Nonlinear Differ. Equ. Appl. (2017) 24:61

1021-9722/17/060001-43
published online October 31, 2017
https://doi.org/10.1007/s00030-017-0484-4

Nonlinear Differential Equations and Applications NoDEA

Iterated Lie brackets for nonsmooth vector fields

To Alberto Bressan.

Ermal Feleqi and Franco Rampazzo

Abstract. If the vector fields f_1, f_2 are locally Lipschitz, the classical Lie bracket $[f_1, f_2]$ is defined only almost everywhere. However, it has been shown that, by means of a set-valued Lie bracket $[f_1, f_2]_{set}$ (which is defined everywhere), one can generalize classical results like the Commutativity theorem and Frobenius' theorem, as well as a Chow–Rashevski's theorem involving Lie brackets of degree 2 (we call 'degree' the number of vector fields contained in a formal bracket). As it might be expected, these results are consequences of the validity of an *asymptotic formula* similar to the one holding true in the regular case. Aiming to more advanced applications—say, a general Chow–Rashevski's theorem or higher order conditions for optimal controls—we address here the problem of defining, for any $m > 2$ and any formal bracket B of degree m, a Lie bracket $B(f_1, \dots, f_m)$ corresponding to vector fields $(f_1, \dots, f_m)$ lacking classical regularity requirements. A major complication consists in finding the right extension of the degree 2 bracket, namely a notion of bracket which admits an asymptotic formula. In fact, it is known that a mere iteration of the construction performed for the case $m = 2$ is not compatible with the validity of an asymptotic formula. We overcome this difficulty by introducing a set-valued bracket $x \mapsto B_{set}(f_1, \dots, f_m)(x)$, defined at

This article is part of the topical collection "Hyperbolic PDEs, Fluids, Transport and Applications: Dedicated to Alberto Bressan for his 60th birthday" guest edited by Fabio Ancona, Stefano Bianchini, Pierangelo Marcati, Andrea Marson.

The authors wish to thank Héctor J. Sussmann (Rutgers University, New Brunwick, NJ) for having generously inspired part of this paper. They are also quite indebted to Rohit Gupta (IMA, University of Minnesota), who helped them with meticulous revisions of the manuscript. This research is partially supported by the Gruppo Nazionale per l' Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM), Italy, and by Padova University Grant PRAT 2015 "Control of Dynamics with Reactive Constraints".

each point x as the convex hull of the set of limits along suitable d-tuples of sequences of points converging to x. The number d depends only on the formal bracket B and is here called the *diff-degree* of B. It counts the maximal order of differentiations involved in $B(g_1, \ldots, g_m)$ (for any smooth m-tuple of vector fields $(g_1, \ldots, g_m)$). The main result of the paper is an asymptotic formula valid for the bracket $B_{set}(f_1, \ldots, f_m)(x)$.

Mathematics Subject Classification. Primary 34A26, 49J52; Secondary 34H05.

Keywords. Nonsmooth vector fields, Iterated Lie brackets, Asymptotic formulas.

1. Introduction

1.1. The problem and the main result

As soon as soon as $f_1 = \sum_{i=1}^n f_1^i \dfrac{\partial}{\partial x^i}$ and $f_2 = \sum_{i=1}^n f_2^i \dfrac{\partial}{\partial x^i}$ are C^1 vector fields on an Euclidean space $\mathbb{R}^n$,[1] the Lie bracket $[f_1, f_2]$ is defined as

$$[f_1, f_2](x) := \sum_{i,j=1}^n \left(\frac{\partial f_2^i}{\partial x^j} f_1^j - \frac{\partial f_1^i}{\partial x^j} f_2^j \right) \frac{\partial}{\partial x^i} \quad \forall x \in \mathbb{R}^n.$$

The Lie bracket turns out to be an intrinsic object. Precisely, though it is constructed by a linear combination of second order differential operators, *is a continuous first order differential operator, namely a continuous vector field.*

As it is well-known, the Lie bracket $[f_1, f_2]$ gauges the noncommutativity of the flows of f_1 and f_2. Indeed, on the one hand the Commutativity Theorem states that the flows of a family $f_1, \ldots, f_m$ commute locally if and only if $[f_i, f_j] \equiv 0$, for all $i, j = 1, \ldots, k$. Furthermore, Frobenius' Theorem establishes that the involutivity condition—namely $[f_i, f_j] \in span\{f_1, \ldots, f_k\}$, $i, j = 1, \ldots, n$ (plus a constant rank hypothesis)—is equivalent to the fact that the flows of $f_1, \ldots, f_k$ give rise to local foliations of dimension k. On the other hand, if $f_1, \ldots, f_k$ are vector fields of class C^∞, one can regard the Lie bracket as a (non-commutative) product, so obtaining the Lie algebra $Lie\{f_1, \ldots, f_k\}$—over the ring of C^∞ functions–generated by $\{f_1, \ldots, f_k\}$. Chow–Rashevski's theorem states that the assumption $Lie\{f_1, \ldots, f_k\}(y) = \mathbb{R}^n \ \forall y \in \mathbb{R}^n$, called *full rank condition*—or, in the PDE's literature, Hörmander condition—guarantees small-time local controllability.[2]

The cornerstone of the mentioned results is an *asymptotic formula* connecting the bracket with a suitable composition of flows of the involved vector

[1] The definitions and results in this paper are valid on a differentiable manifold as well, but, since they involve only *local* notions and estimates it is not restrictive to consider just the case of an Euclidean space.

[2] A k-tuple $(f_1, \ldots, f_k)$ is called *small-time local controllable* from x if, for any $t > 0$, the set of points reachable at t from x via concatenations of forward or backward integral curves of the f_i, $i = 1, \ldots k$, is a neighborhood of x.

fields. Let us use the exponential notation for the flows, namely, if g is a locally Lipschitz vector field, $e^{tg}(y)$ is the solution at t of the Cauchy problem $\dot{x} = g(x)$, $x(0) = y$.

Asymptotic Formula for $[f_1, f_2]$. *For every* $x_* \in \mathbb{R}^n$ *there exists a modulus* $\gamma(\cdot)\colon [0, +\infty[\to [0, +\infty[$, *i.e.,* $\gamma(\cdot)$ *is a nondecreasing function with* $\gamma(0^+) := \lim_{\rho\to 0} \gamma(\rho) = 0$, *such that*

$$\begin{aligned}&\mathrm{dist}\Big(e^{-t_2f_2} \circ e^{-t_1f_1} \circ e^{t_2f_2} \circ e^{t_1f_1}(x) - x \,, t_1t_2[f_1, f_2](x_*)\Big)\\ &\quad \le |t_1t_2|\Big(\gamma\big(|(t_1, t_2)| + |x - x_*|\big)\Big)\end{aligned}$$

for all $t_1, t_2 \in \mathbb{R}$, $x \in \mathbb{R}^n$.

1.1.1. Lie brackets of non-regular vector fields. Notice that there is a discrepancy between the hypotheses on the vector fields needed to construct the Lie bracket and those guaranteeing the existence (and possibly the uniqueness) of local flows. Indeed the asymptotic formula holds true for C^1 vector fields f_1, f_2, while local flows exist unique provided the vector fields are just locally Lipschitz continuous. Therefore, aiming at weakening the regularity hypotheses on the vector fields, one has to address the problem of an extended notion of Lie bracket. Some steps in this direction were taken in [8][3] through the use of the following notion of *set-valued Lie bracket*: if f_1, f_2 are locally Lipschitz vector fields, the set-valued Lie bracket $[f, g]_{set}(\cdot)$ is defined by setting, for every $y \in \mathbb{R}^n$,

$$\begin{aligned}&[f_1, f_2]_{set}(y)\\ &\quad := \overline{co}\Big\{w = \lim_{k\to\infty}[f_1, f_2](y_k)\colon y_k \to \bar{y},\ y_k \in \mathit{Diff}(f_1) \cap \mathit{Diff}(f_2)\Big\}, \qquad (1.1)\end{aligned}$$

where, if h is a vector field, we use $\mathit{Diff}(h) \subset \mathbb{R}^n$ to denote the subset of points of differentiability of h. By Rademacher's theorem, $\mathit{Diff}(f) \cap \mathit{Diff}(g)$ has full measure—hence, it is dense—in $\mathbb{R}^n$. It is straightforward to show that the set-valued map $\bar{y} \mapsto [f, g]_{set}(y)$ is upper semi-continuous[4] with compact convex values. Moreover, $\{[f_1, f_2](y)\} \subseteq [f_1, f_2]_{set}(y)$, at each $y \in \mathbb{R}^n$ where f_1, f_2 are both differentiable, the equality holding true if both Df_1 and Df_2 are continuous at y. By the use of this set-valued bracket, in [7,9] the Commutativity Theorem and Frobenius' Theorem have been extended to families of locally Lipschitz continuous vector fields. For this same class of vector fields an extension of step-2 Chow–Rashevski's Theorem was proved in [8]. Let us point out that set-valued Lie brackets are likely to be useful also for second order necessary conditions to nonsmooth optimal control problems.

There are at least two crucial reasons why the notion of set-valued bracket $[\cdot, \cdot]_{set}$ works well: first, it is *sufficiently small* for proving the Commutativity Theorem and Frobenius' Theorem for the case of locally Lipschitz vector fields, as in [7,9]; secondly , it is *large enough* for applying an open mapping argument in the non-smooth extension of step-2 Chow–Rashevski's theorem [8]. As in

[3] But see also [2] and references therein.

[4] See below the definition of upper semi-continuity of a set-valued map.

the smooth case the main ingredient to prove these results is an formula, which in this case is expressed as follows:

Asymptotic Formula for $[f_1, f_2]_{set}$. *For every* $x_* \in \mathbb{R}^n$ *there exists a modulus* $\gamma(\cdot)\colon [0, +\infty[\to [0, +\infty[$, *such that*

$$\operatorname{dist}\Big(e^{-t_2 f_2} \circ e^{-t_1 f_1} \circ e^{t_2 f_2} \circ e^{t_1 f_1}(x) - x,\ t_1 t_2 [f_1, f_2]_{set}(x_*)\Big) \\ \leq |t_1 t_2| \Big(\gamma\big(|(t_1, t_2)| + |x - x_*|\big)\Big). \tag{1.2}$$

1.2. The main result

The number m of vector fields formally involved in the construction of a bracket B will be called *the degree of* B. In this case we also say that B is a *degree-*m bracket. For instance, the brackets $\big[[X_1, X_2], [X_3, X_4]\big]$, $[X_1, X_2]$, and $[[X_1, X_2], X_3]$ have degree 4, 2, and 3, respectively. In the smooth case, asymptotic formulas for bracket of degree ≥ 3 can obviously be deduced by induction. So a natural issue is: can we do the same in the "non-smooth" case? Let us consider degree-3 brackets: if f_1, f_2 are of class $C^{1,1}$—i.e., their derivatives are locally Lipschitz—then $g := [f_1, f_2]$ is a locally Lipschitz vector field, so, if f_3 is another locally Lipschitz vector field, one might be lead to use the set-valued vector field $[g, f_3]_{set}$ as a (set-valued) definition of the iterated bracket $[[f_1, f_2], f_3]_{set}$. Yet, an example built in [9] shows that an asymptotic formula for the multi-flows associated with $[g, f_3]_{set}$ *fails to be true.* Actually, this example suggests that the bracket $[g, f_3]_{set}$ is *too small.*

In the present paper we propose a generalization of the notion of (set-valued) Lie bracket to the cases when the degree is greater than 2, in such a way that an asymptotic formula holds true. Therefore, in view of the above-mentioned example, this bracket cannot simply be obtained by iteration of the construction performed for degree-2 brackets. In particular, as soon as the degree is greater than 2, a certain number of independent sequences are needed in the definition of the set-valued bracket. This number turns out to coincide with what we call *diff-degree* of the bracket under consideration. Let us anticipate that the diff-degree is strictly smaller than the degree and is not a function of the degree, although it increases with it (see the degree-4 cases below in which the diff-degree is either 3 or 2).

At this introductory level, we prefer avoiding too technical definitions and give the bracket definition and the main result only for the degrees 3 and 4. This, however, will be enough for illustrating the main features of the general.

1.2.1. The degree-3 case.

For every $x_* \in \mathbb{R}^n$, we set

$$\big[[f_1, f_2], f_3\big]_{set}(x_*) \\ := \overline{co}\Big\{w = \lim_{\substack{x_j \to x_* \\ z_j \to x_*}} Df_3(x_j) \cdot [f_1, f_2](z_j) - D[f_1, f_2](z_j) \cdot f_3(x_j),\Big\}, \tag{1.3}$$

where the sequences (x_j) and (z_j) take values in *Diff*(f_3) and *Diff*$([f_1, f_2])$, respectively, and both (x_j) and (z_j) converge to x_*.

The chief difference from the case of the classical bracket consists in the occurrence of *two* independent sequences instead of a single one, i.e. we do not assume $(x_j) = (z_j)$. For this reason, we say that the diff-degree of the formal bracket $[[X_1, X_2], X_3]$ is 2. With reference to the above-mentioned "negative" example, notice that, setting $g := [f_1, f_2]$, one has

$$[g, f_3]_{set}(x_*) \subseteq [[f_1, f_2], f_3]_{set}(x_*),$$

and the inclusion is possibly strict. Together with the asymptotic formula below, this somehow explains the negative result illustrated by the example in [9]. To state the asymptotic formula let us set

$$\begin{aligned}\psi(t_1, t_2) &:= e^{-t_2 f_2} \circ e^{-t_1 f_1} \circ e^{t_2 f_2} \circ e^{t_1 f_1}\\ \psi(t_1, t_2, t_3) &:= e^{-t_3 f_3} \circ (\psi(t_1, t_2))^{-1} \circ e^{t_3 f_3} \circ \psi(t_1, t_2).\end{aligned}$$

Asymptotic Formula for $\big[[f_1, f_2], f_3\big]_{set}$. *For every $x_* \in \mathbb{R}^n$ there exists a modulus $\gamma(\cdot)$ such that*

$$\begin{aligned}&\operatorname{dist}\Big(\psi(t_1, t_2, t_3)(x) - x_*,\ t_1 t_2 t_3 \big[[f_1, f_2],\ f_3\big]_{set}(x_*)\Big)\\ &\quad \le |t_1 t_2 t_3|\, \gamma\big(|(t_1, t_2, t_3)| + |x - x_*|\big)\end{aligned}$$

for all $(t_1, t_2, t_3) \in \mathbb{R}^3$ and $x \in \mathbb{R}^n$. Furthermore, the modulus $\gamma(\cdot)$ can be chosen to be the same for all x_* in a compact subset of $\mathbb{R}^n$.

1.2.2. Two degree-4 cases. If f_1, f_2, f_3, f_4 are vector fields of class $C^{1,1}$ (so that, in particular, the brackets $[f_1, f_2]$, $[f_3, f_4]$ are locally Lipschitz), we set

$$\begin{aligned}&\big[[f_1, f_2],\ [f_3, f_4]\big]_{set}(x_*)\\ &\quad := \overline{co}\left\{ w = \lim_{\substack{x_j \to x_*\\ z_j \to x_*}} D[f_3, f_4](x_j) \cdot [f_1, f_2](z_j) - D[f_1, f_2](z_j) \cdot [f_3, f_4](x_j) \right\},\end{aligned}$$

where the sequences (x_j) and (z_j) take values in *Diff* $([f_1, f_2])$ and *Diff* $([f_3, f_4])$, respectively, and they both converge to x_*.

On the other hand, if the vector fields f_1 and f_2 are of class $C^{2,1}$, f_3 is of class $C^{1,1}$, and f_4 is of class $C^{0,1}$, we set

$$\begin{aligned}&\big[[[f_1, f_2], f_3], f_4\big]_{set}(x_*)\\ &\quad := \overline{co}\Bigg\{ w = \lim_{\substack{x_j \to x_*\\ y_j \to x_*\\ z_j \to x_*}} Df_4(y_j) \cdot \Big(Df_3(x_j) \cdot [f_1, f_2](z_j) - D[f_1, f_2](z_j) \cdot f_3(x_j)\Big)\\ &\qquad - D\Big(Df_3(x_j) \cdot [f_1, f_2](z_j) - D[f_1, f_2](z_j) \cdot f_3(x_j)\Big) \cdot f_4(y_j)\Bigg\},\end{aligned}$$ [5]

where the sequences (x_j) , (y_i), and (z_j) take values in the (full measure) subsets $\mathit{Diff}^{(2)}(f_3)$, $\mathit{Diff}(f_4)$, and $\mathit{Diff}^{(2)}([f_1,f_2])$,[6] respectively, and they all converge to x_*.

Let us remark an important difference between the above brackets: the formal bracket $\big[[[X1,X_2],X_3],\,X_4\big]$ has diff-degree equal to 3, so in the definition of $\big[[[f_1,f_2],f_3],f_4]_{set}(x_*)$ limits are taken along sequences of *triples* converging to (x_*,x_*,x_*). Instead, the formal bracket $\big[[X1,X_2],[X_3,X_4]\big]$ has diff-degree equal to 2, and, accordingly, the definition of $\big[[f_1,f_2],\,[f_3,f_4]\big]_{set}(x_*)$ involves sequences of *pairs* converging to (x_*,x_*).

If one sets

$$\psi(t_1,t_2,t_3,t_4)(x) := \big(\psi(t_3,t_4)\big)^{-1}\circ\big(\psi(t_1,t_2)\big)^{-1}\circ\psi(t_3,t_4)\circ\psi(t_1,t_2)$$

and

$$\tilde\psi(t_1,t_2,t_3,t_4)(x) := \big(\psi(t_1,t_2,t_3)\big)^{-1}\circ e^{-t_4f_4}\circ e^{t_4f_4}\circ\psi(t_1,t_2,t_3)\,,$$

then one obtains the following result:

Asymptotic Formula for $\big[[[f_1,f_2],f_3\big],\,f_4\big]_{set}$. *For every $x_*\in\mathbb{R}^n$ there exists a modulus $\gamma(\cdot)$ such that*

$$\begin{aligned}&\mathrm{dist}\Big(\psi(t_1,t_2,t_3,t_4)(x)-x_*,\; t_1t_2t_3t_4\big[[[f_1,f_2],f_3],f_4]_{set}(x_*)\Big)\\&\quad\le |t_1t_2t_3t_4|\Big(\gamma(|(t_1,t_2,t_3t_4)|+|x-x_*|)\Big)\end{aligned}$$

for all $(t_1,t_2,t_3,t_4)\in\mathbb{R}^4$ and $x\in\mathbb{R}^n$. Furthermore, the modulus $\gamma(\cdot)$ can be chosen to be the same for all x_* in a compact subset of $\mathbb{R}^n$.

An akin result holds for the bracket $\big[[f_1,f_2],\,[f_3,f_4]\big]_{set}$ in relation to the flow $\tilde\psi$.

Asymptotic Formula for $\big[[f_1,f_2],\,[f_3,f_4]\big]_{set}$. *For every $x_*\in\mathbb{R}^n$ there exists a modulus $\gamma(\cdot)$ such that*

$$\begin{aligned}&\mathrm{dist}\Big(\tilde\psi(t_1,t_2,t_3,t_4)(x)-x,\; t_1t_2t_3t_4\big[[f_1,f_2],\,[f_3,f_4]\big]_{set}(x_*)\Big)\\&\quad\le |t_1t_2t_3t_4|\Big(\gamma(|(t_1,t_2,t_3t_4)|+|x-x_*|)\Big)\end{aligned}$$

for all $(t_1,t_2,t_3,t_4)\in\mathbb{R}^4$ and $x\in\mathbb{R}^n$.

