983.

[5] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, P.R. Wolenski, Nonsmooth Analysis and Control Theory, Grad. Texts in Math., Springer, 1998.

[6] M. Kawski, H.J. Sussmann, Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory, in: U. Helmke, D. Praetzel-Wolters, E. Zerz (Eds.), Operators, Systems and Linear Algebra: Three Decades of Algebraic Systems Theory, Kaiserslautern, 1997, Teubner, Stuttgart, 1997, pp. 111–129.

[7] F. Rampazzo, H.J. Sussmann, Set-valued differentials and a nonsmooth version of Chow's theorem, in: Proc. of the 40th IEEE Conf. on Decision and Control, Orlando, FL, December 2001, vol. 3, IEEE Publications, New York, 2001, pp. 2613–2618.

[8] S. Simić, Lipschitz distributions and Anosov flows, Proc. Amer. Math. Soc. 124 (1996) 1869–1877.