h-PRINCIPLE FOR REGULAR CURVES IN BRACKET-GENERATING DISTRIBUTIONS

JAVIER MARTÍNEZ-AGUINAGA AND ÁLVARO DEL PINO

Abstract. We prove that the inclusion of the space of non-singular tangent knots into the space of formal tangent knots is a weak homotopy equivalence for all bracket-generating distributions, as long as the ambient manifold has dimension 4.

The main tool in our arguments is a method, inspired by Control Theory and convex integration, to locally deform a given formal knot to make it tangent. This boils down to introducing "loops" in the projection of the loop onto the distribution.

Contents

1. Statement of the results

Given a smooth manifold M, we say that a smooth section $\mathcal{D}$ of the Grassmann bundle of k-planes is a *k-distribution*. There is a mechanical motivation for considering such an object: we may think of M as the possible states for a particle and of $\mathcal{D}$ as the *admissible directions of motion*. A natural question then is whether any two points in M can actually be connected by an *horizontal path*, i.e. a path whose tangent vectors are contained in $\mathcal{D}$. A sufficient (but not necessary) condition is given by a classic theorem of Chow [?]: any two points in M can be connected if $\mathcal{D}$ is *bracket-generating*. This condition means that any vector in TM can be written as the linear combination (at a point) of vector

2010 *Mathematics Subject Classification.* Primary: 53D10. Secondary: 53D15, 57R17.

fields obtained from $\mathcal{D}$ by finitely many Lie brackets. As such, Chow's theorem is an *infinitesimal* to *global* statement.

Chow's result implies that every homotopy class of loops on M can be represented by an horizontal loop. Let us denote the space of all C^∞-loops by $\Omega(M)$ and let $\Omega(M,\mathcal{D}) \subset \Omega(M)$ be the subspace of $\mathcal{D}$-horizontal ones. We endow both spaces with the C^∞-topology. Then the inclusion

$$\Omega(M,\mathcal{D}) \to \Omega(M)$$

is a π_0-surjection. More recently, it was shown by Z. Ge [12] that this inclusion is a weak homotopy equivalence.

In this paper we will consider a small variation on this theme: Let us write $\mathfrak{Emb}(M,\mathcal{D}) \subset \mathfrak{Imm}(M,\mathcal{D}) \subset \Omega(M,\mathcal{D})$ for the subspaces of horizontal *embedded* and *immersed* loops, respectively. It is natural to ponder whether we can recover properties of $\mathcal{D}$ from the properties of these spaces. A motivating result in this direction was shown by Bennequin [**?**]: the homotopy type of the space of horizontal embedded loops can distinguish distinct 3–dimensional contact structures that are homotopic as plane fields.

We setup the problem as follows: there exists a forgetful map from $\mathfrak{Imm}(M,\mathcal{D})$ to the space of *formal horizontal immersions*:

$$\mathfrak{Imm}^f(M,\mathcal{D}) = \{(\gamma, F) \mid \gamma \in \Omega(M), F \in \mathrm{Mon}(T\mathbb{S}^1, \gamma^*\mathcal{D}).\}$$

An element in $\mathfrak{Imm}^f(M,\mathcal{D})$ is simply a loop γ with a choice of non-vanishing vector field in $\gamma^*\mathcal{D}$ playing the role of the *formal tangent vector field*. If $\mathcal{D}_0$ and $\mathcal{D}_1$ are isomorphic as bundles, the spaces $\mathfrak{Imm}^f(M,\mathcal{D}_0)$ and $\mathfrak{Imm}^f(M,\mathcal{D}_1)$ are homotopy equivalent; it follows that $\mathfrak{Imm}^f(M,\mathcal{D})$ does not encode the *geometric* nature of $\mathcal{D}$. One might wonder whether this is also the case for $\mathfrak{Imm}(M,\mathcal{D})$. If $\mathcal{D}$ is a contact structure the answer is affirmative, since standard h–principle techniques [8] show that $\mathfrak{Imm}(M,\mathcal{D})$ is weakly homotopy equivalent to $\mathfrak{Imm}^f(M,\mathcal{D})$. However, the result fails to hold in general due to the presence of *rigid curves* [**?**, 3].