---

[5] The expression

$$D\Big(Df_3(x_j)\cdot[f_1,f_2](z_j)-D[f_1,f_2](z_j)\cdot f_3(x_j)\Big)\cdot f_4(y_j)\Big)$$

here denotes the quantity

$$\begin{aligned}&D^2f_3(x_j)\cdot[f_1,f_2](z_j)-D^2[f_1,f_2](z_j)\cdot f_3(x_j)+Df_3(x_j)\cdot D[f_1,f_2](z_j)\\&\quad -D[f_1,f_2](z_j)\cdot Df_3(x_j)\,.\end{aligned}$$

[6] $\mathit{Diff}^{(2)}(f_3)$, and $\mathit{Diff}^{(2)}([f_1,f_2])$ denote the set of points where f_3 and $[f_1,f_2]$ are twice differentiable, respectively.

1.3. Organization of the paper

Sections 1.4 and 1.5 are concerned with basic material which will be used throughout the paper. In Sect. 2 we introduce formal Lie brackets of degree $m \geq 1$. In particular, given a bracket B, we define the family of *basic subbrackets* of B, whose cardinality is here called the *diff-degree* of B. Moreover, we consider notions of B-*regularity* for m-tuples of vector fields. Section 3 contains the definition of *set-valued iterated bracket*, which is obtained as a suitable limit of *twisted Lie brackets*. In the main result of the paper, namely the Asymptotic Formula (Theorem 3.7), the set-valued iterated bracket is shown to be related with the infinitesimal behavior of the corresponding multi-flows. In Sect. 4 we give the proof of the Asymptotic Formula. In turn, this proof is based on both the *Exact Integral Formula* for iterated Lie brackets, which is recalled in Sect. 5, and the technical estimates proved in Sect. 6. The latter mainly concern the approximation of twisted Lie brackets by certain *shift-twisted* brackets and, moreover, an error bound for the difference between set-valued brackets and iterated brackets of regularized vector fields.

1.4. Preliminary definitions and notation

Let us use $\mathbb{N}_0$ to denote the set of nonnegative integers, i.e, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

1.4.1. Set-valued maps and vector fields.

Definition 1.1. By a *set-valued map* $F\colon M \rightsquigarrow N$ from a set M to another set N we mean any map from M to $\mathcal{P}(N)$, the family of the subsets of N. If M, N are topological spaces, one says that F is *upper semi-continuous* at a point $x_* \in M$ if, for every neighborhood V of the subset $F(x_*)$ there exists a neighborhood U of x_* such that

$$F(U) := \bigcup_{x \in U} F(x) \subset V.$$

One says that F is *upper semi-continuous* if F is upper semi-continuous at x_* for every $x_* \in M$.

In the case when M, N are metric spaces, we have (see e.g. [1]):

Proposition 1.2. *Assume that M, N are metric spaces and F has closed values (i.e., $F(x)$ is a closed set in N for every $x \in M$). Also, let the graph $gr(F) := \{(x, y)\colon x \in M\ y \in F(x)\}$ be locally compact.*

Then F is upper semi-continuous at x_* if and only if, for every sequence $(x_k)_{k\in\mathbb{N}}$ in M such that $(x_k, y_k) \to (x_*, y)$ for some $y \in N$, with $y_k \in F(x_k) (\subseteq N)$, one has $y \in F(x_*)$. That is, to say, F is upper semi-continuous if and only if $gr(F)$ is closed.

1.4.2. Spaces of vector fields. We recall that a vector field f is said to be *of class C^0* if it is continuous. Furthermore, for $k \geq 1$, a vector field f is said to be *of class C^k* if the derivatives $D^j f$ are continuous for every $j = 0, \dots, k$, while f is said to be *of class $C^{k-1,1}$* if it is of class C^{k-1} and $D^{k-1} f$ is locally Lipschitz continuous.[7] As it is customary, for every integer $k \geq 0$, we shall use C_0^k [resp. $C_0^{k-1,1}$, if $k > 1$] to denote the subset of C^k [resp. $C^{k-1,1}$] whose elements have compact support. If f is of class $C^{k-1,1}$, by Rademacher's theorem, the derivative $D^k f$ exists almost everywhere and is locally bounded. If $f \in C_0^{k-1,1}$ we shall consider the usual norm

$$\|f\|_{k-1,1} = \sum_{j=0}^{k} \|D^j f\|_\infty,$$

$\|\cdot\|_\infty$ denoting the L^∞ norm. Clearly, as soon as $f \in C^k$,

$$\|f\|_{k-1,1} = \|f\|_k := \sum_{j=0}^{k} \|D^j f\|_\infty.$$

1.5. Faà di Bruno's formula

We will make use of the standard multi-index notation, according to which, for a function $\phi\colon \mathbb{R}^n \to \mathbb{R}^q$ and multi-indexes $\alpha, \beta \in (\mathbb{N}_0)^n$ one sets

$$\partial^\alpha \phi := \frac{\partial^{\alpha_1}}{\partial x^1} \cdot \dots \cdot \frac{\partial^{\alpha_n}}{\partial x^n} \phi,$$

where $\beta \leq \alpha$ means $\beta_j \leq \alpha_j$ for every $j = 1, \dots, n$, and

$$|\alpha| := \alpha_1 + \dots + \alpha_n, \qquad \alpha - \beta := (\alpha_1 - \beta_1, \dots, \alpha_n - \beta_n),$$
$$\binom{\alpha}{\beta} := \binom{\alpha_1}{\beta_1} \cdot \binom{\alpha_2}{\beta_2} \cdots \binom{\alpha_n}{\beta_n}.$$

We recall the identity

$$\sum_{|\alpha|=k} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} = 2^k \cdot \binom{k+n-1}{k}, \tag{1.4}$$

which can be found in [6, Problem 2.58].

Let us remind Faà di Bruno's chain rule for higher-order derivatives (see, e.g., [6] for the single-variable case and [3,5] for the many-variable case). We now need some additional notation. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $\alpha = (\alpha_1, \dots, \alpha_n)$, x^α stands for $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$. We introduce an order relation on the set of multi-indices $\mathbb{N}_0^n$. For $\alpha, \beta \in \mathbb{N}_0^n$ we write $\beta \prec \alpha$ if one of the following holds:

1. $|\beta| < |\alpha|$; or
2. $|\beta| = |\alpha|$, $\beta_1 < \alpha_1$; or
3. $|\beta| = |\alpha|$ and there is $i \in \{1, \dots, n-1\}$ such that $\beta_1 = \alpha_1, \dots, \beta_i = \alpha_i$ and $\beta_{i+1} < \alpha_{i+1}$.

---

[7] This definition is *intrinsic*, i.e., the notion of class C^k and of class $C^{k-1,1}$ are not affected by local coordinate changes of classes C^{k+1} and $C^{k,1}$ respectively. In particular, this notion can be equally given on a manifold of class C^{k+1} and $C^{k,1}$, respectively.

Let $\Omega \subset \mathbb{R}^n$ and $D \subset \mathbb{R}^p$ be open sets and let $g\colon \Omega \to \mathbb{R}^p$, $f\colon D \to \mathbb{R}^q$ be k times differentiable functions such that $g(\Omega) \subset D$, where $n, p, q, k \in \mathbb{N}$. Then $f \circ g$ is also k times differentiable and for all $\alpha \in \mathbb{N}_0^n$ (where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$) with $1 \le |\alpha| \le k$

$$\partial^\alpha (f \circ g) = \sum_{|\gamma|=1}^{|\alpha|} (\partial^\gamma f) \circ g \sum_{s=1}^{|\alpha|} \sum_{P_s(\alpha,\gamma)} \alpha! \prod_{j=1}^{s} \frac{(\partial^{\beta^j} g)^{\gamma^j}}{\gamma^j!(\beta^j!)^{|\gamma^j|}}, \tag{1.5}$$

where, for all $s = 1, \ldots, |\alpha|$, $\gamma \in \mathbb{N}_0^p$ with $1 \le |\gamma| \le |\alpha|$, we let

$$P_s(\alpha,\gamma) := \Big\{ (\beta^1, \ldots, \beta^s; \gamma^1, \ldots, \gamma^s) \in (\mathbb{N}_0^n \times \mathbb{N}_0^p)^s \ : \ |\gamma^j| \ge 1, \ \forall j = 1, \ldots, s, \\ \mathbf{0} \prec \beta^1 \prec \cdots \prec \beta^s, \ \sum_{j=1}^{s} \gamma^j = \gamma, \ \sum_{j=1}^{s} |\gamma^j| \beta^j = \alpha \Big\}. \tag{1.6}$$

In the above expression, β^j, like α, is an n-dimensional multi-index. Similarly, γ^j, like γ, is a p-dimensional multi-index.

2. Formal brackets and bracket-regularity of vector fields

2.1. Formal brackets

Given a fixed sequence $\mathbf{X} = (X_1, X_2, \ldots)$ of distinct objects called *variables*, or *indeterminates*, we call *words* the finite ordered strings consisting of the X_i, the *left bracket* [and the *right bracket*],[8] and the comma. We shall use $W(\mathbf{X})$ to denote the set of words. For instance, $X_2X_5X_4$ and $X_3, [X_{13}[,]]X_{61}[$ are words (but we will be mainly concerned with special words, the *iterated brackets*—see Definition 2.1—like $[[X_3, X_4], X_5]$ or $[[[X_4, X_6], X_7], [X_8, X_9]]$).

Given any word $W \in W(\mathbf{X})$, we use $Seq(W)$ to denote the word obtained from W by deleting all left and right brackets and all commas. The **degree** $deg(W)$ of a word $W \in W(\mathbf{X})$ is the cardinality of $Seq(W)$. For instance, if $W = [[[X_4, X_6], X_7], [X_8, X_9]]$, $Seq(W) = X_4X_6X_7X_8X_9$ and $deg(W) = 5$.

Let us give the definition of formal bracket by iteration.

Definition 2.1. We will call **formal bracket of degree** 1 any word of degree one and we will say that the *bracket of two members* W_1, W_2 of $W(\mathbf{X})$ is the word $[W_1, W_2]$ obtained by writing first a left bracket, then W_1, then a comma, then W_2, and then a right bracket.

We call **formal iterated brackets** (or, simply, *brackets*) of $\mathbf{X}$ the elements of the smallest subset $IB(\mathbf{X}) \subset W(\mathbf{X})$ such that:

1. $IB(\mathbf{X})$ contains the brackets of degree 1;
2. whenever W_1 and W_2 belong to $IB(\mathbf{X})$ it follows that $[W_1, W_2] \in IB(\mathbf{X})$;

[8] In this passage the terms "left bracket" and "right bracket" refer to the square parentheses] and [, respectively. They should not be confused with the notion of *formal bracket*—or simply *bracket*—which is introduced below as a special kind of word and is modeled on the Lie bracket.

3. For every element $B \in IB(\mathbf{X})$ one has $seq(B) = X_{\mu+1}, \ldots, X_{\mu+m}$, for some $\mu \geq 0$ and $m > 0$.[9]

Notice that the degree is additive by bracketing operation, i.e.,

$$deg([B_1, B_2]) = deg(B_1) + deg(B_2)$$

for every pair of brackets B_1, B_2.

For every bracket B of degree $m > 1$ there exists a unique pair (B_1, B_2) of brackets such that $B = [B_1, B_2]$. The pair (B_1, B_2) is the *factorization* of B, and the brackets B_1, B_2 are known, respectively, as the *left factor* and the *right factor* of B.

If B is an iterated bracket of degree m, then any substring of B which is itself an iterated bracket is called a *subbracket* of B. We will use $Sbb(B)$ to denote the set of subbrackets of B. Brackets have a nested structure: if S_1, S_2 are subbrackets of B, then either S_1 is a subbracket of S_2, or S_2 is a subbracket of S_1, or S_1 and S_2 are *disjoint* (in the sense that their letter sequences are disjoint).

2.2. Basic brackets and diff-degree

Special subbrackets of degree ≤ 2, here called *basic subbrackets*, will be crucial in establishing the kind of limiting procedure needed for defining set-valued iterated Lie brackets (Sect. 3). The cardinality of the family of basic subbrackets of a bracket B will be called the *diff-degree of* B.

Definition 2.2. Let B be a bracket. A subbracket $S \subset Sbb(B)$ is a **basic subbracket of** B if either $deg(S) = 2$ or $deg(S) = 1$, i.e., $S = X_i$ for some i, and neither $[X_{i-1}, X_i]$ nor $[X_i, X_{i+1}]$ is a subbracket of B. We will call **diff-degree of** B, and write $\mathfrak{Deg}(B)$ the number of basic subbracket of B.

For instance, the subbrackets $[X_2, X_3]$ and X_4 of $[[X_2, X_3], X_4]$ are basic, while X_2 and X_3 are not, so that $\mathfrak{Deg}([[X_2, X_3], X_4]) = 2$. Observe that

$$\mathfrak{Deg}(B) = 1 \iff deg(B) \leq 2.$$

Notice that, for every bracket such that $deg(B) \geq 2$, one has

$$\mathfrak{Deg}(B) \leq deg(B) - 1.$$

For instance,

$$2 = \mathfrak{Deg}([[X_3, X_4], [X_5, X_6]]) = deg([[X_3, X_4], [X_5, X_6]]) - 2$$

and

$$3 = \mathfrak{Deg}([[[X_3, X_4], X_5], X_6]) = deg([[[X_3, X_4], X_5], X_6]) - 1.$$

Notice also that, as soon as $deg(B) > 2$, the diff-degree is additive with respect to factorization: if $B = [B_1, B_2]$, then

$$\mathfrak{Deg}(B) = \mathfrak{Deg}(B_1) + \mathfrak{Deg}(B_2). \tag{2.1}$$

---

[9] We might eliminate the third condition, which, up to renaming variables, is not restrictive. However, it is convenient to retain it for practical reasons connected with factorizations and proofs based on recursion.

The name *diff-degree* refers to the fact that it coincides with the highest order of differentiation showing up in any coordinate representation when the variables of B are replaced by vector fields. Let us see this more precisely. If we plug in smooth vector fields f_j for the indeterminates X_j of a bracket B and interpret the result as a classical Lie bracket—see Definition 2.8—every repeated bracketing of a subbracket involves a certain number of differentiations, and this number depends only on the formal bracket B.

Definition 2.3. If B is a bracket and $S \in Sbb(B)$, let us define $\mathfrak{Deg}(S;B)$ by a backward recursion on S:

$$\mathfrak{Deg}(B;B) := 0, \quad \mathfrak{Deg}(S_1;B) := \mathfrak{Deg}(S_2;B) := 1 + \mathfrak{Deg}([S_1,S_2];B).$$

We shall refer to $\mathfrak{Deg}(S;B)$ as the *the number of differentiations of S in B.*

It is easy to prove that $\mathfrak{Deg}(S;B)$ is equal to the number of right brackets that occur in B to the right of S minus the number of left brackets that occur in B to the right of S. For example, if $B = \big[X_3, [X_4, X_5]\big]$, then $\mathfrak{Deg}([X_4,X_5];B) = 1$, $\mathfrak{Deg}(X_4;B) = 2$, $\mathfrak{Deg}(X_5;B) = 2$, $\mathfrak{Deg}(X_3;B) = 1$.

Clearly, if (B_1, B_2) is the factorization of B, and $seq(B_1) = X_{\mu+1}, \dots, X_{\mu+m_1}$, $seq(B_2) = X_{\mu+m_1+1}, \dots, X_{\mu+m_2}$ for some $\mu \geq 0$, $m_1, m_2 \geq 1$,

$$\mathfrak{Deg}(X_j;B) = \begin{cases} \mathfrak{Deg}(X_j;B_1) + 1 & \text{if } j \in \{\mu+1,\dots,\mu+m_1\} \\ \mathfrak{Deg}(X_j;B_2) + 1 & \text{if } j \in \{\mu+m_1+1,\dots,\mu+m_1+m_2\}. \end{cases}$$

Moreover, if X_j is a subbracket of S and S is a subbracket B, one has

$$\mathfrak{Deg}(X_j;B) = \mathfrak{Deg}(X_j;S) + \mathfrak{Deg}(S;B).$$

The relation between the numbers $\mathfrak{Deg}(X_j;B)$ and the diff-degree $\mathfrak{Deg}(B)$ is as follows:

Lemma 2.4. *For any bracket B one has*

$$\mathfrak{Deg}(B) = \max\Big\{\mathfrak{Deg}(X_j;B)\colon X_j \text{ is an indeterminate of } B\Big\}.$$

Remark 2.5. The diff-degree $\mathfrak{Deg}(B)$ will be the number of (n-dimensional) variables that have to be used in the limiting procedure for defining Lie brackets of vector fields which satisfy weaker regularity hypotheses (see Sect. 3.1). So $n \cdot \mathfrak{Deg}(B)$ might be regarded as the *right dimension* of the space where one can define the new bracket by a density approach. One might say that $\mathfrak{Deg}(B)$ reveals a dimension of the bracket which is *hidden* in the ordinary, i.e., the smooth case.

Finally, let us point out that the diff-degree of a bracket should not be confused with the order of the corresponding Lie bracket as a differential operator. Actually, all iterated Lie brackets of smooth vector fields are vector fields, namely differential operators of the *first order.*

2.3. Bracket-regularity for families of vector fields

To each bracket B with $seq(B) = X_{\mu+1} \dots X_{\mu+m}$, ($\mu \geq 0, m \geq 1$), and each finite sequence $\mathbf{f} = (f_1, \dots, f_\nu)$ of vector fields such that $\nu \geq m + \mu$, we want to associate the expression $B(\mathbf{f})$ obtained by replacing indeterminates with vector fields. For vector fields sufficiently regular, $B(\mathbf{f})$ is obviously the usual iterated Lie bracket. Let us make the required regularity conditions precise:

Definition 2.6. (*Classes* C^{B+k} *and* $C^{B+k-1,1}$) Let B be a bracket of degree $m(\geq 1)$, with $seq(B) = X_{\mu+1} \dots X_{\mu+m}$, ($\mu \geq 0$). Let $\mathbf{f} = (f_1, \dots, f_\nu)$ be a finite sequence of vector fields, with $\nu \geq \mu + m$, and let k be a nonnegative integer. We say that:

- $\mathbf{f}$ *is of class* C^{B+k} if f_j is of class $C^{\mathfrak{Deg}(X_j,B)+k}$ for each $j \in \{1+\mu, \dots, m+\mu\}$;
- $\mathbf{f}$ *is of class* $C^{B+k-1,1}$ if f_j is of class $C^{\mathfrak{Deg}(X_j,B)+k-1,1}$ for each $j \in \{1 + \mu, \dots, m + \mu\}$.

We also write $\mathbf{f} \in C^{B+k}$ and $\mathbf{f} \in C^{B+k-1,1}$, to indicate, respectively, that $\mathbf{f}$ is of class C^{B+k} and $\mathbf{f}$ is of class $C^{B+k-1,1}$.

For example, if $B = \Big[[X_3, X_4], [[X_5, X_6], X_7]\Big]$ and $\mathbf{f} = (f_1, \dots, f_8)$ (so $m = 5$, $\nu = 8$, $\mu = 2$), then $\mathbf{f} \in C^{B+3}$ if and only if $f_3, f_4, f_7 \in C^5$ and $f_5, f_6 \in C^6$, while $\mathbf{f} \in C^{B-1,1}$ if and only if $f_3, f_4, f_7 \in C^{4,1}$, and $f_5, f_5 \in C^{6,1}$. It is easy to verify the following result:

Proposition 2.7. *Assume that we are given* B, k, *and* $\mathbf{f} = (f_1, \dots, f_\nu)$ *as in Definition* 2.6. *Let* (B_1, B_2) *be the factorization of* B. *Then:*

- $\mathbf{f} \in C^{B+k}$ *if and only if* $\mathbf{f} \in C^{B_1+k+1}$ *and* $\mathbf{f} \in C^{B_2+k+1}$;
- $\mathbf{f} \in C^{B+k-1,1}$ *if and only if* $\mathbf{f} \in C^{B_1+k,1}$ *and* $\mathbf{f} \in C^{B_2+k,1}$.

We are now ready to plug vector fields in place of indeterminates in a bracket:

Definition 2.8. For integers $\mu \geq 0$, $m, \nu \geq 1$, such that $\mu + m \leq \nu$, let B be a formal bracket such that $Seq(B) = X_{\mu+1} \dots X_{\mu+m}$ and let $\mathbf{f} = (f_1, \dots, f_\nu)$ be a ν-tuple of continuous vector fields.

- If $S \in Sbb(B)$ has degree 1, i.e., $S = X_j$ for some $j = \mu + 1, \dots, \mu + m$, we define the vector field $S(\mathbf{f})$ by setting
$$S(\mathbf{f}) := X_j(\mathbf{f}) := f_j\,;$$
- If S has degree > 1, so that $S = [S_1, S_2]$, and either $S \neq B$[10] or $S = B$ and $\mathbf{f} \in C^B$, we set
$$S(\mathbf{f}) := [S_1(\mathbf{f}), S_2(\mathbf{f})].$$

Remark 2.9. There is a slight (and not confusing) abuse of the notation $[\cdot, \cdot]$ in the above definition and throughout the whole paper. Indeed, while $[S_1(\mathbf{f}), S_2(\mathbf{f})]$ is a Lie bracket of vector fields, $[S_1, S_2]$ is just a formal bracket.

[10] If S is a proper subbracket of B (i.e., $S \in Sbb(B)\backslash\{B\}$) then $\mathfrak{Deg}(S; B) > 0$, so $\mathfrak{Deg}(S; B) + k - 1 \geq 0$ even if $k = 0$.

The resulting vector field $S(\mathbf{f})$ has the following regularity:

$$S(\mathbf{f}) \in C^{\mathfrak{Deg}(S;B)+k} \qquad \text{if } \mathbf{f} \in C^{B+k} \quad ;$$
$$S(\mathbf{f}) \in C^{\mathfrak{Deg}(S;B)+k-1,1} \quad \text{if } \mathbf{f} \in C^{B+k-1,1}.$$

If S is a proper subbracket of B (i.e., $S \in Sbb(B)\backslash\{B\}$) then $\mathfrak{Deg}(S;B) > 0$, so $\mathfrak{Deg}(S;B) + k - 1 \geq 0$ even if $k = 0$. However, the inductive construction of $S(\mathbf{f})$ can be pursued even for $S = B$ as soon as $\mathbf{f} \in C^B$, in which case we obviously set

$$B(\mathbf{f})(x) = [B_1(\mathbf{f}), B_2(\mathbf{f})](x), \tag{2.2}$$

(B_1, B_2) being the factorization of B. On the other hand, *when* $\mathbf{f}$ *is just of class* $C^{B-1,1}$, we still adopt definition (2.2) for $B(\mathbf{f})$ *at the points where both* $B_1(\mathbf{f})$ *and* $B_2(\mathbf{f})$ *are differentiable, i.e., almost everywhere.*

Let us collect the results we already have, based on classical, single-valued, Lie brackets:

Proposition 2.10. *Assume that we are given data B, k, ν, and $\mathbf{f} = (f_1, \dots, f_\nu)$ as in Definition 2.6. Then:*

- *If $\mathbf{f} \in C^{B+k}$, then $B(\mathbf{f})$ is a vector field of class C^k;*
- *If $\mathbf{f} \in C^{B+k-1,1}$ and $k \geq 1$ then $B(\mathbf{f})$ is a vector field of class $C^{k-1,1}$;*
- *If $\mathbf{f} \in C^{B+k-1,1}$ and $k = 0$ then $B(\mathbf{f})$ is bounded measurable vector field (defined almost everywhere).*

On ν-tuples of $\mathbf{f} \in C^{B+k-1,1}$ having compact support, we will consider the norms:

$$\|\mathbf{f}\|_{B+k} := \|\mathbf{f}\|_{B+k-1,1} := \sum_{S \in Sbb(B)} \|S(\mathbf{f})\|_{\mathfrak{Deg}(S,B)+k}.$$ [11]

Set-valued brackets If $\mathbf{f} = (f_1, \dots, f_\nu)$ is of class $C^{B-1,1}$ and $B = [B_1, B_2]$, so that $B_1(\mathbf{f})$ and $B_2(\mathbf{f})$ are locally Lipschitz vector fields, one might expect that the correct defn of "set-valued bracket" at a point x would be obtained by

(a) *Observing that $B(\mathbf{f})(x)$ is defined almost everywhere (by Rademacher's theorem);*
(b) *Taking the limits of $B(\mathbf{f})(x+h)$ as h goes to* 0;
(c) *Defining the bracket to be the convex hull of the set of these limits.*

However, such a definition, while proving quite effective for brackets of degree 2 (see the Introduction and [7–9]), would encounter a serious objection: *the asymptotic formula for the multi-flows corresponding to brackets of degree* > 2 *(see Definition 3.6 and Theorem 3.7) would fail to be valid, as it is illustrated by an example in* [9, Section 7].

[11] Of course, we set

$$\|B(\mathbf{f})\|_{\mathfrak{Deg}(B,B)+k} = \|B(\mathbf{f})\|_k := \sum_{j=0}^{k-1} \|D^j(|B(\mathbf{f}))\|_0 + \|D^k(B(\mathbf{f}))\|_{L^\infty}.$$

In particular, the limits obtained by the above procedure are *not enough*: we will see that step (b) should be extended to suitable limiting processes performed along k-tuples of sequences, where k is the diff-degree of B. Let us make this precise by first introducing the notion of *twisting of a bracket*.

2.4. Twistings, shift-twistings, basic twistings

Let $DIFF^\infty(\mathbb{R}^n, \mathbb{R}^n)$ and $Id_{\mathbb{R}^n}$ denote the family of C^∞ diffeomorphisms of $\mathbb{R}^n$ and the identity map on $\mathbb{R}^n$, respectively.

Definition 2.11. Let B be a bracket of degree m. A **twisting** θ **of** B is any map

$$\theta\colon Sbb(B) \to DIFF^\infty(\mathbb{R}^n, \mathbb{R}^n)$$

such that $\theta(X_j) = Id_{\mathbb{R}^n}$ if X_j is not a basic subbracket of B.

Let us define the **identity twisting** as the twisting $\iota\colon Sbb(B) \to DIFF^\infty(\mathbb{R}^n, \mathbb{R}^n)$ such that

$$\iota(S) = Id_{\mathbb{R}^n}$$

for every subbracket S of B.

Definition 2.12. For any $h \in \mathbb{R}^n$, let us define the **shift** $\tau_h : \mathbb{R}^n \to \mathbb{R}^n$ by setting

$$\tau_h(x) = x + h \quad \text{forall } x \in \mathbb{R}^n.$$

A twisting Θ of B is called a **shift-twisting** if, for any subbracket S of B there exists $h \in \mathbb{R}^n$ such that

$$\Theta(S) = \tau_h.$$

In particular, by taking $h = 0$ one obtains

$$\iota(S) = \tau_0(S) = Id_{\mathbb{R}^n}$$

for every subbracket S of B.

In the proof of the asymptotic formula for multi-flows we will make use of particular shift-twistings one builds by considering, for a chosen $x \in \mathbb{R}^n$, the increments of a particular twisting θ at x. Precisely:

Definition 2.13. For any bracket B, any twisting θ of B, and any $x \in \mathbb{R}^n$, we consider the shift-twisting θ_x defined as follows:

1. If $S = B$, we set $\theta_x(B) := \tau_{\left(\theta(B)(x)-x\right)}$. Namely,

$$\theta_x(B)(y) := \tau_{\left(\theta(B)(x)-x\right)}(y) = y + \big(\theta(B)(x) - x\big) \quad \text{for all } y \in \mathbb{R}^n.$$

2. Let S be a subbracket of B with factorization (S_1, S_2), namely $S = [S_1, S_2]$. We set

$$\theta_x(S_i) := \tau_{\left(\theta(S_i)\circ\theta(S)(x)-x\right)}, \quad i = 1, 2.$$

The shift-twisting θ_x will be called the **shift-twisting approximation of** θ **based on** x.

The asymptotic formula is concerned with an estimate holding true in a neighborhood of a given point x_*. Therefore in what follows we shall need a notion of *convergence* of twistings, which we define by introducing a metric structure on twistings as follows.

Let $A \subset \mathbb{R}^n$ be a bounded set. We set for every θ, η twistings of B and every integer $k \geq 0$,

$$\mathbf{d}_{B+k}(\theta,\, \eta) := \sum_{S \in Sbb(B)} \Big(\|\theta(S) - \eta(S)\|_{\mathfrak{Deg}(S,B)+k,\, A} \\ + \|(\theta(S))^{-1} - (\eta(S))^{-1}\|_{\mathfrak{Deg}(S,B)+k,\, \theta(S)(A) \cup \eta(S)(A)} \Big) \; (< +\infty) \tag{2.3}$$

where, for every integer $j \geq 0$ and every C^∞ map $\Phi \colon \mathbb{R}^n \to \mathbb{R}^n$, we have used the notation

$$\|\Phi\|_{j,A} := \sum_{\ell=0}^{j} \|D^\ell \Phi\|_{\infty,A} \; (< +\infty),$$

where

$$\|D^\ell \Phi\|_{\infty,A} = \sup_{x \in A} |D^\ell \Phi(x)|.$$

Remark 2.14. Clearly $\mathbf{d}_{B+k}$ depends on the subset A, but we are going to keep A fixed in the sequel. Typically, given a point x_*, we shall think of A as the ball of unit radius centered at x_*. Let us point out that $\mathbf{d}_{B+k}$ enjoys the properties of a pseudometric. However, if we introduce an equivalence relation on twistings by declaring *equivalent* two twistings that coincide on A, then $\mathbf{d}_{B+k}$ is a distance on the quotient set.

Definition 2.15. We will say that a sequence $(\theta^\ell)_{\ell \in \mathbb{N}}$ of twistings of B C^{B+k}-*converges* (on a given bounded set A) to a twisting θ if

$$\mathbf{d}_{B+k}\left(\theta^\ell,\, \theta\right) \to 0$$

as $\ell \to \infty$.

Since diffeomorphisms act on vectors through their derivatives, we will also make use of the notion of the *derivative* of a given twisting θ of a bracket B:

Definition 2.16. Let θ be a twisting of a bracket B. The **derivative** of θ, which is here denoted by $D\theta$, is the function on $Sbb(B) \times \mathbb{R}^n$ defined by

$$D\theta(S)(x) := D(\theta(S))(x)$$

for each subbracket S of B and each $x \in \mathbb{R}^n$. In particular, $D\theta(S)(x)$ is a linear isomorphism of $\mathbb{R}^n$.

Notice that if Θ is a shift-twisting, one has

$$D\Theta(S)(x) := Id_{\mathbb{R}^n}$$

for any subbracket S of B and any $x \in \mathbb{R}^n$.

We will make use also of a notion of *convergence for derivatives of twistings*, by setting, for every pair of twistings θ, η and every non-negative integer k,

$$\begin{aligned}\widehat{\mathbf{d}}_{B+k}(D\theta,\, D\eta) := \sum_{S\in Sbb(B)} \Big(& \|D(\theta(S)) - D(\eta(S))\|_{\mathfrak{Deg}(S,B)+k,\, K} \\ & + \|D(\theta^{-1}(S)) - D(\eta^{-1}(S))\|_{\mathfrak{Deg}(S,B)+k,\, \theta(S)(K)\cup\eta(S)(K)}\Big) \\ & (< +\infty). \end{aligned} \tag{2.4}$$

Definition 2.17. Let $(\theta^\ell)_{\ell\in\mathbb{N}}$ be a sequence of twistings of a given bracket B. We say that the sequence $(D\theta^\ell)_{\ell\in\mathbb{N}}$ C^{B+k}*-converges* to the derivative $D\theta$ of a twisting θ if

$$\widehat{\mathbf{d}}_{B+k}\left(D\theta^\ell,\, D\theta\right) \to 0$$

as $\ell \to \infty$.

We will be mostly interested in sequences $(\theta^\ell)_{\ell\in\mathbb{N}}$ of twistings of B such that, for some $k \geq 0$,

$$\mathbf{d}_{B+k}(\theta^\ell, \iota) \to 0 \quad \text{and} \quad \widehat{\mathbf{d}}_{B+k}(D\theta^\ell, D\iota) \to 0,$$

as ℓ tends to infinity.

Remark 2.18. For any shift-twisting Θ of B and any integer $k \geq 0$, one has

$$\mathbf{d}_B(\Theta, \iota) = \mathbf{d}_{B+k}(\Theta, \iota) = \sum_{S\in Sbb(B)} |\Theta(S)(0)|,$$

$$\widehat{\mathbf{d}}_B(D\Theta, \iota) = \widehat{\mathbf{d}}_{B+k}(D\Theta, \iota) = 0.$$

2.5. Twisted Lie brackets of vector fields

Let us recall the notion of *action of a diffeomorphism on a* C^1 *vector field*:

Definition 2.19. (*Action of diffeomorphisms on vector fields*) Given a vector field $f\colon \mathbb{R}^n \to \mathbb{R}^n$ and $\Psi \in DIFF(\mathbb{R}^n, \mathbb{R}^n$, we define $\Psi^\# f$, the *action of* Ψ *on* f, to be the vector field given by

$$x \mapsto \Psi^\# f(x) := D\Psi^{-1}(\Psi(x)) \cdot (f \circ \Psi(x)) \quad \text{for all} \quad x \in \mathbb{R}^n.$$

Remark 2.20. Notice that a vector field f is defined almost everywhere if and only if $\Psi^\# f$ is defined almost everywhere.

As it is well-known, the above defined action preserves the Lie bracket:

$$\Psi^\#[f, g](x) = [\Psi^\# f, \Psi^\# g](x) \tag{2.5}$$

for all vector fields f, g of class C^1. Of course this extends to less regular vector fields, provided one considers only the points x such that both sides of (2.5) are classically defined.

Remark 2.21. If Ψ is a shift, i.e., $\Psi = \tau_h$ for some $h \in \mathbb{R}^n$, then $D\Psi(x) = Id_{\mathbb{R}^n}$, for all $x \in \mathbb{R}^n$, so that

$$(\Psi^\# f)(x) = f \circ \tau_h(x) = f(x + h) \quad \text{for all } x \in \mathbb{R}.$$

We may deform, or *twist*, a Lie bracket $B(\mathbf{f})$ as follows:

Definition 2.22. (*Twisted Lie bracket*) Let $m \geq 1$ and let B be a bracket of degree m. Let $\mu \geq 0$ be an integer such that $Seq(B) = X_{\mu+1} \dots X_{\mu+m}$, and consider a ν-tuple $\mathbf{f} = (f_1, \dots, f_\nu)$, $\nu \geq m + \mu$of vector fields of class $C^{B-1,1}$. For every $x \in \mathbb{R}^n$ and every twisting θ, let us define the **θ-twisted Lie bracket** $B^\theta(\mathbf{f})(x)$ through recursion on the degree m:

(i) if $m = 1$, we set

$$B^\theta(\mathbf{f})(x) := D(\theta(B)^{-1}) \cdot B(\mathbf{f})(\theta(B)(x)) = \theta(B)^{\#} B(\mathbf{f})(x);$$

(ii) If $m > 1$, and (B_1, B_2) is the factorization of B, using θ_i to denote the restriction of θ on $Sbb(B_i)$ for $i = 1, 2$, we set

$$B^\theta(\mathbf{f}) := \theta(B)^{\#}[B_1^{\theta_1}(\mathbf{f}), B_2^{\theta_2}(\mathbf{f})].$$

Remark 2.23. From Remark 2.20 it follows that $B^\theta(\mathbf{f})$ is defined almost everywhere.

Example. Let $B = [[X_1, X_2], X_3]$ and let θ be a twisting of B. Let us set $\Psi_i = \theta(X_i)$, $i = 1, 2, 3$, $\Psi_{12} = \theta([X_1, X_2])$, and $\Psi = \theta(B)$, so that, in particular, $\theta(X_1) = \theta(X_2) = Id_{\mathbb{R}^n}$ since X_1, X_2 are not basic subbrackets of B.

Then

$$\begin{aligned} B^\theta(\mathbf{f}) &= \Psi^{\#}\Big[\Psi_{12}^{\#}\Big[\Psi_1^{\#} f_1,\ \Psi_2^{\#} f_2\Big],\ \Psi_3^{\#} f_3\Big] \\ &= \big(D(\Psi)^{-1} \circ \Psi\big) \cdot \Big(\Big[\ \big(D(\Psi_{12})^{-1} \circ \Psi_{12}\big) \cdot \big([f_1, f_2] \circ \Psi_{12}\big), \\ &\qquad \big(D(\Psi_3)^{-1} \circ \Psi_3\big) \cdot (f_3 \circ \Psi_3)\Big] \circ \Psi\Big). \end{aligned}$$

In order to construct set-valued Lie brackets, it will be sufficient to consider only the subclass of *basic* twistings, that are defined as follows:

Definition 2.24. We say that a twisting $\theta\colon Sbb(B) \to DIFF^\infty(\mathbb{R}^n, \mathbb{R}^n)$ is a **basic twisting** if its restriction to the non-basic subbrackets coincides with the identity twisting (i.e., $\theta(S) = Id_{\mathbb{R}^n}$ as soon as S is not a basic subbracket.

Furthermore, in the proof of the asymptotic formula we will replace the twisting θ appearing in the integral exact formula (see (5.4)) with the *basic twisting* $\theta_{\mathfrak{b}}$ *associated with* θ, which acts only on basic subbrackets and is *equivalent* to θ, by which we mean that

$$B^\theta(\mathbf{f})(x) = B^{\ \theta_{\mathfrak{b}}}(\mathbf{f})(x). \tag{2.6}$$

for any ν-tuple $\mathbf{f} = (f_1, \dots, f_\nu) \in C^{B-1,1}$ and almost every $x \in \mathbb{R}^n$. Let us point out that an equality like (2.6) turns out to be a manifestation of the bracket-preserving property of actions of diffeomorphisms on vector fields.

Let us define the *basic twisting* $\theta_{\mathfrak{b}}$ *associated with* θ in the general case. For this purpose, if B is a bracket of degree m, for each subbracket $S \in Sbb(B)$, let us define Σ_S, the **subbracket chain** of S in B, to be the finite sequence of subbrackets

$$\Sigma_S = \Big(B_{S,0}, \dots, B_{S,\mathfrak{Deg}(S,B)}\Big),$$

where, if we let, for $r = 0, \ldots, \mathfrak{Deg}(S, B)$, $B_{S,r}$ be the unique subbracket of B such that $B_{S,r} \in Sbb(B)$ and $\mathfrak{Deg}(B_{S,r}, B) = r$ (notice that $B_{S,0} = S$ and $B_{S,\mathfrak{Deg}(S,B)} = B$).

Definition 2.25. For a given twisting θ let us define the **basic twisting** $\theta_{\mathfrak{b}}$ **associated with** θ by letting

1. $\theta_{\mathfrak{b}}(S) = Id_{\mathbb{R}^n}$ as soon S is a subbracket of B which is not basic[12];
2. $\theta_{\mathfrak{b}}(S) := \theta(B_{S,\mathfrak{Deg}(S,B)}) \circ \cdots \circ \theta(B_{S,0})$ for any basic subbracket S.

Notice that a twisting θ is a basic twisting if and only if $\theta = \theta_{\mathfrak{b}}$.

Proposition 2.26 below will be essential in the proof of the Asymptotic Formula (Theorem 3.7). Its trivial proof is omitted.