A rigid horizontal curve is a curve that, up to reparametrisation, possesses no C^∞-deformations relative to its endpoints. As such, these curves are isolated and conform exceptional components within the space of all horizontal immersions with given boundary conditions. *Rigid loops* (i.e. those whose deformations are just reparametrisations) also exist and, because of this, the inclusion $\mathfrak{Imm}(M,\mathcal{D}) \to \mathfrak{Imm}^f(M,\mathcal{D})$ is in general not bijective at the level of connected components; see [20, Remark 23].

Being rigid is the most extreme case of being *singular*: this means that the *endpoint map* of the curve is not submersive, so the curve has fewer deformations than expected; see Subsection **??**. The spaces of singular and rigid curves have a *geometric* and not a *topological* nature, since small perturbations of $\mathcal{D}$ (even within a given class of distributions) can radically change their homotopy type [20, Theorem 27]. Additionally, it was shown in [**?**, Corollary 7] that *regular* curves (i.e. non-singular) are C^∞–generic. Our Theorem **??** discusses this genericity in more quantitative terms.

These considerations motivate us to discard the singular curves when we study the space $\mathfrak{Imm}(M,\mathcal{D})$. We will henceforth write $\mathfrak{Imm}^{\mathrm{r}}(M,\mathcal{D})$ for the space of regular horizontal loops. Our first result reads:

Theorem 1. *Let $(M,\mathcal{D})$ be a manifold endowed with a bracket-generating distribution. Then the inclusion*

$$\mathfrak{Imm}^{\mathrm{r}}(M,\mathcal{D}) \to \mathfrak{Imm}^f(M,\mathcal{D})$$

is a weak homotopy equivalence.

M. Gromov posed this as an exercise in [**?**, p. 84] without mentioning the regularity assumption. It was only later with the work [3] that the need for regularity became apparent. This result is already known in the contact case [8] and in the Engel case [20].

One can proceed in a similar manner when studying embeddings. We write $\mathfrak{Emb}^f(M,\mathcal{D})$ for the space of *formal* horizontal embedded loops (see Subsection **??**) and $\mathfrak{Emb}^{\mathrm{r}}(M,\mathcal{D})$ for the space of regular horizontal embeddings. Then our second (and main) result reads:

Theorem 2. *Let* $(M, \mathcal{D})$ *be a bracket-generating distribution with* $\dim(M) \geq 4$*. Then the inclusion*

$$\mathfrak{Emb}^{\mathrm{r}}(M, \mathcal{D}) \to \mathfrak{Emb}^{f}(M, \mathcal{D})$$

is a weak homotopy equivalence.

Note that the dimensional assumption is sharp, since the result is known to be false in 3-dimensional contact topology [**?**].

Theorem 2 was already known in certain instances: in the Engel case [**?**] and in the higher-dimensional contact setting [8, p. 128] It is worth pointing out that both the present paper and [**?**] follow the general structure of an h-principle, but the key geometrical construction is different. Our main ingredient is the definition of a (locally defined) *generalised Lagrangian projection* in which the distribution can be understood as a connection; see Subsection **??**. Reading [12] brought to our attention that this generalised Lagrangian projection is commonly used (not under this name) in the control theory literature.

1.1. Structure of the paper.

Section 2 provides an overview of distributions and the bracket-generating condition.

Section 3 explains the main ingredients of the proof. The key new idea in this paper is our definition of a tangle (Subsection 3.4): this is a local model for tangent curves that is homotopically trivial and performs a motion in the direction of a given iterated Lie bracket. This can be regarded as a generalisation of 1-dimensional convex integration in which the flower (which integrates to the tangle) produces the desired motion of higher order. In the Appendix, some auxiliary results on iterated Lie brackets are stated and proven.

In Section 4 we put together the previous ingredients in order to conclude the proof.

Acknowledgments:

2. The language of distributions

In this section we introduce the definitions of interest in the paper, as well as the background material we need. A standard reference regarding the theory of distributions is [**?**]. All the h-principle language and results we use can be found in [**?**, 8]. At the end of the section we will look at singular horizontal curves, reviewing some of the contents of [**?**].

2.1. Differential systems.

The following definition generalises the notion of distribution:

Definition 3. *Let* M *be a smooth manifold. A* ***differential system*** $\mathcal{D}$ *is a* C^∞*-submodule of the space of smooth vector fields.*

Given any smooth distribution on M we may construct a differential system by taking its smooth sections. Conversely, a differential system $\mathcal{D}$ arises from a distribution if the dimension of its span $\mathcal{D}(p) \subset T_pM$ remains constant as $p \in M$ varies. In this manner, we think of differential systems as singular distributions; we will often abuse notation and use $\mathcal{D}$ to denote both.

Remark 4. *A subtle point here is that, when* M *is not compact, other sensible choices do exist. For instance, one might consider the compactly supported sections of the distribution instead* []*. Our main focus is on closed manifolds, so this will not be relevant for us.*

In the study of differential systems (and distributions) a classical question is whether a given germ of differential system can be put in a standard form. Equivalently, we wonder whether two given germs can be mapped to one another by a diffeomorphism germ. One can often find obstructions to the existence of such a diffeomorphism by showing that there is some invariant distinguishing the two differential systems. We will now show that the behaviour of the differential system with respect to the Lie bracket provides many such invariants. In some sense, all these invariants measure the failure of the differential system to be a Lie subalgebra of the space of all smooth vector fields.

2.2. **Growth vector and associated flag.** Let us introduce some terminology. We say that the string a, depending on the variable a, is a **formal bracket expression** of length 0. Similarly, we say that the string $[a_0, a_1]$, depending on the variables a_0 and a_1, is a formal bracket expression of length 1. Inductively, we define a formal bracket expression of length n to be a string of the form $[A(a_0, \cdots, a_j), B(a_{j+1}, a_n)]$ with $0 < j < n$ and A and B formal bracket expressions of lengths j and $n - j - 1$, respectively.

Given a differential system $\mathcal{D}$ we define its **(small) associated flag** as the sequence of differential systems

$$\mathcal{D}_0 \subset \mathcal{D}_1 \subset \mathcal{D}_2 \subset \cdots$$

in which $\mathcal{D}_i$ is the C^∞-span of the vector fields of the form $A(v_1, \cdots, v_j)$, $j \leq i$, where the v_k are vector fields in $\mathcal{D}$ and A is a formal bracket expression of length j. As such, $\mathcal{D}_0 = \mathcal{D}$. The reader should note that sometimes in the literature the associated flag is defined in a slightly different way []; see Example 6 below. The definition we give is well-suited to our purposes and, arguably, it is the most natural one from the perspective of Lie algebras.

If M is compact and the differential system is analytic, one can show that the associated flag eventually stabilises (see [2], Proposition pp. 13–14); i.e. $\mathcal{D}_{i+1} = \mathcal{D}_i$ for all $i \geq m$, for some integer m. However, this does not need to be true anymore if either of the assumptions is dropped; see Example 6.

Given a point $p \in M$ one can produce a flag of vector spaces using the associated flag:

$$\mathcal{D}_0(p) \subset \mathcal{D}_1(p) \subset \mathcal{D}_2(p) \subset \cdots$$

Where $\mathcal{D}_i(p)$ denotes the span of $\mathcal{D}_i$ at the point p. This flag produces a non-decreasing sequence of integers

$$(\dim(\mathcal{D}_0(p)), \dim(\mathcal{D}_1(p)), \dim(\mathcal{D}_2(p)), \cdots)$$

which in general depends on p. This sequence is called the **growth vector** of $\mathcal{D}$ at p. If the growth vector does not depend on the point, we will say that the differential system $\mathcal{D}$ is of *constant growth*. If this is the case, the differential systems in the associated flag all arise from distributions. Some examples of families of constant growth are (regular) foliations, contact structures, Engel structures, and Goursat structures.