Proposition 2.26. *If B is a bracket and θ is a twisting of B, then θ is* equivalent *to its associated basic twisting $\theta_{\mathfrak{b}}$, by which we mean that*

$$B^{\theta}(\mathbf{f})(x) = B^{\,\theta_{\mathfrak{b}}}(\mathbf{f})(x) \tag{2.7}$$

for all $\mathbf{f} \in C^{B-1,1}$ and x such that both members are defined.[13]

Example. Consider the formal bracket $B = [[X_1, X_2], X_3]$, and let θ be a twisting of B. So $\theta_{\mathfrak{b}}$ is defined by setting

$$\theta_{\mathfrak{b}}(X_1) = \theta(X_2) = Id_{\mathbb{R}^n}, \quad \theta_{\mathfrak{b}}(X_3) = \theta([[X_1, X_2], X_3]) \circ \theta(X_3),$$
$$\theta_{\mathfrak{b}}([X_1, X_2]) = \theta([[X_1, X_2], X_3]) \circ \theta([X_1, X_2]), \quad \theta_{\mathfrak{b}}([[X_1, X_2], X_3]) = Id_{\mathbb{R}^n}$$

It is straightforward to verify directly that $\theta_{\mathfrak{b}}$ is equivalent to θ, namely $B^{\theta}(\mathbf{f})(x) = B^{\,\theta_{\mathfrak{b}}}(\mathbf{f})(x)$.

Indeed, if $\mathbf{f} \in C^{B-1,1}$ and $\Psi_3 := \theta(X_3)$, $\Psi_{12} := \theta([X_1, X_2])$, and $\Psi := \theta(B)$, one has

$$\begin{aligned}
&B^{\theta}(\mathbf{f})(x) \\
&= D\left(\Psi\right)^{-1}(\Psi(x)) \cdot \Bigg(\Bigg[\left(D(\Psi_{12})^{-1} \circ \Psi_{12}\right) \cdot \left(\Big[f_1, f_2\Big] \circ \Psi_{12}\right), \\
&\qquad \left(D(\Psi_3)^{-1} \circ \Psi_3\right) \cdot \left(f_3 \circ \Psi_3\right) \Bigg] \circ \Psi(x) \Bigg) \\
&= \Bigg[\left(D(\Psi_{12} \circ \Psi)^{-1} \circ \Psi_{12} \circ \Psi\right) \cdot \left(\Big[f_1, f_2\Big] \circ (\Psi_{12} \circ \Psi)\right), \\
&\qquad \left(D(\Psi_3 \circ \Psi)^{-1}\right) \cdot \Big(f_3 \circ (\Psi_3 \circ \Psi)\Big) \Bigg](x) \quad = B^{\,\theta_{\mathfrak{b}}}(\mathbf{f})(x).
\end{aligned}$$

When one considers a basic shift-twisting of a bracket B (instead of a general twisting), the notion of twisted Lie bracket simplifies into that of “shifted Lie bracket”. In particular, unlike a general twisted Lie bracket, a shifted Lie bracket depends only on d parameters in $\mathbb{R}^n$, where $d := \mathfrak{Deg}(B)$.

[12] Either S has degree > 2, or $S = X_j$ for some j and X_j is not a factor of a subbracket of degree 2 of S.

[13] We remind that the set of such x has full measure.

Definition 2.27. Let B be a bracket of degree $m \geq 1$ and let us order the basic subbrackets of B, e.g., lexicographically: we use $S_1, \ldots, S_d$ to denote them. For every $(h_1, \ldots, h_d) \in (\mathbb{R}^n)^d$, let $\Theta^{(h_1,\ldots h_d)}$ denote the shift-twisting given by

$$\Theta^{(h_1,\ldots h_d)}(S_j)(x) = \tau_{h_j}(x) \ = x + h_j, \quad \text{for all } j = 1, \ldots, d, \ \text{ for all } x \in \mathbb{R}^n.$$

If $\mu \geq 0$ is the integer such that $Seq(B) = X_{\mu+1} \ldots X_{\mu+m}$, let $\mathbf{f} = (f_1, \ldots, f_\nu)$, $\nu \geq m+\mu$, be an ν-tuple of vector fields of class $C^{B-1,1}$. For $x \in \mathbb{R}^n$ and almost every value of the parameter $(h_1, \ldots, h_d) \in (\mathbb{R}^n)^d$, the twisted bracket at x

$$B^{\Theta^{(h_1,\ldots h_d)}}(\mathbf{f})(x)$$

will be called the $(h_1, \ldots h_d)$-**shifted bracket at** x.

By merely applying Definition 2.22 (and keeping in mind the additivity of the diff-degree) it turns out that $B^{\Theta^{(h_1,\ldots h_d)}}(\mathbf{f})(x)$ can be defined through recursion on the diff-degree:

Proposition 2.28. *For all $x \in \mathbb{R}^n$ one has*

$$B^{\Theta^{(h)}}(\mathbf{f})(x) := \left(\Theta^{(h)}(B)\right)^{\#} B(\mathbf{f})(x) = B(\mathbf{f})(x+h)$$

if $\mathfrak{Deg}(B) = 1$ and $h \in \mathbb{R}^n$. Moreover, if $1 < d := \mathfrak{Deg}(B)$ and B has factorization (B_1, B_2), with $d_1 := \mathfrak{Deg}(B_1)$, $d_2 := \mathfrak{Deg}(B_2)$ (so that $d_1 + d_2 = d$), one has

$$B^{\Theta^{(h_1,\ldots h_d)}}(\mathbf{f})(x) = \left[B_1^{\Theta^{(h_1,\ldots h_{d_1})}}(\mathbf{f}),\ B_2^{\Theta^{(h_{d_1+1},\ldots h_d)}}(\mathbf{f})\right](x)$$

for all $(h_1, \ldots, h_d)$ in a full measure subset of $(\mathbb{R}^n)^d$ depending on x.

Example.

1. Let us consider the formal bracket $B = [X_2, [X_3, X_4]]$, and $\mathbf{f} = (f_1, \ldots, f_6)$, with f_2 a vector field of class $C^{0,1}$ (i.e., locally Lipschitz) and f_3, f_4 vector fields of class $C^{1,1}$. One has $\mathfrak{Deg}(B) = 2$, and the basic subbrackets are $S_1 = X_2$ and $S_2 = [X_3, X_4]$. So, for every value of the parameter $(h_1, h_2) \in (\mathbb{R}^n)^2$ one has $\Theta^{(h_1,h_2)}(S_1) = \tau_{h_1}$ and $\Theta^{(h_1,h_2)}(S_2) = \tau_{h_2}$. Finally, for all (h_1, h_2) such that $(Df_2(x+h_1), D[f_3, f_4](x+h_2))$ is defined (i.e., almost everywhere, see Remark 2.20), one has

$$\begin{aligned}\big[X_2, [X_3, X_4]\big]^{\Theta^{(h_1,h_2)}}(\mathbf{f})(x) &= \big[f_2 \circ \tau_{h_1},\ [f_3, f_4] \circ \tau_{h_2}\big](x) \\ &= D[f_3, f_4](x + h_2) \cdot f_2(x + h_1) \\ &\quad -Df_2(x + h_1) \cdot [f_3, f_4](x + h_2).\end{aligned}$$

2. Similarly, if f_2, f_3, f_4, f_5 are vector fields of class $C^{1,1}$,

$$\begin{aligned}\big[[X_2, X_3], [X_4, X_5]\big]^{\Theta^{(h_1,h_2)}}(\mathbf{f})(x) &= \big[[f_2, f_3] \circ \tau_{h_1},\ [f_4, f_5] \circ \tau_{h_2}\big](x) \\ &= D[f_4, f_5](x + h_2) \cdot [f_2, f_3](x + h_1) \\ &\quad -D[f_2, f_3](x + h_1) \cdot [f_4, f_5](x + h_2).\end{aligned}$$

   Notice that $\mathfrak{Deg}\left(\big[[X_2, X_3], [X_4, X_5]\big]\right) = 2$, i.e., in this case the diff-degree is 2, so that the shifted bracket depends on $2n$ parameters.

3. One more n-dimensional parameter is needed for the bracket $\big[[X_2,[X_3, X_4]],X_5\big]$, . More precisely, $\mathfrak{Deg}\left(\big[[X_2,[X_3,X_4]],X_5]\big]\right)=3$, and, if f_2 is of class $C^{1,1}$, f_3, f_4 are of class $C^{2,1}$, and f_5 is of class $C^{0,1}$, for every $(h_1,h_2,h_3)\in(\mathbb{R}^n)^3$ one has

$$\begin{aligned}\big[[X_2,[X_3,X_4]],X_5\big]^{\Theta^{(h_1,h_2,h_3)}}(\mathbf{f})(x) &= \big[[f_2\circ\tau_{h_1},[f_3,f_4]\circ\tau_{h_2}],f_5\circ\tau_{h_3}\big](x)\\ &= Df_5(x+h_3)\cdot\Big(D[f_3,f_4](x+h_2)\cdot f_2(x+h_1)\\ &\quad -Df_2(x+h_1)\cdot[f_3,f_4](x+h_2)\Big)\\ &\quad -D\Big(D[f_3,f_4](x+h_2)\cdot f_2(x+h_1)\\ &\quad -Df_2(x+h_1)\cdot[f_3,f_4](x+h_2)\Big)\cdot f_5(x+h_3).\end{aligned}$$

3. Set-valued brackets and the Asymptotic Formula

3.1. The set-valued bracket

For every $x\in\mathbb{R}^n$, we will obtain the **set-valued bracket** as the convex hull of the limits of the shift-twisted brackets as the shift $(h_1,\dots,h_d)$ goes to zero. If E is a subset of a $\mathbb{R}$-linear space, let us use $co(E)$ to denote the convex hull of E.

Definition 3.1. (*SET-VALUED BRACKET*) Let B be a bracket and let $m := deg(B)$, $d := \mathfrak{Deg}(B)$. If $\mu\ge 0$ and $m\ge 1$ are integers such that $Seq(B) = X_{\mu+1}\dots X_{\mu+m}$ and $\mathbf{f}=(f_1,\dots,f_\nu)$, where $\nu\ge m+\mu$ is a ν-tuple of vector fields of class $C^{B-1,1}$, we define the **set-valued bracket** $B_{set}(\mathbf{f})$ at a point $x\in\mathbb{R}$ as the convex hull of all limits of "shift-twisted" brackets at x as the nd-dimensional parameter $(h_{j_1},\dots,h_{j_d})$ goes to zero. Expressly,

$$B_{set}(\mathbf{f})(x) := co\Big\{\lim_{(h_{j_1},\dots,h_{j_d})\to 0} B^{\Theta^{(h_{j_1},\dots h_{j_d})}}(\mathbf{f})(x)\Big\} \tag{3.1}$$

where, for any $x\in\mathbb{R}^n$, limits are taken along all sequences $\big((h_{j_1},\dots,h_{j_d})\big)_{j\in\mathbb{N}}\subset(\mathbb{R}^n)^d$ converging to zero and such that the vector $B^{\Theta^{(h_{j_1},\dots,h_{j_d})}}(\mathbf{f})(x)$ is defined.[14]

Example. In view of Example 2.5 we have that, if $B=[X_2,[X_3,X_4]]$ and $\mathbf{f}=(f_1,\dots,f_6)$, with f_2 of class $C^{0,1}$ and f_3,f_4 of class $C^{1,1}$,

$$\begin{aligned}\big[f_2,[f_3,f_4]\big]_{set}(x) = co\Big\{&\lim_{(h_{j_1},h_{j_2})\to 0} D[f_3,f_4](x+h_{j_2})\cdot f_2(x+h_{j_1})\\ &-Df_2(x+h_{j_1})\cdot[f_3,f_4](x+h_{j_2})\Big\},\end{aligned}$$

where, we have written $[f_2,[f_3,f_4]]_{set}(x)$ in place of $\big[X_2,[X_3,X_4]\big]_{set}(\mathbf{f})(x)$.

[14] As remarked above, $B^{\Theta^{(h_{j_1},\dots,h_{j_d})}}(\mathbf{f})(x)$ is defined for almost every $(h_{j_1},\dots,h_{j_d})$.

Similarly, if f_2, f_3, f_4, f_5 are vector fields of class $C^{1,1}$ and $B = \big[[X_2, X_3], [X_4, X_5]\big]$,

$$\big[[f_2, f_3], [f_4, f_5]\big]_{set}(x) = co\Big\{ \lim_{(h_{j_1}, h_{j_2}) \to 0} D[f_4, f_5](x + h_2) \cdot [f_2, f_3](x + h_{j_1}) - D[f_2, f_3](x + h_{j_1}) \cdot [f_4, f_5](x + h_{j_2}) \Big\}.$$

Furthermore, if f_2 is of class $C^{1,1}$, f_3, f_4 are of class $C^{2,1}$, and f_5 is of class $C^{0,1}$, for every $(h_1, h_2, h_3) \in (\mathbb{R}^n)^3$ one has

$$\begin{aligned}
&\big[[f_2, [f_3, f_4]], f_5\big]_{set}(x) \\
&\quad = co\Big\{ \lim_{(h_{j_1}, h_{j_2}, h_{j_3}) \to 0} \Big(Df_5(x + h_{j_3}) \cdot \Big(D[f_3, f_4](x + h_{j_2}) \cdot f_2(x + h_{j_1}) \\
&\qquad - Df_2(x + h_{j_1}) \cdot [f_3, f_4](x + h_{j_2}) \Big) D\Big(D[f_3, f_4](x + h_{j_2}) \cdot f_2(x + h_{j_1}) \\
&\qquad - Df_2(x + h_{j_1}) \cdot [f_3, f_4](x + h_{j_2}) \Big) \cdot f_5(x + h_{j_3}) \Big) \Big\}.
\end{aligned}$$

Example. Let f_1, f_2 be the vector fields in $\mathbb{R}^2$ defined by

$$f_1(x, y) := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f_2(x, y) := \begin{pmatrix} 0 \\ \alpha(x) \end{pmatrix} \quad \text{for all } (x, y) \in \mathbb{R}^2, \tag{3.2}$$

where, for some $m \geq 2$ $\alpha\colon \mathbb{R} \to \mathbb{R}$ is a function of class $C^{m-2,1}$.

At differentiability points of $D^{m-2}\alpha$ one has

$$\underbrace{[f_1, [f_1, [\cdots, [f_1, f_2]]]]}_{\text{bracket of degree } m}(x, y) = D^{m-1}\alpha(x) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Therefore,

$$\begin{aligned}
\underbrace{[f_1, [f_1, [\cdots, [f_1, f_2]]]]_{set}}_{\text{bracket of degree } m}(x, y) &= \partial_C \big(D^{m-2}\alpha\big)(x) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&:= \left\{ \lambda \begin{pmatrix} 0 \\ 1 \end{pmatrix} : \lambda \in \partial_C \big(D^{m-2}\alpha\big)(x) \right\},
\end{aligned}$$

where, for any locally Lipschitz function $\phi\colon \mathbb{R} \to \mathbb{R}$, $\partial_C \phi(x)$ denotes Clarke's generalized derivative[15] Needless to say that if α is of class C^{m-1} in some neighborhood of a point x deprived of x itself, then

$$\underbrace{[f_1, [f_1, [\cdots, [f_1, f_2]]]]_{set}}_{\text{bracket of degree } m}(x, y) = [a, b] \begin{pmatrix} 0 \\ 1 \end{pmatrix} := \left\{ \begin{pmatrix} 0 \\ \lambda \end{pmatrix} : \lambda \in [a, b] \right\},$$

[15] We remind that for a locally Lipschitz map $\phi\colon \mathbb{R} \to \mathbb{R}$ the **Clarke's generalized derivative** at a point $x \in \mathbb{R}$ is

$$\partial_C \phi(x) := co\Big\{ \lim_{j \to \infty} D\phi(x_j) : (x_j)_{j \in \mathbb{N}} \text{ sequence of differentiability points of } \phi \text{ such that } x_j \to x \text{ as } j \to \infty \Big\}.$$

$$a := \min\left\{\alpha^{(m-1)}(x^-), \alpha^{(m)}(x^+)\right\}, \quad b := \max\left\{\alpha^{(m-1)}(x^-), \alpha^{(m-1)}(x^+)\right\}.$$

Proposition 3.2. *The set-valued map $x \mapsto B_{set}(\mathbf{f})(x)$ is upper semi-continuous with convex, compact, non-empty values.*

Proof. The fact that for every $x \in \mathbb{R}^n$ is compact convex follows directly from the definition (and by the fact that the set-valued map is $x \mapsto B(\mathbf{f})(x)$ is locally bounded). Furthermore, let us observe that the map

$$(x, h_1, \ldots, h_d) \mapsto B^{\Theta^{(h_1,\ldots h_d)}}(\mathbf{f})(x)$$

is defined almost everywhere and bounded measurable. More precisely, this map is defined on a set of the form

$$\bigcup_{x\in\mathbb{R}^n} \Big(\{x\} \times E_x\Big),$$

where, for every $x \in \mathbb{R}$, $E_x \subset (\mathbb{R}^n)^d$ has full Lebesgue measure. Moreover, since $\tau_h(x) = \tau_{h+x-y}(y)$ for all $x, y, h \in \mathbb{R}^n$, one easily obtains that

$$\begin{aligned} B_{set}(\mathbf{f})(x) &= co\Big\{v\colon v = \lim_{(h_{j_1},\ldots,h_{j_d})\to 0} B^{\Theta^{(h_{j_1},\ldots h_{j_d})}}(\mathbf{f})(x)\Big\} \\ &= co\Big\{v\colon v = \lim_{(y,h_{j_1},\ldots,h_{j_d})\to (x,0)} B^{\Theta^{(h_{j_1},\ldots h_{j_d})}}(\mathbf{f})(x)\Big\} \end{aligned}$$

which, in turn, implies that the map $x \mapsto B_{set}(\mathbf{f})(x)$ is upper semi-continuous on $\mathbb{R}^n$. □

Remark 3.3. Notice that as soon as $\mathbf{f}$ is of class C^B at $\bar{x} \in \mathbb{R}^n$, one recovers the classical, single-valued bracket, namely $B_{set}(\mathbf{f})(\bar{x}) = \{B(\mathbf{f})(\bar{x})\}$.

Remark 3.4. On one hand, the set-valued bracket $B_{set}(\mathbf{f})(x)$ is *small enough*: for instance, it is contained in the set one would obtain by formally replacing the classical derivatives with Clarke's generalized derivatives. This allows idempotency, namely the fact that for every locally Lipschitz vector field f one has

$$[f, f]_{set} \equiv \{0\}$$

to hold true (which would be not the case if we used Clarke's generalized derivatives, as one can easily check by considering the vector field $f(x) = |x|$ in $\mathbb{R}$). This same "smallness of the bracket" allows to prove a "Frobenius type Theorem" (see [7]), namely the characterization of local integrability through (suitably rephrased) involutivity.

However, for degree > 2, *a too small bracket*[16] would not be fit for generating sufficient conditions giving controllability, as in Chow–Rashevski's Theorem. This explains why the definition of set-valued bracket involves limits along d-tuples of $\mathbb{R}^n$-sequences, where d is the diff-degree of the bracket, rather than

[16] For instance, the "bracket" one would obtain by mimicking the case of the degree 2.

along limits of single $\mathbb{R}^n$-sequences (as in the degree two case). Among the consequences of such a construction one has that

$$\big[[g,h],[g,h]\big]_{set} \supseteq [f,f]_{set} = \{0\} \tag{3.3}$$

for any pair of vector fields g,h of class $C^{1,1}$ (so that $f := [g,h]$ is locally Lipschitz). Furthermore, one can exhibit examples (see [9, Section 7]) where the inclusion is strict.

Remark 3.5. A relation between the set-valued bracket and Clarke's generalized derivative is recovered as soon as one of the two terms of the factorization is slightly more regular. Precisely, if $B = [B_1, B_2]$, $\mathbf{f}_1$ is of class C^{B_1+1}, and $\mathbf{f}_2$ is of class $C^{B_2,1}$, then:

1. The vector fields $F_1 = B_1(\mathbf{f}_1)$ and $F_2 = B_1(\mathbf{f}_2)$ are of class C^1 and $C^{0,1}$, respectively;
2. The pair of vector fields $\mathbf{f} = (\mathbf{f}_1, \mathbf{f}_2)$ is of class $C^{B-1,1}$;
3. One has

$$B(\mathbf{f}) = \partial_C F_2 \cdot F_1 - \nabla F_1 \cdot F_2 = \Big\{ M \cdot F_1 - \nabla F_1 \cdot F_2 \ : \ M \in \partial_C F_2 \Big\},$$

where $\partial_C F_2$ denotes the Clarke's generalized derivative of F_2.

3.2. The asymptotic formula

We will now state the main result of the paper, an *asymptotic formula*, namely an approximation at a point x of a certain composition of flows related to a given bracket $B(\mathbf{f})$ and to times $t_1, \ldots, t_m$ by $t_1 \cdots t_m B(\mathbf{f})(x)$.

To begin with, let us establish the multi-flows we will be interested in.

Definition 3.6. Let B be a bracket of degree $m \geq 1$ and let $\mu \geq 0$ be an integer such that $Seq(B) = X_{\mu+1} \ldots X_{\mu+m}$. For some $\nu \geq m + \mu$, let $\mathbf{f} = (f_1, \ldots, f_\nu)$ be an ν-tuple of Lipschitz vector fields of class $C^{B-1,1}$. For every $(t_1, \ldots, t_m) \in \mathbb{R}^m$ we define the multi-flow $x \mapsto \Psi_B^{\mathbf{f}}(t_1, \ldots, t_m)(x)$ recursively as follows:

(i) if $deg(B) = m = 1$, i.e., $B = X_j$ for some $j \in \mathbb{N}$, we set

$$\Psi_B^{\mathbf{f}}(t)(x) \ := \ e^{tf_j}(x), \quad \text{for all } x \in \mathbb{R}^n,$$

i.e., $\Psi_B^{\mathbf{f}}(t)(x)$ is the value at t of the solution to the Cauchy problem $\dot{y} = f_j(y),\ y(0) = x$;

(ii) if $deg(B) = m > 1$ and (B_1, B_2) is the factorization of B, we set

$$\begin{aligned}\Psi_B^{\mathbf{f}}(t_1, \ldots, t_m)(x) := &\big(\Psi_{B_2}^{\mathbf{f}}(t_{m_1+1}, \ldots, t_m)\big)^{-1} \circ \big(\Psi_{B_1}^{\mathbf{f}}(t_1, \ldots, t_{m_1})\big)^{-1} \\ &\circ \Psi_{B_2}^{\mathbf{f}}(t_{m_1+1}, \ldots, t_m) \ \circ \Psi_{B_1}^{\mathbf{f}}(t_1, \ldots, t_{m_1})(x)\end{aligned}$$

for all $x \in \mathbb{R}^n$, where $m_1 := deg(B_1)(< m)$.

Example.

1. If $B = [X_1, X_2]$ and $\mathbf{f} = (f, g)$, then

$$\Psi_B^{\mathbf{f}}(t_1, t_2)(x) = e^{-t_2 g} \circ e^{-t_1 f} \circ e^{t_2 g} \circ e^{t_1 f}(x).$$

2. If $B=[X_1,[X_2,X_3]]$ and $\mathbf{f}=(f,g,h)$, then
$$\Psi_B^{\mathbf{f}}(t_1,t_2,t_3)(x)=e^{-t_2g}\circ e^{-t_3h}\circ e^{t_2g}\circ e^{t_3h}\circ e^{-t_1f}\circ e^{-t_3h}\circ e^{-t_2g}\\ \circ e^{t_3h}\circ e^{t_2g}\circ e^{t_1f}(x)$$
3. if $B=[[X_1,X_2],[X_3,X_4]]$ and $\mathbf{f}=(f,g,h,k)$, then
$$\Psi_B^{\mathbf{f}}(t_1,t_2,t_3,t_4)(x)=e^{-t_3h}\circ e^{-t_4k}\circ e^{t_3h}\circ e^{t_4k}\circ e^{-t_1f}\circ e^{-t_2g}\circ e^{t_1f}\circ e^{t_2g}\\ \circ e^{-t_4k}\circ e^{-t_3h}\circ e^{t_4k}\circ e^{t_3h}\circ e^{-t_2g}\circ e^{-t_1f}\circ e^{t_2g}\circ e^{t_1f}(x).$$

Theorem 3.7. (ASYMPTOTIC FORMULA) *Let B a bracket of degree $m\geq 2$ and let $n\in\mathbb{N}$. Then there exists a number K, depending only on m and n, such that, for every $\mathbf{f}=(f_1,\dots,f_\mu)\in C_0^{B-1,1}$, one has*
$$\begin{aligned}&\operatorname{dist}\Big(\Psi_B^{\mathbf{f}}(t_1,\dots,t_m)(x)-x,\ t_1\cdots t_m\cdot B_{set}(\mathbf{f})(x_*)\Big)\\ &\leq |t_1\cdots t_m|\Bigg(\gamma\Big(K\|\mathbf{f}\|_{B-1,1}|(t_1,\dots,t_m)|+|x-x_*|\Big)\\ &\quad+K\cdot\|\mathbf{f}\|_{B-1,1}^{m+1}\cdot|(t_1,\dots,t_m)|\Bigg)\end{aligned}\tag{3.4}$$
for all $x,x_*\in\mathbb{R}^n$ and $(t_1,\dots,t_m)\in\mathbb{R}^m$ with $|(t_1,\dots,t_m)|\|\mathbf{f}\|_{B-1,1}\leq 1$, where γ is the modulus of continuity of the map $x\mapsto B(\mathbf{f})(x)$ at x_* defined by formula (4.6) below. In particular,
$$\lim_{(t_1,\dots,t_m,x)\to(0,x_*)}\operatorname{dist}\left(\frac{\Psi_B^{\mathbf{f}}(t_1,\dots,t_m)(x)-x}{t_1\cdots t_m},\ B_{set}(\mathbf{f})(x_*)\right)=0.\tag{3.5}$$

4. Proof of Theorem 3.7 (asymptotic formula)

The proof of Theorem 3.7 will be based on the exact integral formula, recalled in Sect. 5, and on a series of technical estimates which are proved in Sect. 6. We will subdivide the proof in four successive steps. First, we will apply the integral exact formula proved in [4] to some regularizations of the vector fields $f_1,\dots,f_m$. Secondly, we will approximate the twisted brackets contained in the integral formulas with shift-twisted brackets. This will be utilized in the third step, where we will obtain an estimate for the distance of the multi-flows relative to the regularized vector fields from the set-valued Lie brackets. The proof will be finalized by letting the regularization parameter go to zero.

Step 1. *The integral formula for the regularized vector fields.*

Let $\phi\colon\mathbb{R}^n\to[0,+\infty[$ be a mollifier—i.e., ϕ is a C^∞, compactly supported function with L^1-norm equal to 1—and, for every $\zeta\geq 0$, let us consider the regularized vector fields
$$f_i^\zeta(x):=\int_{\mathbb{R}^n}f_i(x+\zeta h)\phi(h)\,dh\quad\forall x\in\mathbb{R}^n.$$
Correspondingly, let $\Psi_B^{\mathbf{f}^\zeta}$ be the multi-flow associated with the bracket B and the m-tuple $\mathbf{f}^\zeta=(f_1^\zeta\dots,f_m^\zeta)$ of regularized vector fields, as in Definition 3.6.

Since the vector fields f_i^ζ, $i = 1, \dots, m$ are of class C_0^∞, we can apply the exact integral formula (5.7) below (see Sect. 5), which gives

$$\begin{aligned} &\Psi_B^{\mathbf{f}^\zeta}(t_1, \dots, t_m)(x) - x \\ &\quad = \int_0^{t_1} \cdots \int_0^{t_m} B^{\theta^\zeta}(\mathbf{f}^\zeta) \left(\Psi_B^{\mathbf{f}^\zeta}(t_1, \dots, t_{m-1}, s_m)(x) \right) ds_1 \dots ds_m. \end{aligned} \tag{4.1}$$

Here, in connection with the choice $(g_1, \dots, g_m) = (f_1^\zeta, \dots, f_m^\zeta)$, θ^ζ stands for the (*basic*) twisting utilized in Theorem 5.10 below, namely

$$\theta^\zeta := \ \theta_{\mathfrak{b}}{}^{(t_1, \dots, t_{m_1-1}, t_{m_1+1}, \dots, t_{m-1}, s_m, s_1, \dots, s_{m-1})}.$$

Step 2. *Approximations through shift-twistings.*

When, on the right hand-side of (4.1), the twisting θ^ζ is replaced by the shift approximation $\theta^\zeta_{x_\zeta}$ of θ^ζ at the point $x_\zeta = \Psi_B^{\mathbf{f}^\zeta}(t_1, \dots, t_{m-1}, s_m)(x)$, $s_m \in [0, t_m]$ for $t_m \geq 0$ and $s_m \in [t_m, 0]$ for $t_m < 0$ (see Definition 2.13) one obtains the integral

$$\int_0^{t_1} \cdots \int_0^{t_m} B^{\theta^\zeta_{x_\zeta}}(\mathbf{f}^\zeta)(x_\zeta)\, ds_1 \dots ds_m.$$

We wish to estimate the corresponding error, namely the quantity

$$\left| \Psi_B^{\mathbf{f}^\zeta}(t_1, \dots, t_m)(x) - x - \int_0^{t_1} \cdots \int_0^{t_m} B^{\theta^\zeta_{x_\zeta}}(\mathbf{f}^\zeta)(x_\zeta) ds_1 \dots ds_m \right|.$$

Making use of Proposition 6.6 below for $k = 0$ and for the m-tuple $\mathbf{g}$ given by the regularization $\mathbf{f}^\zeta$, we get

$$\left| B^{\theta^\zeta}(\mathbf{f}^\zeta)(x_\zeta) - B^{\theta^\zeta_x}(\mathbf{f}^\zeta)(x_\zeta) \right| \leq c_0(m) \|\mathbf{f}^\zeta\|_B^m \cdot \mathbf{d}_{B+1}(\theta, \iota),\text{[17]} \tag{4.2}$$

where $c_0(m)$ is the number defined in (6.21) below.

Hence, for $x_\zeta = \Psi_B^{\mathbf{f}^\zeta}(t_1, \dots, t_{m-1}, s_m)(x)$,

$$\begin{aligned} &\left| \Psi_B^{\mathbf{f}^\zeta}(t_1, \dots, t_m)(x) - x - \int_0^{t_1} \cdots \int_0^{t_m} B^{\theta^\zeta_{x_\zeta}}(\mathbf{f}^\zeta)(x_\zeta) ds_1 \dots ds_m \right| \\ &\quad \leq \left| \Psi_B^{\mathbf{f}^\zeta}(t_1, \dots, t_m)(x) - x - \int_0^{t_1} \cdots \int_0^{t_m} B^{\theta^\zeta}(\mathbf{f}^\zeta)(x_\zeta)\, ds_1 \dots ds_m \right| \\ &\qquad + \int_0^{t_1} \cdots \int_0^{t_m} \left| B^{\theta^\zeta}(\mathbf{f}^\zeta)(x_\zeta) - B^{\theta^\zeta_x}(\mathbf{f}^\zeta)(x_\zeta) \right| ds_1 \dots ds_m \\ &\quad = \int_0^{t_1} \cdots \int_0^{t_m} \left| B^{\theta^\zeta}(\mathbf{f}^\zeta)(x_\zeta) - B^{\theta^\zeta_x}(\mathbf{f}^\zeta)(x_\zeta) \right| ds_1 \dots ds_m \\ &\quad \leq c_0(m) \cdot \|\mathbf{f}^\zeta\|_B^m \cdot \widehat{\mathbf{d}}_B(D\theta, D\iota)|t_1| \cdots |t_m| \\ &\quad \leq c_0(m)k(m) \cdot \|\mathbf{f}^\zeta\|_B^{m+1} \cdot |(t_1, \dots, t_m)| \cdot |t_1| \cdots |t_m|, \end{aligned} \tag{4.3}$$

where $k(m)$ denotes the least number such that, for every bracket B of degree m, $\mathbf{f} \in C^{B-1,1}$, $\zeta > 0$, the estimate

$$\mathbf{d}_{B+1}(\theta^\zeta, \iota) \leq k(m) \|\mathbf{f}^\zeta\|_B |(t_1, \dots, t_m)| \tag{4.4}$$

[17] See (2.3) for the definition of $\mathbf{d}_{B+1}(\theta, \iota)$.

holds true for all $(t_1,\dots,t_m)\in\mathbb{R}^m$ with $\|\mathbf{f}^\zeta\|_B|(t_1,\dots,t_m)|\le 1$. Notice that the existence of $k(m)$ follows, in particular, from the fact that $\theta^\zeta(S)=\theta^\zeta(t_1,\dots,t_{m_1-1},t_{m_1+1},\dots,t_{m-1},s_m,s_1,\dots,s_{m_1})(S)$, $S\in Sbb(B)$, are products of flows of the vector fields $f_1^\zeta,\dots,f_m^\zeta$ as defined by (5.1), (5.2) and by Lemma 6.1.

Therefore we obtain

$$\left|\Psi_B^{\mathbf{f}^\zeta}(t_1,\dots,t_m)(x)-x-\int_0^{t_1}\cdots\int_0^{t_m}B^{\theta^\zeta_{x^\zeta}}(\mathbf{f}^\zeta)(x)\,ds_1\dots ds_m\right| \\ \le c_0(m)k(m)\cdot\|\mathbf{f}^\zeta\|_B^{m+1}\cdot|(t_1,\dots,t_m)|\cdot|t_1|\cdots|t_m|. \tag{4.5}$$

Step 3. *Approximation of the regularized multi-flow through a set-valued bracket.* Now we wish to utilize (4.5) to estimate the distance

$$\mathrm{dist}\left(\Psi_{B(\mathbf{f}^\zeta)}(t_1,\dots,t_m)(x)-x,\ t_1\cdots t_m B_{set}(\mathbf{f})(x_*)\right)$$

for all $x,x_*\in\mathbb{R}^n$ and $(t_1,\dots,t_m)\in\mathbb{R}^m$; we will make (x,ζ) tend to $(x_*,0)$ and get the thesis). For this purpose, let us introduce the function $\gamma\colon[0,+\infty[\,\to[0,+\infty[$,

$$\gamma(\rho):=\sup_\Theta\left\{\mathrm{dist}\left(B^\Theta(\mathbf{f})(x),\ B_{set}(\mathbf{f})(x_*)\right)\right\}, \tag{4.6}$$

where the supremum is taken over all *basic shift-twistings* Θ *of* B and all points $x\in\mathbb{R}^n$ such that

(a) x lies in the domain of $B^\Theta(\mathbf{f})$;
(b) $|x-x_*|+\mathbf{d}_B(\Theta,\iota)\le\rho$.[18]

According to Lemma 6.5 below, the function γ is a *modulus*, i.e., γ is nondecreasing and

$$\lim_{\rho\to0^+}\gamma(\rho)=0.$$

Furthermore, by Proposition 6.8, there exists a constant $C_0\ge0$ such that, for all basic shift-twistings Θ and all $\zeta>0$, one has

$$\mathrm{dist}\left(B^\Theta(\mathbf{f}^\zeta)(x),\ B_{set}(\mathbf{f})(x_*)\right)\le\gamma\big(\mathbf{d}_B\Theta,\iota)+|x-x_*|+\zeta\big)+C_0\zeta\qquad\forall x\in\mathbb{R}^n. \tag{4.7}$$

By applying (4.7) to the shift-twisting $\Theta=\theta^\zeta_{x_\zeta}$ and using (4.5), we obtain

$$\mathrm{dist}\left(\Psi_B^{\mathbf{f}^\zeta}(\mathbf{t})(x)-x,\ t_1\cdots t_m B(\mathbf{f})(x_*)\right) \\ \le|t_1\cdots t_m|\Big(\gamma\big(\mathbf{d}_B(\theta^\zeta_{x_\zeta},\iota)+|x-x_*|+\zeta\big)+C_0\zeta \\ +c_0(m)k(m)\cdot\|\mathbf{f}\|^{m+1}_{B-1,1}\cdot|(t_1,\dots,t_m)|\Big). \tag{4.8}$$

Since in view of (4.4) one has

$$\mathbf{d}_B(\theta^\zeta_{x_\zeta},\iota)\le\mathbf{d}_B(\theta^\zeta,\iota)\le k(m)\|\mathbf{f}^\zeta\|_B|(t_1,\dots,t_m)|,$$

[18] See Remark 2.18 for the definition of $\mathbf{d}_B(\Theta,\iota)$.

from (4.8) one gets

$$\begin{aligned}
&\operatorname{dist}\left(\Psi_{B(\mathbf{f}^\zeta)}(\mathbf{t})(x)-x,\, t_1\cdots t_m B(\mathbf{f})(x_*)\right)\\
&\quad\le |t_1\cdots t_m|\Bigg(\gamma\Big(k(m)\|\mathbf{f}^\zeta\|_B|(t_1,\ldots,t_m)|+|x-x_*|+\zeta\Big)\\
&\qquad + C_0\zeta + c_0(m)k(m)\cdot\|\mathbf{f}\|^{m+1}_{B-1,1}\cdot|(t_1,\ldots,t_m)|\Bigg) \qquad (4.9)
\end{aligned}$$

for all $x\in\mathbb{R}^n$, $(t_1,\ldots,t_m)\in\mathbb{R}^m$, $\zeta>0$ with $\|\mathbf{f}^\zeta\|_B|(t_1,\ldots,t_m)|\le 1$.

Step 4. *Conclusion of the proof.*

Taking the limit on both sides of (4.9) as the parameter ζ goes to zero, we get the thesis of the theorem, with $K = k(m)+c_0(m)k(B)$. □

5. Integral formulas for bracket-generating C^∞ multi-flows

Let us briefly recall the (exact) integral formulas for bracket-generating multi-flows of C^∞ vector fields, which were established in [4] by generalizing to the degree >2 the results in [8]. When not otherwise specified *in this section we shall assume that all vector fields are of class* C^∞ *and* complete (by which we mean that their flows are globally defined, a hypothesis which can be recovered by means of standard "cut-off" arguments).

5.0.1. Integrating brackets. If a bracket B of degree $m\ge 1$ is given together with a m-tuple $\mathbf{g}=(g_1,\ldots,g_m)$ of smooth vector fields, the *integrating bracket* of $B(\mathbf{g})$ is a twisted bracket, depending on $2(m-1)$ parameters, obtained through suitable diffeomorphisms which are compositions of the g_i's flows.

Let m be an integer ≥ 1, and let B be a bracket of degree m, which means $Seq(B) = X_1\ldots X_m$. Let $\mathbf{g}=(g_1,\ldots,g_m)$ be an m-tuple of C^∞ complete vector fields.

Definition 5.1. Let us choose

$$(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})\in\mathbb{R}^{2m-2},$$

and let us consider the $(2m-2)$-parameter twisting

$$\theta_{\mathbf{g}} = \theta_{\mathbf{g}}^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}$$

defined as follows:

- For all $j=1,\ldots,m$, we set
$$\theta_{\mathbf{g}}(X_j) := Id_{\mathbb{R}^n}. \qquad (5.1)$$
- If $S\in Sbb(B)$, with $k:=deg(S)\ge 2$, $S=(S_1,S_2)$, $k_1 := deg(S_1)(<k)$, and $Seq(S)=(X_{r+1},\ldots,X_{r+k})$ for some $r\in\{0,1,\ldots,m-2\}$, we set

$$\begin{aligned}
&\theta_{\mathbf{g}}(S)\\
&:= \Psi^{\mathbf{g}}_{S_2}(t_{r+k_1+1},\ldots,t_{r+k-1},s_{r+k})\circ\Psi^{\mathbf{g}}_{S_1}(t_{r+1},\ldots,t_{r+k_1-1},s_{r+k_1}). \qquad (5.2)
\end{aligned}$$

Remark 5.2. Notice that $\theta_{\mathbf{g}}=\iota$ when $m=1$. Moreover, when one of the indexes m_1,m_2 is equal one, formula (5.2) has to be interpreted as follows:

- If $m_1 = m_2 = 1$,
$$\theta_{\mathbf{g}} = \theta_{\mathbf{g}}^{s_2,s_1} := e^{s_2 g_2} \circ e^{s_1 g_1}.$$
- If $m_1 > 1$ and $m_2 = 1$,
$$\theta_{\mathbf{g}} := e^{s_{m_1+1} g_{m_1+1}} \circ \Psi_{B_1}^{\mathbf{g}}(t_1, \ldots, t_{m_1-1}, s_{m_1})$$
- If $m_1 = 1$ and $m_2 > 1$,
$$\theta_{\mathbf{g}} := \Psi_{B_2}^{\mathbf{g}}\left(t_2 \ldots, t_{m_2}, s_{1+m_2}\right) \circ e^{s_1 f_1}$$

Definition 5.3. For every value of the parameter
$$(t_1, \ldots, t_{m_1-1}, t_{m_1+1}, \ldots, t_{m-1}, s_m, s_1, \ldots, s_{m-1}) \in \mathbb{R}^{2m-2},$$
the vector field $B^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}(\mathbf{g})$ defined by setting, for every $x \in \mathbb{R}^n$,
$$\begin{aligned} &B^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}(\mathbf{g})(x) \\ &:= B^{\theta_{\mathbf{g}}^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}}(\mathbf{g})(x) \end{aligned}$$
is called the **integrating bracket corresponding to B and $\mathbf{g}$**.