The following notion is central to us:

Definition 5. *A differential system $(M, \mathcal{D})$ is **bracket-generating** if, for every p and every $v \in T_pM$, there is an integer j such that $v \in \mathcal{D}_j(p)$.*

That is, every direction in M can be generated from the admissible directions of motion by iterated Lie brackets. A key property is that the integer j is a lower-semicontinuous function of p. It follows that if M is compact the integer j can be taken to be independent of p and v.

In the following example, in which we look at the *Martinet distribution*, we discuss the different definitions of growth vector in the literature:

Example 6. *Fix a positive integer k and let us consider the differential system*

$$(\mathbb{R}^3, \mathcal{D} = \ker(dy - z^k dx) = \langle \partial_z, \partial_x + z^k \partial_y \rangle).$$

The differential systems in the associated flag $\{\mathcal{D}_i\}_{0 \leq i \leq k}$, are spanned by ∂_z, ∂_x, and $z^{k-i}\partial_y$. In particular, $\mathcal{D}_i(p) = \mathcal{D}_{i'}(p)$ for all $0 < i, i' < k$ and all $p \in \mathbb{R}^3$, despite the fact that the underlying set of generators of the C^∞-module is different. Additionally, $\mathcal{D}_k = TM$. In Figure 1 we show the case $k = 2$.

One can modify slightly this example and consider instead the differential system $(\mathbb{R}^3, \mathcal{D} = \ker(dy - f(z)dx))$ with f a function vanishing to all orders along $z = 0$ and otherwise satisfying $f'(z) \neq 0$. Then, the associated flag of C^∞-modules never stabilises, but the sequence of vector spaces $\{\mathcal{D}_i(p)\}_{i=1,\cdots,\infty}$ is constant for each p.

In the literature the (small) associated flag is sometimes defined as follows:

$$\tilde{\mathcal{D}}_0 = \mathcal{D} \subset \tilde{\mathcal{D}}_1 = [\mathcal{D}_0, \mathcal{D}_0] \subset \cdots \subset \tilde{\mathcal{D}}_{i+1} = [\tilde{\mathcal{D}}_i, \mathcal{D}_0] \subset \cdots$$

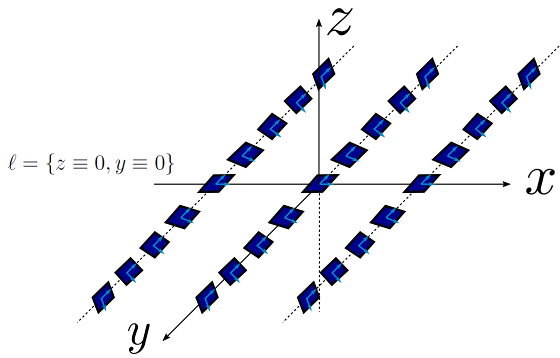


FIGURE 1. The structure $\ker(dy - z^2 dx)$ in $\mathbb{R}^3$.

Where $[\tilde{\mathcal{D}}_i, \mathcal{D}_0]$ denotes the C^∞-span of all Lie brackets of vector fields tangent to $\tilde{\mathcal{D}}_i$ and $\mathcal{D}_0$ (as subspaces of the tangent space at each point). In all the previous examples, one can readily check that $\tilde{\mathcal{D}}_1 = \mathcal{D}_1$ and that $\tilde{\mathcal{D}}_2$ agrees with TM. In particular, a distribution can be bracket-generating in the second sense but not in the first. □

Assumption 7. *We will henceforth assume that the differential system $\mathcal{D}$ we start with is a distribution. We make no assumptions regarding the other elements of the associated flag.*

2.3. Curvature.

Distributions with the same growth vector may still behave very differently. We will now define another pointwise invariant called the *curvature*.