Example 5.4. If $B = X_1$,
$$B(\mathbf{g})(x) = g_1(x).$$

Example 5.5. If $B = [X_1, X_2]$,
$$\begin{aligned} B(\mathbf{g})^{(s_2,s_1)}(x) &= [g_1, g_2]^{(s_2,s_1)}(x) \\ &= D\left(e^{s_1 g_1} \circ e^{s_2 g_2}\right)^{-1} \cdot [g_1,\ g_2]\left(e^{s_1 g_1} \circ e^{s_2 g_2}\right)(x). \end{aligned}$$

Example 5.6. If $B = \Big[[X_1, X_2], X_3\Big]$,
$$\begin{aligned} B(\mathbf{g})^{(t_1,s_3,s_1,s_2)}(x) &= \Big[[g_1, g_2], g_3\Big]^{(t_1,s_3,s_1,s_2)}(x) \\ &= \Psi^{\#}\Big[[g_1, g_2]^{(s_2,s_1)},\ g_3\Big](x) \\ &= D\Psi^{-1} \cdot \Big[D\left(\Psi_{12}^{-1}\right) \cdot \big[g_1,\ g_2\big] \circ \Psi_{12},\ g_3\Big]\left(\Psi(x)\right), \end{aligned} \tag{5.3}$$
where we have set
$$\begin{aligned} \Psi &:= e^{-s_2 g_2} \circ e^{-t_1 g_1} \circ e^{s_2 g_2} \circ e^{t_1 g_1} \circ e^{s_3 g_3}, \\ \Psi_{12} &:= e^{s_1 g_1} \circ e^{s_2 g_2}. \end{aligned}$$

Example 5.7. If $B = \Big[[X_1, X_2], [X_3, X_4]\Big]$,
$$\begin{aligned} B(\mathbf{g})^{(t_1,t_3,s_4,s_1,s_2,s_3)}(x) &= \Big[[g_1, g_2], [g_3, g_4]\Big]^{(t_1,t_3,s_4,s_1,s_2,s_3)}(x) \\ &= \Psi^{\#}\Big[[g_1, g_2]^{(s_2,s_1)},\ [g_3, g_4]^{(s_4,s_3)}\Big](x) \\ &= D\Psi^{-1} \cdot \Big[D\left(\Psi_{12}^{-1}\right) \cdot [g_1,\ g_2] \circ \Psi_{12}, D\left(\Psi_{34}^{-1}\right) \\ &\quad \cdot [g_3,\ g_4] \circ \Psi_{34}\Big]\left(\Psi(x)\right), \end{aligned}$$

where

$$\Psi := e^{-s_2g_2} \circ e^{-t_1g_1} \circ e^{s_2g_2} \circ e^{t_1g_1} \circ e^{-s_4g_4} \circ e^{-t_3g_3} \circ e^{s_4g_4} \circ e^{t_3g_3},$$
$$\Psi_{12} := e^{s_1g_1} \circ e^{s_2g_2}, \quad \Psi_{34} := e^{s_3g_3} \circ e^{s_4g_4}.$$

5.0.2. Integral formulas. Let B be a formal bracket of degree $m \geq 1$, let m_1 be the degree of the first bracket of the factorization of B, and let $\mathbf{g} = (g_1, \ldots, g_m)$ be a m-tuple of C^∞ vector fields.

Theorem 5.8. (Integral representation) *For every m-tuple $\mathbf{t} = (t_1, \ldots, t_m) \in \mathbb{R}^m$ one has*

$$\begin{aligned}\Psi_B^{\mathbf{g}}&(t_1, \ldots, t_m)(x) - x \\ &= \int_0^{t_1} \cdots \int_0^{t_m} B^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}(\mathbf{g}) \\ &\quad \left(\Psi_B^{\mathbf{g}}\left(t_1, \ldots, t_{m-1}, s_m\right)(x)\right) ds_1 \ldots ds_m. \end{aligned} \tag{5.4}$$

Remark 5.9. These formulas generalize the familiar integral representation of ODE's. Indeed, if $B = X_1$, $\mathbf{g} = g_1$ one gets

$$e^{t_1g_1}(x) - x = \Psi_B^{\mathbf{g}}(t_1)(x) - x = \int_0^{t_1} B(\mathbf{g})(\Psi_B^{\mathbf{g}}(s)(x))ds = \int_0^{t_1} g_1(e^{sg_1}(x))ds.$$

Furthermore, in relation with the brackets considered in the previous three examples, the general formulas read

$$\Psi^{\mathbf{g}}_{[X_1,X_2]}(t_1,t_2)(x) = x + \int_0^{t_1}\int_0^{t_2} [g_1,g_2]^{(s_2,s_1)}\left(\Psi^{\mathbf{g}}_{[X_1,X_2]}(t_1,s_2)(x)\right) ds_1\, ds_2, \tag{5.5}$$

$$\begin{aligned}&\Psi^{\mathbf{g}}_{[[X_1,X_2],X_3]}(t_1,t_3,t_3)(x) \\ &\quad = x + \int_0^{t_1}\int_0^{t_2}\int_0^{t_3} [[g_1,g_2],g_3]^{(t_1,s_3,s_1,s_2)}\left(\Psi^{\mathbf{g}}_{[[X_1,X_2],X_3]}(t_1,t_2,s_3)(x)\right) ds_1\, ds_2\, ds_3,\end{aligned} \tag{5.6}$$

and

$$\begin{aligned}\Psi^{\mathbf{g}}_{[[X_1,X_2],[X_3,X_4]]}&(t_1,t_2,t_3,t_4))(x) - x \\ &= \int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\int_0^{t_4} [[g_1,g_2],[g_3,g_4]]^{(t_1,t_3,s_4,s_1,s_2,s_3)} \\ &\quad \left(\Psi^{\mathbf{g}}_{[[X_1,X_2],[X_3,X_4]]}(t_1,t_2,t_3,s_4)(x)\right) ds_1\, ds_2\, ds_3\, ds_4,\end{aligned}$$

respectively.

One can replace the twisting

$$\theta^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}$$

with the corresponding basic twisting

$$\theta_{\mathfrak{b}}{}^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})},$$

so to obtain the (parametrized) vector field

$$y \mapsto B_{\mathfrak{b}}^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}(\mathbf{g})(y)$$
$$:= B^{\theta_{\mathfrak{b}}^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}}(\mathbf{g})(y).$$

By Proposition 2.26 and Theorem 5.8 one gets an integral formula involving only basic twistings:

Theorem 5.10. (Integral representation with basic twistings) *For every m-tuple $(t_1,\ldots,t_m)\in\mathbb{R}^m$ one has*

$$\begin{aligned}&\Psi_B^{\mathbf{g}}(t_1,\ldots,t_m)(x)-x\\ &\quad=\int_0^{t_1}\cdots\int_0^{t_m}B_{\mathfrak{b}}^{(t_1,\ldots,t_{m_1-1},t_{m_1+1},\ldots,t_{m-1},s_m,s_1,\ldots,s_{m-1})}(\mathbf{g})(x)\\ &\quad\big(\Psi_B^{\mathbf{g}}\left(t_1,\ldots,t_{m-1},s_m\right)(x)\big)\,ds_1\ldots ds_m.\end{aligned} \tag{5.7}$$

6. Some estimates

In this section we state and prove a few technical results which have been utilized in the proof of the Theorem 3.7.

Lemma 6.1. *Let k,n be positive integers. Let f be a vector field on $\mathbb{R}^n$ of class C^k and having compact support. Then, for every $t\in\mathbb{R}$, the flow map $x\mapsto e^{tf}(x)$ is (well-defined and) of class C^k. In addition, there exists a constant L depending only on k and n such that, for every $\ell=0,\ldots,k$,*

$$|D_x^\ell\big(e^{tf}(x)-\iota\big)(x,t)|\le L\|f\|_k|t| \tag{6.1}$$

for all $(x,t)\in\mathbb{R}^n\times\mathbb{R}$ verifying $\|f\|_k|t|\le 1$.

The regularity of the flow-map is a classical result, while the proof of (6.1) can be easily deduced from standard arguments of the theory of o.d.e.'s.

Lemma 6.2. *Let $\mathbb{R}^n\ni x\mapsto A(x)\in\mathbb{R}^{n\times n}$ and $\mathbb{R}^n\ni x\mapsto g(x)\in\mathbb{R}^n$ be, respectively, a matrix-valued smooth function and a vector-valued smooth function. Then for all $k\in\mathbb{N}_0$*

$$\|D^k(A\cdot g)\|_\infty\le(k+1)2^k\binom{k+n-1}{k}\|A\|_k\|g\|_k. \tag{6.2}$$

The norm $\|A\|_k$ of a matrix- or a vector-valued function $\mathbb{R}^n\supset\Omega\ni x\mapsto A(x)=(A_{ij}(x))_{1\le i\le m,\,1\le j\le n}\in\mathbb{R}^{m\times n}$, $m,n\in\mathbb{N}$, defined on some open set Ω, is by definition

$$\|A\|_k:=\sum_{\ell=0}^k\|D^\ell A\|_\infty,$$

where, for $0\le\ell\le k$,

$$\|D^\ell A\|_\infty:=\sup_{x\in\Omega}\left(\sum_{|\alpha|=\ell}\sum_{\substack{1\le i\le m\\1\le j\le n}}|\partial^\alpha A_{ij}(x)|\right).$$

Clearly, with these notation, $\|A\|_0$ *and* $\|A\|_\infty$ *denote the same quantity.*

Proof. For $k=0$ the proof is easy so let $k \geq 1$. We have

$$\begin{aligned}\sum_{|\alpha|=k}\left|\partial^\alpha\left(\sum_{j=1}^n\binom{\alpha}{\beta}A_{ij}(x)g^j(x)\right)\right| &\leq \sum_{i,j=1}^n\sum_{|\alpha|=k}\sum_{\beta\leq\alpha}\binom{\alpha}{\beta}\left|\partial^{\alpha-\beta}A_{ij}(x)\partial^\beta g^j(x)\right| \\ &\leq \sum_{|\alpha|=k}\sum_{\beta\leq\alpha}\binom{\alpha}{\beta}\|A\|_k\|g\|_k \\ &= 2^k\binom{k+n-1}{k}\|A\|_k\|g\|_k,\end{aligned}$$

where the last equality follows from the combinatorial identity (1.4). □

Lemma 6.3. *Let* $g\colon \Omega\subset\mathbb{R}^n\to\mathbb{R}^p$, $f\colon D\subset\mathbb{R}^p\to\mathbb{R}^q$ *be functions of class* C^k *and* $g(\Omega)\subset D$, *where* $n,p,q,k\in\mathbb{N}$ *and* Ω, D *are open sets. Then* $f\circ g$ *is of class* C^k *as well, and moreover, the estimate*

$$\|D^k(f\circ g)\|_0 \leq k(k+1)^{(n+p)(k+1)}\,\|Df\|_{k-1}(1+\|Dg\|_{k-1})^k. \tag{6.3}$$

holds true. In particular,

$$\|f\circ g\|_k \leq (k+1)^{(n+p)(k+1)+1}\,\|f\|_k(1+\|Dg\|_{k-1})^k. \tag{6.4}$$

Proof. By FÃăa di Bruno's formula (see (1.5)), for every natural number $k\geq 1$, we have

$$\begin{aligned}|D^k(f\circ g)| &= \sum_{|\alpha|=k}|\partial^\alpha(f\circ g)| \\ &= \sum_{|\alpha|=k}\sum_{|\gamma|=1}^{|\alpha|}|(\partial^\gamma f)\circ g|\sum_{s=1}^{|\alpha|}\sum_{P_s(\alpha,\gamma)}\alpha!\prod_{j=1}^s\frac{|(\partial^{\beta^j}g)^{\gamma^j}|}{\gamma^j!(\beta^j!)^{|\gamma^j|}},\end{aligned} \tag{6.5}$$

where, for all $\alpha\in\mathbb{N}_0^n$, $\gamma\in\mathbb{N}_0^p$ with $1\leq|\gamma|\leq|\alpha|$ and for all $s=1,\dots,k$, the set $P_s(\alpha,\gamma)\subset(\mathbb{N}_0^n\times\mathbb{N}_0^p)^s$ is defined by (1.6).

Let us observe that, provided $|\beta^j|\leq k$ for whatever γ^j, one has

$$|(\partial^{\beta^j}g)^{\gamma^j}| \leq (1+|\partial^{\beta^j}g_j|)^{|\gamma^j|} \leq (1+|D^{|\beta^j|}g|)^{|\gamma^j|} \leq (1+\|Dg\|_{k-1})^{|\gamma^j|}.$$

Hence, for every $(\beta^1,\dots,\beta^s;\gamma^1,\dots,\gamma^s)\in P_s(\alpha,\gamma)$, since $\sum_{j=1}^s|\gamma^j| = |\gamma|$, we have

$$\prod_{j=1}^s|(\partial^{\beta^j}g)^{\gamma^j}| \leq (1+\|Dg\|_{k-1})^{|\gamma|}.$$

Continuing with the estimate in (6.5), we obtain

$$|D^k(f\circ g)| \leq C_k\,\|Df\|_{k-1}(1+\|D^g\|_{k-1})^k, \tag{6.6}$$

where

$$C_k := \sum_{|\alpha|=k}\sum_{|\gamma|=1}^{|\alpha|}\sum_{s=1}^{|\alpha|}\sum_{P_s(\alpha,\gamma)}1.$$

Since the cardinality $|P_s(\alpha,\gamma)|$ of $P_s(\alpha,\gamma)$ verifies

$$|P_s(\alpha,\gamma)| \le (|\alpha|+1)^{(n+p)s} \le (k+1)^{(n+p)k},$$

we get

$$C_k \le k(k+1)^{(n+p)(k+1)}, \tag{6.7}$$

which, together with (6.6), implies (6.3). □

In order to state Lemma 6.4 below we set, for every nonnegative integer k,

$$p_k(1) := a_k {b_k}^2 (n+2)^{2k}, \quad c_k := p_k(1)(k+1), 2^{2k+n-1} \tag{6.8}$$

and, for every natural number $m \ge 2$,

$$p_k(m) := (2c_{k+m})^{m-1} p_{k+m}(1)^m, \tag{6.9}$$

where a_k, b_k are the constants appearing, respectively, in estimates (6.2) and (6.4) for $p = n$, i.e.,

$$a_k = (k+1)2^k \binom{k+n-1}{k}, \quad b_k := (k+1)^{2n(k+1)+1}. \tag{6.10}$$

Lemma 6.4. *Let $m \ge 1$ be an integer. For every bracket B of degree m, integer $k \ge 0$, m-tuple $\mathbf{g} = (g_1, \dots, g_m) \in C^{B+k}$ and every twisting θ, one has*

$$\|B^\theta(\mathbf{g})\|_k \le p_k(m) \cdot \|\mathbf{g}\|^m_{B+k} \tag{6.11}$$

provided $\mathbf{d}_{B+k}(\theta,\, \iota) \le 1$.

Proof. We prove the result by induction on $m = deg(B)$.

Consider the case $m = 1$, so that $B = X_i$ for some $i \in \mathbb{N}$. Since the twisting θ consists of a single diffeomorphism, we write simply θ instead of $\theta(B)$. If $\mathbf{g} = (g_1, \dots, g_m)$ for some $m \ge i$, we have $B^\theta(\mathbf{g}) = \theta^\# g = \theta^\# g - g \circ \theta + g \circ \theta = \big(D(\theta^{-1} - Id_{\mathbb{R}}) \circ \theta\big) \cdot g \circ \theta + g \circ \theta$.

By Lemma 6.2, we obtain

$$\begin{aligned} \|B^\theta(\mathbf{g})\|_k &\le \| \big(D(\theta^{-1} - Id_{\mathbb{R}^n}) \circ \theta\big) \cdot (g \circ \theta) \|_k + \|g \circ \theta\|_k \\ &\le a_k \, \|\big(D(\theta^{-1} - Id_{\mathbb{R}^d})\big) \circ \theta\|_k \, \|g \circ \theta\|_k + \|g \circ \theta\|_k, \end{aligned} \tag{6.12}$$

while, by applying by Lemma 6.3, we get

$$\|D\big(\theta^{-1} - Id_{\mathbb{R}^n}\big) \circ \theta\|_k \le b_k \, \|D(\theta^{-1} - Id_{\mathbb{R}^n})\|_k (1 + \|D\theta\|_{k-1})^k, \tag{6.13}$$

$$\|g \circ \theta\|_k \le b_k \, \|g\|_k (1 + \|D\theta\|_{k-1})^k, \tag{6.14}$$

where a_k, b_k are the constants defined in (6.10). In addition

$$\|D\theta\|_{k-1} \le \|D(\theta - Id_{\mathbb{R}^n})\|_{k-1} + \|D(Id_{\mathbb{R}^n})\|_{k-1} \le \mathbf{d}_{B+k}(\theta, \iota) + n, \tag{6.15}$$

$$\|D(\theta^{-1} - Id_{\mathbb{R}^n})\| \le \mathbf{d}_{B+k+1}(\theta, \iota) \tag{6.16}$$

(see (2.3) for the definitions of $\mathbf{d}_{B+k}(\theta,\, \iota)$, $\mathbf{d}_{B+k+1}(\theta,\, \iota)$). Combining together the estimates (6.12), (6.13), (6.14), (6.15), (6.16) we deduce (6.11) with $p_k(1)$ given by (6.8).

Now let $m > 1$. We shall prove that, assuming the thesis to hold true for the brackets of degree $< m$, it holds true also for brackets of degree m. So let B be a bracket of $deg(B)=m$ and let (B_1,B_2) the canonical factorization of B, i.e., $B=[B_1,B_2]$. Let the subbrackets B_1, B_2 have degree m_1 and m_2, respectively, so that, in particular $m_1+m_2=m$. Let θ be a twisting of B verifying $\mathbf{d}_{B+k}(\theta,\mathbf{0})\le 1$. We denote by θ_i, $i=1,2$, the restrictions of θ to $Subb(B_i)$, respectively. Let $\Psi=\theta(B)$. Let also $\mathbf{g}\in C^{B+k}$. We have (see Definitions 2.22 and 2.19)

$$B^{\theta}(\mathbf{g})(x)=D\Psi^{-1}(\Psi(x))\cdot\left[B_1^{\theta_1}(\mathbf{g}),B_2^{\theta_2}(\mathbf{g})\right](\Psi(x)).$$

For the sake of brevity let us set $\hat g_1=B_1^{\theta_1}(\mathbf{g})$, $\hat g_2=B_2^{\theta_2}(\mathbf{g})$.

As in the case $m=1$[19] one has

$$\|B^{\theta}(\mathbf{g})\|_k\le p_k(1)\|[\hat g_1,\hat g_2]\|_k. \tag{6.17}$$

Since $[\hat g_1,\hat g_2]=D\hat g_2\cdot\hat g_1-D\hat g_1\cdot\hat g_2$, by Lemma 6.2, it follows that

$$\|[\hat g_1,\hat g_2]\|_k\le(k+1)2^{2k+n-1}\left(\|\hat g_1\|_{k+1}\|\hat g_2\|_k+\|\hat g_2\|_{k+1}\|\hat g_1\|_k.\right) \tag{6.18}$$

By the inductive hypothesis, for $i=1,2$,

$$\begin{aligned}\|\hat g_i\|_k&\le p_k(m_i)\ \left(\|\mathbf{g}\|_{B_i+k}\right)^{m_i},\\ \|\hat g_i\|_{k+1}&\le p_{k+1}(m_i)\ \left(\|\mathbf{g}\|_{B_i+k+1}\right)^{m_i}\end{aligned}$$

since $\mathbf{d}_{B_i+k}(\theta_i,\,\iota)\le\mathbf{d}_{B_i+k+1}(\theta_i,\,\iota)\le\mathbf{d}_{B+k}(\theta,\,\iota)\le 1$.

By these estimates and by (6.18), since $\|\mathbf{g}\|_{B_i+k}\le\|\mathbf{g}\|_{B_i+k+1}\le\|\mathbf{g}\|_{B+k}$, $\mathbf{d}_{B_i+k+1}(\theta_i,\,\iota)\le\mathbf{d}_{B_i+k+2}(\theta_i,\,\iota)\le\mathbf{d}_{B+k+1}(\theta,\,\iota)$, it follows that

$$\|[\hat g_1,\hat g_2]\|_k\le(k+1)2^{2k+n-1}\left(p_{k+1}(m_1)p_k(m_2)+p_{k+1}(m_2)p_k(m_1)\right)\left(\|\mathbf{g}\|_{B+k}\right)^{m_1+m_2}$$

From this estimate, (6.17), $m_1+m_2=m$, and the fact that

$$p_k(m)\ge p_k(1)(k+1)2^{2k+n-1}\left(p_{k+1}(m_1)p_k(m_2)+p_{k+1}(m_2)p_k(m_1)\right) \qquad \square$$

it follows that (6.11) holds true.

6.1. Shift-brackets and set-valued brackets

Consider the function $\gamma\colon[0,+\infty[\,\to[0,+\infty[$ defined by setting

$$\gamma(\rho):=\sup_{\Theta}\operatorname{dist}\left(B^{\Theta}(\mathbf{f})(x),\,B_{set}(\mathbf{f})(x_*)\right),\quad\forall\rho\ge 0$$

where the supremum is taken over all *basic shift-twistings* Θ *of* B and all points $x\in\mathbb{R}^n$ such that

(a) x lies in the domain of $B^{\Theta}(\mathbf{f})$ (which has full measure, since $\mathbf{f}\in C^{B-1,1}$);
(b) $|x-x_*|+\mathbf{d}_B(\Theta,\iota)\le\rho$.

Namely, $\gamma(\cdot)$ gives the largest error in the evaluation of the set-valued bracket at x_* by any Θ-shifted bracket when the latter is evaluated at x such that $|x-x_*|+\mathbf{d}_B(\Theta,\iota)\le\rho$.

[19] Notice that if a twisting, say η, of a degree-1 bracket X_i maps X_i into the diffeomorphism Ψ, one has $\mathbf{d}_{X_i+k}(\eta,\,\iota)\le\mathbf{d}_{X_i+k}(\theta,\,\iota)\le 1)$ for every k.

Lemma 6.5. *The function* γ *is a modulus, i.e.,*

$$\gamma(0^+) := \lim_{\rho\to 0^+} \gamma(\rho) = 0.$$

$$\mathbf{d}_B(\Theta, \iota) = \sum_{\substack{S \in Sbb(B) \\ S \ \ is\ basic}} |\Theta(S)(0)|,$$

Proof. Though this lemma is a nothing but a manifestation of the definition and upper semi-continuity of the set-valued map $x \mapsto B^{\Theta}(\mathbf{f})(x)$, for the sake of self-consistency we provide a proof. Clearly the right-limit $\gamma(0^+)$ exists, for γ is a nondecreasing function. If we had $\gamma(0^+) > 0$, then we could find an $\varepsilon > 0$ and a sequence $\rho_j \to 0^+$ such that $\gamma(\rho_j) > \varepsilon > 0$ for all $j \in \mathbb{N}$. Therefore, for each $j \in \mathbb{N}$, there would exist a basic shift-twisting Θ_j and $x_j \in \mathbb{R}^n$ such that

$$\text{dist}\left(B^{\Theta_j}(\mathbf{f})_{cl}(x_j),\, B_{set}(\mathbf{f})(x_*)\right) > \varepsilon, \tag{6.19}$$

while

$$|x - x_*| + \mathbf{d}_B(\Theta_j, \iota) \leq \rho_j.$$

Clearly the sequence $\left\{B^{\Theta_j}(\mathbf{f})_{cl}(x_j)\right\}_{j\in\mathbb{N}}$ is bounded. Thus, there exists a subsequence, still denoted $\left\{B^{\Theta_j}(\mathbf{f})_{cl}(x_j)\right\}_{j\in\mathbb{N}}$, such that $B^{\Theta_j}(\mathbf{f})(x_j) \to v$ as $j \to \infty$. We now slightly change the sequence (Θ_j) into $(\tilde{\Theta}_j)$ by setting $\tilde{\Theta}_j(S) = \Theta_j(S) + \tau_{x_j - x_*}$ for all $j \in \mathbb{N}$ and any basic subbracket S of B. One has $\mathbf{d}_B(\tilde{\Theta}_j, \iota) \to 0$ as $j \to \infty$. Therefore, by the definition of set-valued bracket (see Def. 3.1), $v \in B(\mathbf{f})(x_*)$. This gives a contradiction, for, passing to the limit in (6.19) one gets

$$0 = \text{dist}\left(v,\, B_{set}(\mathbf{f})(x_*)\right) \geq \varepsilon > 0.$$

□

6.2. Approximation of twisted brackets by shifted brackets

If B is a bracket, for any twisting θ of B and any m-tuple of vector fields $\mathbf{g} = (g_1, \ldots, g_m)$ of class C^B, we have

$$B^{\theta_x}(\mathbf{g})(x), \quad x \in \mathbb{R}^n, \tag{6.20}$$

is obtained by shifting $B(\mathbf{g})$ through the shift-twisting approximation θ_x of θ based on x (see Definition 2.13). Informally, we can say that $B^{\theta_x}(\mathbf{g})(x)$ is obtained from the expression of the twisted bracket $B^{\theta}(\mathbf{g})$ evaluated at x by removing all terms containing derivatives of $\theta(S)$ and $\theta(S)^{-1}$.

If (B_1, B_2) is the factorization of B and θ_1, θ_2 are the restrictions of θ on, resp., $Subb(B_1)$, $Subb(B_2)$, we have

$$B^{\theta_x}(\mathbf{g})(x) = \left[B_1^{(\theta_1)_x}(\mathbf{g}),\, B_2^{(\theta_2)_x}(\mathbf{g})\right](\theta(B)(x)).$$

In order to state Proposition 6.6, we need to define the following numbers $c_k(m)$, $m \in \mathbb{N}$, $k \in \mathbb{Z}, k \geq 0$:

$$c_k(1) := 2^k(k-1)k^{2nk}(2+n)^{k-1} + k(k+1)^{2n(k+1)}(n+2)^{n+k}$$

$$c_k(m) := \max\left\{2^k\binom{k+n-1}{k}2c_{k+1}(m_1)c_{k+1}(m_2)+c_k(1)p_k(m)\right\} \quad \forall m \geq 2, \tag{6.21}$$

where the maximum is computed over all pairs $(m_1, m_2) \in \mathbb{N}^2$ such that $m_1 + m_2 = m$.

Proposition 6.6. *Let us choose $m \in \mathbb{N}$ and an integer $k \geq 0$. Let B be a bracket of degree m and let θ be a basic twisting verifying*

$$\mathbf{d}_{B+k+1}(\theta,\, \iota) \leq 1.$$

Then, for each m-tuple $\mathbf{g} = (g_1, \ldots, g_m) \in C^{B+k}$, one has

$$\left|D^\ell B^\theta(\mathbf{g})(x) - D^\ell B^{\theta_x}(\mathbf{g})(x)\right| \leq c_k(m)\|\mathbf{g}\|^m_{B+k} \cdot \mathbf{d}_{B+k}\left(\theta, \iota\right) \tag{6.22}$$

for all $0 \leq \ell \leq k$, where $c_k(m)$ has been defined in (6.21), and distances $\mathbf{d}_{B+k}\left(\theta, \iota\right)$ are defined as in (2.3) on a compact neighborhood K of x.

Proof. The result will be proved by induction on the degree m of B. First let us examine the case when m is equal to 1.

If $m = deg(B) = 1$, then $B = X_i$ for some natural number i. Since the twisting θ consists of one single diffeomorphism, it is not confusing to write θ instead of $\theta(B)$. In addition we observe that

$$\partial^\alpha\big(B^{\theta_x}(\mathbf{g})\big)(x) = \partial^\alpha g_i(\theta(x)).$$

Since by Leibniz's rule

$$\partial^\alpha B^\theta(\mathbf{g}) = \sum_{0\neq\beta\leq\alpha}\binom{\alpha}{\beta}\partial^{\alpha-\beta}\big(D\theta^{-1}\circ\theta\big)\cdot\partial^\beta\big(g\circ\theta\big) + \partial^\alpha\big(g\circ\theta\big),$$

by Lemma 6.3, and in particular, by estimate (6.3), we have

$$\begin{aligned}
&\left|\partial^\alpha B^\theta(x) - \partial^\alpha\big(g\circ\theta\big)(x)\right| \\
&\quad \leq 2^k(k-1)k^{2nk}\|D\theta^{-1}\|_{k-1}\left(1+\|D\theta\|_{k-1}\right)^{k-1}\|\mathbf{g}\|_k \\
&\quad \leq 2^k(k-1)k^{2nk}(2+n)^{k-1}\,\mathbf{d}_{B+k+1}(\theta,\iota)\,\|\mathbf{g}\|_k,
\end{aligned} \tag{6.23}$$

where we have used the fact $\|D\theta\|_{k-1} \leq \mathbf{d}_{B+k+1}(\theta,\iota) + n \leq 1+n$.

By Fàa di Bruno's formula (1.5)

$$\begin{aligned}
&\partial^\alpha\big(g\circ\theta\big)(x) - \partial^\alpha g(\theta(x)) \\
&\quad = \sum_{|\gamma|=|\alpha|}(\partial^\gamma g)\circ\theta\sum_{p_{|\alpha|}(\alpha,\gamma)}\alpha!\left(\prod_{j=1}^{|\alpha|}\frac{\big(\partial^{\beta^j}\theta\big)^{\gamma^j}}{\gamma^j!(\beta^j!)^{\gamma^j!}} - \prod_{j=1}^{|\alpha|}\frac{\big(\partial^{\beta^j}Id_{\mathbb{R}^n}\big)^{\gamma^j}}{\gamma^j!(\beta^j!)^{\gamma^j!}}\right).
\end{aligned} \tag{6.24}$$

By the elementary inequality

$$\left|\prod_{i=1}^n a_i - \prod_{i=1}^n b_i\right| \leq \Big(\max_{i=1,\ldots,n}(|a_i|+|b_i|)\Big)^{n-1}\sum_{i=1}^n|a_i-b_i| \tag{6.25}$$

valid for all $(a_1,\ldots,a_n,b_1,\ldots,b_n)\in\mathbb{R}^{2n}$, and from (6.24), we deduce that

$$\begin{aligned}&\left|\partial^\alpha\big(g\circ\theta\big)(x)-\partial^\alpha g(\theta(x))\right|\\ &\quad\le\sum_{|\gamma|=1}^{|\alpha|}\left|\partial^\gamma g\big(\theta(x)\big)\right|\sum_{s=1}^{|\alpha|}\sum_{P_s(\alpha,\gamma)}\alpha!\left|\prod_{j=1}^{s}\frac{\big(\partial^{\beta^j}\theta(x)\big)^{\gamma^j}}{\gamma^j!(\beta^j!)^{\gamma^j!}}-\prod_{j=1}^{s}\frac{\big(\partial^{\beta^j}Id_{\mathbb{R}^n}\big)^{\gamma^j}}{\gamma^j!(\beta^j!)^{\gamma^j!}}\right|\\ &\quad\le\|\mathbf{g}\|_{B+k+1}\left(\mathbf{d}_{B+k+1}(\theta,\,\iota)+n\right)^{n-1}\cdot\\ &\qquad\times\sum_{|\gamma|=1}^{|\alpha|}\sum_{s=1}^{|\alpha|}\sum_{P_s(\alpha,\gamma)}\sum_{j=1}^{n}\left|\big(\partial^{\beta^j}\theta(x)\big)^{\gamma^j}-\big(\partial^{\beta^j}Id_{\mathbb{R}^n}\big)^{\gamma^j}\right|.\end{aligned}$$

Since

$$\left|x^\gamma-y^\gamma\right|\le(|x|+|y|+1)^{|\gamma|}\quad\text{for all } x,y\in\mathbb{R}^n,\ \gamma\in\mathbb{N}_0^n,$$

we obtain

$$\begin{aligned}&\left|\partial^\alpha\big(g\circ\theta\big)(x)-\partial^\alpha g(\theta(x))\right|\\ &\quad\le\|\mathbf{g}\|_{B+k+1}\left(\mathbf{d}_{B+k+1}(\theta,\,\iota)+n\right)^{n-1}\left(\mathbf{d}_{B+k+1}(\theta,\,\iota)+n+1\right)^{k}\\ &\qquad\cdot\mathbf{d}_{B+k+1}(\theta,\iota)\sum_{|\gamma|=1}^{|\alpha|}\sum_{s=1}^{|\alpha|}\sum_{P_s(\alpha,\gamma)}\sum_{j=1}^{n}1.\end{aligned}$$

As in the proof of Lemma 6.3 and using the fact that $\mathbf{d}_{B+k+1}(\theta,\iota)\le 1$ we obtain

$$\left|\partial^\alpha\big(g\circ\theta\big)(x)-\partial^\alpha g(\theta(x))\right|\le k(k+1)^{2n(k+1)}(n+2)^{n+k}\,\mathbf{d}_{B+k+1}(\theta,\iota)\,\|\mathbf{g}\|_{B+k+1}. \tag{6.26}$$

From estimates (6.23) and (6.26) we deduce (6.22) for $m=1$.

So let us assume the result true for all brackets of degree $\le m-1$, where $m\ge 2$, and prove it for brackets of degree m. Let B be a bracket of degree m with factorization (B_1,B_2) for some brackets B_1,B_2 with $deg(B_1)=m_1$, $deg(B_2)=m_2$, $m_1+m_2=m$. Let $\mathbf{g}=(g_1,\ldots,g_m)\in C^{B+k}$, let θ be a twisting of B and let θ_x denote the corresponding shift-twisting of B, see Definition 2.13.

For $i=1,2$, let θ_i denote the restriction of θ to B_i, $h_i=B_i^{\theta_i}(\mathbf{g})$ and $\tilde h_i=B_i^{(\theta_i)_x}(\mathbf{g})$.

Clearly, $h_i,\tilde h_i\in C^{k+1}$ and by the inductive hypothesis, for all $j\in\{0,\ldots,k+1\}$,

$$\left|D^jh_i(x)-D^j\tilde h_i(x)\right|\le c_{k+1}(m_i)\,\mathbf{d}_{B_i+k+1}(\theta_i,\iota) \tag{6.27}$$

for $i=1,2$ because $\mathbf{d}_{B_i+k+1}(\theta_i,\iota)\le\mathbf{d}_{B+k+2}(\theta_i,\iota)\le\mathbf{d}_{B+k+1}(\theta,\iota)\le 1$.

Since we can write

$$\begin{aligned}[h_1,h_2]-[\tilde h_1,\tilde h_2]&=[h_1-\tilde h_1,h_2]+[\tilde h_1,h_2-\tilde h_2]\\ &=Dh_2\cdot(h_1-\tilde h_1)-D(h_1-\tilde h_1)\cdot h_2+D(h_2-\tilde h_2)\cdot\tilde h_1\\ &\quad-D\tilde h_1\cdot(h_2-\tilde h_2),\end{aligned}$$

using Leibniz's product differentiation rule and (6.27) we conclude that for all $j = 0, \ldots, k$

$$\begin{aligned}&\left|D^j[h_1,h_2](x) - D^j[\tilde{h}_1,\tilde{h}_2](x)\right|\\ &\quad\le 2^k\binom{k+n-1}{k} 2\, c_{k+1}(m_1)\, c_{k+1}(m_2)\, \mathbf{d}_{B_i+k+1}(\theta,\iota).\end{aligned}$$

Thus we have

$$\begin{aligned}&\left|D^j[h_1,h_2](\Phi(x)) - D^j B^{\theta_x}(\mathbf{g})(x)\right|\\ &\quad\le 2^k\binom{k+n-1}{k} 2c_{k+1}(m_1)c_{k+1}(m_2)\, \mathbf{d}_{B_i+k+1}(\theta,\iota).\end{aligned} \tag{6.28}$$

From the case $m = 1$ we have

$$\left|D^j\left(\Phi^{\#}[h_1,h_2]\right)(x) - D^j[h_1,h_2](\Phi(x))\right| \le c_k(1)\, \mathbf{d}_{B+k+1}(\theta,\iota)\, \|[h_1,h_2]\|_k.$$

This together with Lemma 6.4 gives

$$\left|D^j\left(\Phi^{\#}[h_1,h_2]\right)(x) - D^j[h_1,h_2](\Phi(x))\right| \le c_k(1)p_k(m)\, \mathbf{d}_{B+k+1}(\theta,\iota)\, \|\mathbf{g}\|_{B+k}. \tag{6.29}$$

Estimates (6.28), (6.29) imply (6.22). □

6.3. Estimating the noncommutativity of bracket twisting and vector field regularization

For every pair of integers $m \ge 1$ and $k \ge 0$, let us define the numbers $r_k(m), s_k(m)$, and $w_k(m)$ inductively by setting:

(i) $s_k(1) = r_k(1) = 0$;
(ii) $s_k(2) = r_k(2)$;
(iii) $w_k(1) = w_k(2) = 0$;
(iv) $r_0(m) = \max\{4p_1(m_1)\cdot p_1(m_2) \ : \ (m_1,m_2)\in\mathbb{N}^2,\ m_1+m_2 = m\}$ if $m \ge 2$.

Furthermore, for every $m \ge 2$ and every $k \ge 0$, we set

$$r_k(m) := \max\left\{n^2\cdot 2^k\cdot\binom{k+n-1}{k}\cdot\left(4p_{k+1}(m_1)\cdot p_{k+1}(m_2)\right)\right\},$$

where the maximum is taken over all pairs $(m_1,m_2)\in\mathbb{N}^2$ such that $m_1+m_2 = m$, while, for any $m \ge 3$, $w_k(m)$ $s_k(m)$ are defined recursively by

$$\begin{aligned}w_k(B) := \max\Bigg\{& n2^k\binom{k+n-1}{k}\cdot(p_{k+1}(m_2)\cdot s_k(m_1)\\ &+p_k(m_2)\cdot s_{k+1}(m_1) + p_k(m_1)\cdot s_{k+1}(m_2) + p_{k+1}(m_1)\cdot s_k(m_2))\Bigg\},\end{aligned} \tag{6.30}$$

where the maximum is taken over all pairs $(m_1,m_2)\in\mathbb{N}^2$ such that $m_1+m_2 = m$,

$$s_k(m) := w_k(m) + r_k(m);$$

here the $p_k(m)$'s are defined as in (6.8), (6.9).