Let $(M, \mathcal{D})$ be a manifold equipped with a distribution. Let $p \in M$, and $u \in \mathcal{D}(p)$. A local vector field $\tilde{u} \in \mathfrak{X}(\mathcal{O}p(p))$ is a **local extension** of u (with respect to $\mathcal{D}$) if $\tilde{u}(p) = u$ and $\tilde{u}(q) \in \mathcal{D}(q)$ for every $q \in \mathcal{O}p(p)$.

Definition 8. *The **curvature** of the distribution $\mathcal{D}$ is the bundle morphism:*

$$\Gamma(\mathcal{D}) : \wedge^2 \mathcal{D}(p) \to T_pM/\mathcal{D}(p)$$
$$(u, v) \mapsto [\tilde{u}, \tilde{v}](p) + \mathcal{D}(p)$$

where $\tilde{u}$ and $\tilde{v}$ are local extensions of u and v, respectively.

Lemma 9. *The curvature is well-defined.*

Proof. We have to check that the definition does not depend on the chosen extension. Take a local basis $\{X_1, \cdots, X_k\}$ of $\mathcal{D}$, so we may write $\tilde{u} = \sum_i f_i X_i$ and $\tilde{v} = \sum_j g_j X_j$. Then we compute:

$$[\tilde{u}, \tilde{v}] = \left[\sum_i f_i X_i, \sum_j g_j X_j \right] = \sum_{i,j} [f_i X_i, g_j X_j] = \sum_{i,j} f_i g_j [X_i, X_j] - dg_j (f_i X_i)(X_j) - df_i (g_j X_j)(X_i).$$

Now we observe that $dg_j (f_i X_i)(X_j) + df_i (g_j X_j)(X_i) \in \mathcal{D}$ and that $\sum_{i,j} f_i g_j [X_i, X_j](p)$ does not depend on the extension. The argument concludes by noting that the map is bilinear and antisymmetric, due to the properties of the Lie bracket, so indeed it descends to $\wedge^2 \mathcal{D}(p)$. □

By Frobenius' theorem, $\mathcal{D}$ is integrable if and only if the rank of its curvature is zero. Indeed, this rank measures how far the distribution is from being integrable.

Remark 10. *Consider a manifold M equipped with a bracket–generating distribution $\mathcal{D}_0$ with associated flag*

$$\mathcal{D}_0 \subset \mathcal{D}_1 \subset \cdots \subset \mathcal{D}_i \subset \cdots \subset \mathcal{D}_m \subset TM$$

in which $\mathcal{D}_i$ is a distribution of dimension k_i. Then we can define, as in Definition 8, the following morphisms

$$\Gamma_i : \mathcal{D}_0 \otimes \mathcal{D}_i \to \mathcal{D}_{i+1}/\mathcal{D}_i$$
$$(u, v) \mapsto [\tilde{u}, \tilde{v}] + \mathcal{D}_i.$$

From now on we will abuse notation and write $[u, v](p)$ for $[\tilde{u}, \tilde{v}](p) + \mathcal{D}(p)$ in $T_pM/\mathcal{D}(p)$.

3. Ingredients of the proof

3.1. **Horizontal immersions and embeddings.** As we explained in the introduction, the dynamical origin of distributions motivates us to look into those curves that are tangent to them:

Definition 11. *Let $(M,\mathcal{D})$ be a manifold endowed with a distribution. A map/immersion/embedding $\gamma:\mathbb{S}^1\to M$ is said to be **horizontal** if $\gamma'(t)\in\mathcal{D}_{\gamma(t)}$ for all $t\in\mathbb{S}^1$.*

Let $\mathfrak{Imm}(M)$ be the space of smooth immersions of $\mathbb{S}^1$ into M equipped with the C^∞-topology. We write $\mathfrak{Emb}(M)$ for the subspace of embeddings. We are interested in those immersions/embeddings that are horizontal, so we define the following subspaces:

$$\mathfrak{Imm}(M,\mathcal{D})=\left\{\gamma\in\mathfrak{Imm}(M):\gamma'(t)\in\mathcal{D}_{\gamma(t)}\quad\forall t\in\mathbb{S}^1\right\}$$
$$\mathfrak{Emb}(M,\mathcal{D})=\left\{\gamma\in\mathfrak{Emb}(M):\gamma'(t)\in\mathcal{D}_{\gamma(t)}\quad\forall t\in\mathbb{S}^1\right\}$$

An immersion is a map satisfying a certain differential relation, namely, having a differential of maximal rank. By decoupling this differential relation we can define *formal* analogues of these spaces.