Lemma 6.7. *For any bracket B of degree m, any integer $k \geq 0$, any basic shift-twisting Θ and for any m-tuples $\mathbf{f} = (f_1, \dots, f_\mu)$ of vector fields on $\mathbb{R}^n$ of class $C^{B-1+k,1}$ and with compact support, one has*

$$\left\| D^k B^\Theta(\mathbf{f}^\zeta) - \left(D^k B^\Theta(\mathbf{f})\right)^\zeta \right\| \leq s_k(m) \cdot \|\mathbf{f}\|^m_{B+k} \cdot \zeta \quad \forall \zeta > 0. \tag{6.31}$$

Proof. We shall prove the result by recursion on the degree of B. The first step of induction relies on basic properties of convolution. Indeed, assume $m = deg(B) = 1$, namely $B = X_j$ for some $j = 1 \dots, m$. Let $\mathbf{f} \in C^{B-1+k,1}$ for some integer $k \geq 0$, which means nothing but $B(\mathbf{f}) = f_j$ and $f_j \in C^{k,1}$. The diffeomorphism $\Theta(B)$ is a translation, so that $\Theta(B)(x) = x + h$ for all $x \in \mathbb{R}^n$, where $h := \Theta(B)(0)$. Since both translation and differentiation commute with convolution, one has

$$\begin{aligned} D^k B^\Theta(\mathbf{f}^\zeta)(x) &= D^k(f_j^\zeta(\cdot + h))(x) = (D^k f_j^\zeta)(x + h) \\ &= (D^k f_j(\cdot + h))^\zeta(x) = (D^k B^\Theta(\mathbf{f}))^\zeta(x) \end{aligned}$$

for every $x \in \mathbb{R}$. Hence (6.31) holds true when $m = 1$, with $s_k(1) = 0$ for all $k \geq 0$.

Now let m be an integer > 1 and let us assume the result true for all brackets of degree $\leq m - 1$: we shall prove that it holds true for brackets of degree m as well. Let B a bracket of degree m and let $B = (B_1, B_2)$ be its factorization. Let $m_1 = deg(B_1)$, $m_2 = deg(B_2)$; clearly $m_1 + m_2 = m$. Since $\mathbf{f}$ is of class $C^{B-1+k,1}$, $\mathbf{f}$ is also of class $C^{B_i+k,1}$ for every $i = 1, 2$. Therefore, in view of the inductive hypothesis, for every integer $k \geq 0$, $i = 1, 2$, one has

$$\begin{aligned} &\left\| D^\ell B_i^{\Theta_i}((\mathbf{f}^\zeta)) - D^\ell \left(B_i^{\Theta_i}(\mathbf{f})\right)^\zeta \right\| \leq s_\ell(m_i) \Big\| \cdot \Big\| \mathbf{f} \Big\|^{m_i}_{B+\ell} \cdot \zeta \\ &\quad i = 1, 2, \ \ell = 0, \dots, k+1 \end{aligned} \tag{6.32}$$

for all $\zeta > 0$, where $s_k(m_i)$ are positive constants depending only on m_i, k, n, and Θ_i is the shift-twisting of B_i obtained by restricting B to $Subb(B_i)$.

Furthermore,[20]

$$\begin{aligned} B^\Theta(\mathbf{f})(x) &= D(\Theta(B))^{-1}(\Theta(B(x)) \cdot [B_1^\Theta(\mathbf{f}), B_2^\Theta(\mathbf{f})](\Theta(B)(x)) \\ &= [B_1^{\Theta_1}(\mathbf{f}), B_2^{\Theta_2}(\mathbf{f})](x + h), \quad \forall x \in \mathbb{R}^n, \end{aligned}$$

where $h := \Theta(B)(0)$, and, similarly,

$$B^\Theta(\mathbf{f}^\zeta)(x) = [B_1^{\Theta_1}(\mathbf{f}^\zeta), B_2^{\Theta_2}(\mathbf{f}^\zeta)](x + h) \quad \forall x \in \mathbb{R}^n.$$

For every integer $k \geq 0$, let us consider the inequality

$$\begin{aligned} &\left\| D^k B^\Theta(\mathbf{f}^\zeta) - D^k \left(B^\Theta(\mathbf{f})\right)^\zeta \right\| \\ &\quad = \left\| D^k [B_1^{\Theta_1}(\mathbf{f}^\zeta), B_2^{\Theta_2}(\mathbf{f}^\zeta)] - D^k [B_1^{\Theta_1}(\mathbf{f}), B_2^{\Theta_2}(\mathbf{f})]^\zeta \right\| \end{aligned}$$

---

[20] We remind that $\Theta(B)$ is a translation, for Θ is a shift-twisting. Actually, since Θ is a also basic twisting, $\Theta(B)$ is the identity on $\mathbb{R}^n$ (so $h = 0$) as soon as $deg(B) = m > 2$.

$$\leq \underbrace{\left\|D^k[B_1^{\Theta_1}(\mathbf{f}^\zeta), B_2^{\Theta_2}(\mathbf{f}^\zeta)] - D^k[(B_1^{\Theta_1}(\mathbf{f}))^\zeta, (B_2^{\Theta_2}(\mathbf{f}))^\zeta]\right\|}_{(\mathbf{I})}$$
$$+ \underbrace{\left\|D^k[(B_1^{\Theta_1}(\mathbf{f}))^\zeta, (B_2^{\Theta_2}(\mathbf{f}))^\zeta] - D^k[B_1^{\Theta_1}(\mathbf{f}), B_2^{\Theta_2}(\mathbf{f})]^\zeta\right\|}_{(\mathbf{II})} \tag{6.33}$$

Claim: one has

$$\begin{aligned}(\mathbf{I}) &= \left\|D^k[B_1^{\Theta_1}(\mathbf{f}^\zeta), B_2^{\Theta_2}(\mathbf{f}^\zeta)] - D^k[(B_1^{\Theta_1}(\mathbf{f}))^\zeta, (B_2^{\Theta_2}(\mathbf{f}))^\zeta]\right\| \\ &\leq w_k(m) \cdot \left\|\mathbf{f}\right\|_{B+k}^m \cdot \zeta,\end{aligned} \tag{6.34}$$

and

$$\begin{aligned}(\mathbf{II}) &= \left\|D^k[(B_1^{\Theta_1}(\mathbf{f}))^\zeta, (B_2^{\Theta_2}(\mathbf{f}))^\zeta] - D^k[B_1^{\Theta_1}(\mathbf{f}), B_2^{\Theta_2}(\mathbf{f})]^\zeta\right\| \\ &\leq r_k(m) \cdot \left\|\mathbf{f}\right\|_{B+k}^m \cdot \zeta\end{aligned} \tag{6.35}$$

where $w_k(m), r_k(m) \geq 0$ has been defined above.

To prove (6.34), observe that by Leibniz rule we have

$$\begin{aligned}(\mathbf{I}) &= \left\|D^k[B_1^{\Theta_1}(\mathbf{f}^\zeta), B_2^{\Theta_2}(\mathbf{f}^\zeta)] - D^k[(B_1^{\Theta_1}(\mathbf{f}))^\zeta, (B_2^{\Theta_2}(\mathbf{f}))^\zeta]\right\| \\
&= \sum_{|\alpha|=k} \Bigg\|\partial^\alpha \Bigg(DB_2^{\Theta_2}(\mathbf{f}^\zeta) \cdot (B_1^{\Theta_1}(\mathbf{f}^\zeta) - (B_1^{\Theta_1}(\mathbf{f}))^\zeta) \\
&\quad - D\left(B_1^{\Theta_1}(\mathbf{f}^\zeta\right) - \left(B_1^{\Theta_1}(\mathbf{f}))^\zeta\right) \cdot B_2^{\Theta_2}(\mathbf{f}^\zeta) \\
&\quad + D\left(B_2^{\Theta_2}(\mathbf{f}^\zeta) - (B_2^{\Theta_2}(\mathbf{f}))^\zeta\right) \cdot (B_1^{\Theta_1}(\mathbf{f}))^\zeta - D\left((B_1^{\Theta_1}(\mathbf{f}))^\zeta\right) \\
&\quad \cdot \left(B_2^{\Theta_2}(\mathbf{f}^\zeta) - (B_2^{\Theta_2}(\mathbf{f}))^\zeta\right)\Bigg)\Bigg\| \\
&\leq \sum_{|\alpha|=k} \Bigg\| \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \Bigg(\partial^{\alpha-\beta}\left(DB_2^{\Theta_2}(\mathbf{f}^\zeta\right) \cdot \partial^\beta\left(B_1^{\Theta_1}(\mathbf{f}^\zeta) - (B_1^{\Theta_1}(\mathbf{f}))^\zeta\right) \\
&\quad - \partial^{\alpha-\beta}\left(D\left(B_1^{\Theta_1}(\mathbf{f}^\zeta\right) - (B_1^{\Theta_1}(\mathbf{f}))^\zeta\right) \cdot \partial^\beta\left(B_2^{\Theta_2}(\mathbf{f}^\zeta\right) \\
&\quad + \partial^{\alpha-\beta}\left(D\left(B_2^{\Theta_2}(\mathbf{f}^\zeta\right) - (B_2^{\Theta_2}(\mathbf{f}))^\zeta\right) \cdot \partial^\beta\left((B_1^{\Theta_1}(\mathbf{f}))^\zeta\right)\end{aligned}$$

$$
\begin{aligned}
&\quad - \partial^{\alpha-\beta}\left(D\left(B_1^{\Theta_1}(\mathbf{f})\right)^{\zeta}\right)\cdot \partial^{\beta}\left(B_2^{\Theta_2}(\mathbf{f}^{\zeta}) - (B_2^{\Theta_2}(\mathbf{f}))^{\zeta}\right)\Bigg)\Bigg\| \\
&= \sum_{|\alpha|=k}\sum_{\beta\le\alpha}\binom{\alpha}{\beta}\Bigg(p_{k+1}(m_2)\cdot\Big\|\mathbf{f}\Big\|^{m_2}_{B_2+k+1}\cdot s_k(m_1)\cdot\Big\|\mathbf{f}\Big\|^{m_1}_{B_1+k} \\
&\quad + p_k(m_2)\Big\|\mathbf{f}\Big\|^{m_2}_{B_2+k} s_{k+1}(m_1)\Big\|\mathbf{f}\Big\|^{m_1}_{B_1+k+1} \\
&\quad + p_k(m_1)\cdot\Big\|\mathbf{f}\Big\|^{m_2}_{B_1+k}\cdot s_{k+1}(m_2)\cdot\Big\|\mathbf{f}\Big\|^{m_1}_{B_2+k+1} \\
&\quad + p_{k+1}(m_1)\cdot\Big\|\mathbf{f}\Big\|^{m_2}_{B_1+k+1}\cdot s_k(m_2)\cdot\Big\|\mathbf{f}\Big\|^{m_1}_{B_2+k}\Bigg)\cdot\zeta \\
&\le n\binom{k+n-1}{k}2^k\Big(p_{k+1}(m_2)+p_k(m_2)+p_k(m_1)+p_{k+1}(m_1)\Big) \\
&\quad \cdot\Big\|\mathbf{f}\Big\|^{m}_{B+k}\cdot\zeta \\
&\le w_k(m)\cdot\Big\|\mathbf{f}\Big\|^{m}_{B+k}\cdot\zeta,
\end{aligned}
$$

where $w_k(m)$ is defined in (6.30).

Now let us prove estimate (6.35). For notational simplicity let us set $V_i(x) := B_i^{\Theta}(\mathbf{f})(x)$, $i = 1, 2$, $x \in \mathbb{R}^n$. One easily checks that

$$
\begin{aligned}
E^{\zeta}(x) &:= \Big[V_1^{\zeta}, V_2^{\zeta}\Big](x) - \Big[V_1, V_2\Big]^{\zeta}(x) \\
&= \int_{\mathbb{R}^n}\phi(h)\Big(DV_2(x+\zeta h)\big(V_1^{\zeta}(x) - V_1(x+\zeta h)\big) \\
&\qquad + DV_1(x+\zeta h)\big(V_2^{\zeta}(x) - V_2(x+\zeta h)\big)\Big)\,dh
\end{aligned}
$$

for all $x \in \mathbb{R}^n$. The estimate (6.35) when $k = 0$[21] is promptly obtained, because $|V_i^{\zeta}(x) - V_i(x+\zeta h)| \le 2\|V_i\|_1\cdot\zeta$, $i = 1, 2$, so, by Lemma 6.4,

$$
\|E^{\zeta}(x)\| \le 4\|V_1\|_1\|V_2\|_1\zeta \le 4p_1(m_1)p_1(m_2)\|\mathbf{f}\|^m_B\cdot\zeta.
$$

Now let $k \ge 1$. If $V_i(x) = \displaystyle\sum_{j=1}^{n} V_i^j(x)\frac{\partial}{\partial x^j}$, $i = 1, 2$, by Leibniz's product differentiation rule and the fact that convolution commutes with differentiation, we obtain, for every multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ such that $|\alpha| = k$,

[21] See also [8] for the case when $m = 2$ and $k = 0$.

$$
\begin{aligned}
&\partial^\alpha E^\zeta(x)\\
&\quad = \int_{\mathbb{R}^n} \phi(h) \sum_{\beta\le\alpha}\sum_{j=1}^n \binom{\alpha}{\beta}\Bigg(\partial^{\alpha-\beta}\frac{\partial V_2}{\partial x^j}(x+\zeta h)\cdot\Big((\partial^\beta V_1^j)^\zeta(x)-\partial^\beta V_1^j(x+\zeta h)\Big)\\
&\qquad -\partial^{\alpha-\beta}\frac{\partial V_1}{\partial x^j}(x+\zeta h)\cdot\Big((\partial_\beta V_2^j)^\zeta(x)-\partial^\beta V_2^j(x+\zeta h)\Big)\Bigg)\,dh.
\end{aligned}
$$

Then, by Lemma 6.4,

$$
\Big|(\partial^\beta V_i^j)^\zeta(x)-\partial^\beta V_i^j(x+\zeta h)\Big| \le 2\,\|V_i\|_{k,1}\cdot\zeta \le 2p_{k+1}(m_i)\|(\mathbf{f})\|^{m_i}_{B_i+k+1}
$$

and

$$
\left\|\partial^{\alpha-\beta}\frac{\partial V_i}{\partial x^j}\right\| \le p_{k+1}(m_i)\|(\mathbf{f})\|^{m_i}_{B_i+k+1}
$$

for $i=1,2$, $j\in\{1,\dots,n\}$, $|\alpha|\le k$, $\beta\le\alpha$, from which we obtain

$$
\begin{aligned}
(\mathbf{II}) &= \Big\|D^\ell[(B_1^{\Theta_1}(\mathbf{f}))^\zeta,(B_2^{\Theta_2}(\mathbf{f}))^\zeta]-D^\ell[B_1^{\Theta_1}(\mathbf{f}),B_2^{\Theta_2}(\mathbf{f})]^\zeta\Big\|\\
&\le \sum_{|\alpha|=k}\sum_{\beta\le\alpha}\sum_{j=1}^n\binom{\alpha}{\beta}\Bigg(2p_{k+1}(m_2)\|(\mathbf{f})\|^{m_2}_{B_2+k+1}p_{k+1}(m_1)\|(\mathbf{f})\|^{m_1}_{B_1+k+1}\\
&\qquad +2p_{k+1}(m_1)\|(\mathbf{f})\|^{m_1}_{B_2+k+1}p_{k+1}(m_2)\|(\mathbf{f})\|^{m_1}_{B_2+k+1}\|\Bigg)\cdot\zeta\\
&\le n^2\cdot 2^k\cdot\binom{k+n-1}{k}\cdot\Big(4p_{k+1}(m_1)\cdot p_{k+1}(m_2)\Big)\cdot\|(\mathbf{f})\|^m_{B+k}\cdot\zeta\\
&= r_k(m)\cdot\|(\mathbf{f})\|^m_{B+k}\cdot\zeta, \qquad (6.36)
\end{aligned}
$$

for some $r_k(m)\ge 0$ depending only on k,m,n, so the estimate (6.35) for $(\mathbf{II})$ is proved as well. Putting (6.33), (6.34), and (6.35) together, one gets the thesis. □

6.4. Approximation of the set-valued bracket by shifted brackets of regularized fields

Proposition 6.8. *Let B be an iterated bracket of degree $m\ge 1$, and let $\mathbf{f}=(f_1,\dots,f_m)$ be of class $C^{B-1,1}$ with compact support. Let γ be the function defined in (4.6) and consider a point $x_*\in\mathbb{R}^n$. Then, for all shift twistings Θ and all $x\in\mathbb{R}^n$, and $\zeta>0$, one has*

$$
\mathrm{dist}\,\big(B^\Theta(\mathbf{f}^\zeta)(x),\,B(\mathbf{f})(x_*)\big)\le\gamma\big(\mathbf{d}_B(\Theta,\iota)+|x-x_*|+\zeta\big)+C\zeta \qquad (6.37)
$$

where $C:=s_0(B)(\|\mathbf{f}\|_B)^m$

Proof. By Lemma 6.7

$$\left|B^{\Theta}(\mathbf{f}^{\zeta})(x) - \left(B^{\Theta}(\mathbf{f})\right)^{\zeta}(x)\right| \leq s_0(B)\|\mathbf{f}\|_B^m \cdot \zeta \tag{6.38}$$

for all shift -twistings Θ. On the other hand $\left(B^{\Theta}(\mathbf{f})\right)^{\zeta}(x)$ is an average of vectors of the form $B^{\Theta}(\mathbf{f})(z)$ for points $z = x + \zeta h$, $|h| \leq 1$. Therefore, by $|z - x_*| \leq |x - x_*| + \zeta$ and by the definition of γ it follows that

$$\operatorname{dist}\left(\left(B^{\Theta}(\mathbf{f})\right)^{\zeta}(x),\, B(\mathbf{f})(x_*)\right) \leq \gamma(\mathbf{d}_B(\Theta, \iota) + |x - x_*| + \zeta). \tag{6.39}$$

By (6.38) and (6.39) we finally obtain the thesis. □

References

[1] Aubin, J.-P., Cellina, A.: Differential inclusions: set-valued maps and viability theory, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 264. Springer, Berlin (1984)

[2] Calcaterra, C.: Basic properties of nonsmooth Hörmander's vector fields and Poincaré's inequality. Commun. Math. Anal. **11**(4), 1–40 (2011)

[3] Constantine, G., Savits, T.: A multivariate Faa di Bruno formula with applications. Trans. Am. Math. Soc. **348**(2), 503–520 (1996)

[4] Feleqi, E., Rampazzo, F.: Integral representations for bracket-generating multiflows. Discrete Contin. Dyn. Syst. **35**(9), 4345–4366 (2015)

[5] Encinas, L.H., Masqué, J.M.: A short proof of the generalized Faà di Bruno's formula. Appl. Math. Lett. **16**(6), 975–979 (2003)

[6] Mariconda, C., Tonolo, A.: Discrete Calculus, La Matematica per il 3+2. Springer, Berlin (2016)

[7] Rampazzo, F.: Frobenius-type theorems for Lipschitz distributions. J. Differ. Equ. **243**(2), 270–300 (2007)

[8] Rampazzo, F., Sussmann, H.: Set-valued differentials and a nonsmooth version of Chow–Rashevski's theorem. In: Proceedings of the 40th IEEE Conference on Decision and Control, Orlando, FL, Dec 2001, pp. 2613–2618 (2001)

[9] Rampazzo, F., Sussmann, H.: Commutators of flow maps of nonsmooth vector fields. J. Differ. Equ. **232**(1), 134–175 (2007)

Ermal Feleqi
School of Mathematics
Cardiff University
Senghennydd Road
Cardiff CF24 4AG
UK
e-mail: FeleqiE@cardiff.ac.uk

Franco Rampazzo
Dipartimento di Matematica
Universitá degli Studi di Padova
Via Trieste 63
35121 Padua PD
Italy
e-mail: rampazzo@math.unipd.it

Received: 20 February 2017.
Accepted: 10 October 2017.