The space of **formal horizontal immersions**:

$$\mathfrak{Imm}^f(M,\mathcal{D})=\left\{(\gamma,F):\gamma\in\mathfrak{Maps}(\mathbb{S}^1,M),\quad F\in\mathrm{Mon}_{\mathbb{S}^1}(T\mathbb{S}^1,\gamma^*\mathcal{D})\right\}.$$

The space of **formal horizontal embeddings**:

$$\mathfrak{Emb}^f(M,\mathcal{D})=\Big\{\big(\gamma,(F_s)_{s\in[0,1]}\big):\quad\gamma\in\mathfrak{Emb}(M),\quad F_s\in\mathrm{Mon}_{\mathbb{S}^1}(T\mathbb{S}^1,\gamma^*TM),$$
$$F_0=\gamma',\quad F_1\in\gamma^*\mathcal{D}\Big\}.$$

There are natural inclusions:

$$\mathfrak{Imm}(M,\mathcal{D})\to\mathfrak{Imm}^f(M,\mathcal{D}),\qquad\mathfrak{Emb}(M,\mathcal{D})\to\mathfrak{Emb}^f(M,\mathcal{D}).$$

by setting $F=\gamma'$ in the first case and $F_s=\gamma'$ in the second. These inclusions are continuous maps when all these spaces are endowed with the C^∞-topology. The central question in the *h-principle* philosophy is whether these inclusions are weak homotopy equivalences. Note that the formal spaces have an *algebro-topological* nature, whereas the original spaces are defined in *geometric* terms.

h-principle proofs are local in nature. As such, we will need to consider horizontal curves parametrised by intervals. If I is a 1-manifold, we write

$$\mathfrak{Imm}(I,M,\mathcal{D})\to\mathfrak{Imm}^f(I,M,\mathcal{D}),\qquad\mathfrak{Emb}(I,M,\mathcal{D})\to\mathfrak{Emb}^f(I,M,\mathcal{D})$$

for the spaces of horizontal immersions, formal horizontal immersions, embeddings, and formal horizontal embeddings of I into $(M,\mathcal{D})$.

3.2. ε–**horizontal immersions and embeddings.** It is usually harder to work with closed differential relations (such as being horizontal) than with open ones. As such, one might want to weaken the notion of horizontality to that of ε-horizontality, obtaining an open relation. ε-horizontal curves, just like actual horizontal curves, can be manipulated in terms of their projections to the distribution; this will be explained in Subsection **??**.

Fix a riemannian metric g in M. We write $\angle$ for the positive angle, in terms of the metric g, between any two given linear subspaces at each T_pM.

Given a constant $\pi/2>\varepsilon>0$ we define the space of **ε-horizontal embeddings** as follows:

$$\mathfrak{Emb}^\varepsilon(M,\mathcal{D})=\{\gamma\in\mathfrak{Emb}(M):\angle(\gamma',\mathcal{D})<\varepsilon\}.$$

We can also consider its formal analogue, the space of **formal ε-horizontal embeddings**:

$$\mathfrak{Emb}^f(M,\mathcal{D})=\Big\{\big(\gamma,(F_s)_{s\in[0,1]}\big):\quad\gamma\in\mathfrak{Emb}(M),\quad F_s\in\mathrm{Mon}_{\mathbb{S}^1}(T\mathbb{S}^1,\gamma^*TM),$$
$$F_0=\gamma',\quad\angle(F_1,\gamma^*\mathcal{D})<\varepsilon\Big\}.$$

There is a natural inclusion $\mathfrak{Emb}^\varepsilon(M,\mathcal{D})\to\mathfrak{Emb}^{f,\varepsilon}(M,\mathcal{D})$. It is a classic result due to M. Gromov that the h-principle holds in this setting:

Proposition 12. *Let $\mathcal{D}$ be a distribution of rank greater or equal to 2. The inclusion $\mathfrak{Emb}^\varepsilon(M,\mathcal{D})\to\mathfrak{Emb}^{f,\varepsilon}(M,\mathcal{D})$ is a weak homotopy equivalence.*

Proof. The proof follows from the convex integration theorem for open and ample relations (Theorem [8, 18.4.1 p.171]). The relation is clearly open. The ampleness follows from the fact that the only coordinate principal subspace (see [8, page 167]) is the fiber itself, which intersected with the fiberwise defining relation gives conical sets of the form

$$\{v \in \mathbb{R}^n : \angle(v, \mathcal{D}) < \varepsilon\}.$$

They are connected (because $\mathcal{D}$ has at least rank 2) and their convex hull is the whole $\mathbb{R}^n$; they are therefore ample. ☐

Additionally the formal spaces $\mathfrak{Emb}^{f,\varepsilon}(M, \mathcal{D})$ and $\mathfrak{Emb}^f(M, \mathcal{D})$ are easily shown to be homotopy equivalent:

Lemma 13. *The inclusion $\mathfrak{Emb}^f(M, \mathcal{D}) \to \mathfrak{Emb}^{f,\varepsilon}(M, \mathcal{D})$ is a homotopy equivalence.*

Proof. Just note that the fiberwise orthogonal riemannian projection of F_1 onto $\mathcal{D}$ provides a homotopy inverse. ☐

Corollary 14. *The spaces $\mathfrak{Emb}^\varepsilon(M, \mathcal{D})$ and $\mathfrak{Emb}^f(M, \mathcal{D})$ are weakly homotopy equivalent.*

We can summarise the previous results in the following commutative diagram:

$$\begin{array}{ccc} \mathfrak{Emb}(M, \mathcal{D}) & \longrightarrow & \mathfrak{Emb}^f(M, \mathcal{D}) \\ \downarrow & & \downarrow \cong \\ \mathfrak{Emb}^\varepsilon(M, \mathcal{D}) & \xrightarrow{\cong} & \mathfrak{Emb}^{f,\varepsilon}(M, \mathcal{D}) \end{array}$$

It follows that if we want to relate the spaces $\mathfrak{Emb}(M, \mathcal{D})$ and $\mathfrak{Emb}^f(M, \mathcal{D})$, it is sufficient to understand the inclusion $\mathfrak{Emb}(M, \mathcal{D}) \hookrightarrow \mathfrak{Emb}^\varepsilon(M, \mathcal{D})$.

These h-principle results are also relative in the parameter, relative in the domain, and satisfy C^0-closeness. More precisely:

Proposition 15. *Let K be a compact manifold. Let $(M, \mathcal{D})$ be a manifold endowed with a distribution of rank greater or equal to 2. Suppose we are given a map $(\gamma, F_s) : K \to \mathfrak{Emb}^f([0,1], M, \mathcal{D})$ satisfying:*

- *$(\gamma, F_s)(k)|_{\mathcal{O}p(\{0,1\})}$ is a ε-horizontal embedding for all $k \in K$,*
- *$(\gamma, F_s)(k) \in \mathfrak{Emb}^\varepsilon([0,1], M, \mathcal{D})$ for $k \in \mathcal{O}p(\partial K)$.*

Then there exists a map $(\tilde{\gamma}, \tilde{F}_s) : K \to \mathfrak{Emb}^\varepsilon([0,1], M, \mathcal{D})$ satisfying:

- *$(\tilde{\gamma}, \tilde{F}_s)$ is homotopic to (γ, F_s) as a map into $\mathfrak{Emb}^f([0,1], M, \mathcal{D})$,*
- *this homotopy is relative in the parameter (i.e. relative to $k \in \mathcal{O}p(\partial K)$) and in the domain (i.e relative to $t \in \mathcal{O}p(\{0,1\})$),*
- *$\tilde{\gamma}$ is C^0-close to γ.*

One can define, analogously, the space of ε-horizontal immersions $\mathfrak{Imm}^\varepsilon(M, \mathcal{D})$:

$$\mathfrak{Imm}^\varepsilon(M, \mathcal{D}) = \{\gamma \in \mathfrak{Imm}(M) : \angle(\gamma', \mathcal{D}) < \varepsilon\}.$$

From the arguments above it follows that:

Lemma 16. *Let $\mathcal{D}$ be a distribution of rank greater or equal to 2. The inclusion $\mathfrak{Imm}^\varepsilon(M, \mathcal{D}) \to \mathfrak{Imm}^f(M, \mathcal{D})$ is a weak homotopy equivalence.*

This h-principle is also relative in the parameter, relative in the domain, and C^0-close.

3.3. **Adapted-coordinate charts.** When studying Legendrian (i.e. horizontal) knots in standard contact $(\mathbb{R}^3, \xi_{\rm std} = \ker(dy - zdx))$, one often uses their so-called Lagrangian projection:

$$\pi_L \colon \mathbb{R}^3 \to \mathbb{R}^2$$
$$(x, y, z) \mapsto (x, z)$$

which projects the distribution ξ at each point $p \in \mathbb{R}^3$ isomorphically onto the tangent space $T_{\pi_L(p)}\mathbb{R}^2$ at the image point. This implies that any Legendrian knot projects to an immersed planar curve. From the projected curve one can recover the y-coordinate by integration:

$$z(t) = z(t_0) + \int_{t_0}^{t} x(s)y'(s)ds.$$

Note that the integral on the right-hand side, when evaluated over the whole curve, computes the area it bounds due to Stokes' theorem. This turns the problem of manipulating Legendrian knots into a problem about planar curves satisfying an area constraint.

In this subsection we introduce the notion of a (local) *generalised Lagrangian projection* for a general bracket–generating distribution:

Definition 17. *Let M be an n-dimensional manifold and fix a point $p \in M$. Let $\mathcal{D} = \mathcal{D}_0 \subset \mathcal{D}_1 \subset \cdots \subset \mathcal{D}_m = TM$ be the flag (of differential systems) associated to the rank-k bracket-generating distribution $\mathcal{D}$. Let k_i be the rank of $\mathcal{D}_i(p)$. Write $\pi_L : \mathbb{R}^n \to \mathbb{R}^k$ for the projection onto the first k coordinates.*

A set of local coordinates $\phi : (\mathbb{D}_\varepsilon(p), p) \to (\mathbb{R}^n, 0)$ is said to be ***adapted*** *to $\mathcal{D}$ if the following properties hold:*

(a) $\phi_*\mathcal{D}_i(p) = \langle \partial_1, \cdots, \partial_{k_i} \rangle$.
(b) $d_q\pi_L : \phi_*\mathcal{D} \to \mathbb{R}^k$ *is an isomorphism at each point $q \in \mathcal{O}p(p)$.*
(c) *For each $j \in \{k_{i-1} + 1, \cdots, k_i\}$ there is a formal bracket expression of length i satisfying:*

$$\partial_j(p) = A_j(X_{l_1^j}, \cdots, X_{l_i^j})(p);$$

where X_l is the lift to $\mathcal{D}$ of the vector field ∂_l in $\mathbb{R}^k$.

The projection $\pi_L \circ \phi$ is said to be a local ***Lagrangian projection****. The framing $\{X_1, \cdots, X_n\}$ consisting on the liftings of the vector fields ∂_i in the base is called the* ***adapted framing****.*

A few remarks are in order. Property (b) states that $\phi_*\mathcal{D}$ is an Ehresmann connection over $\mathbb{R}^k$; from this, it follows that any ∂_l with $l \in \{1, \cdots, k\}$ can be uniquely lifted to $\phi_*\mathcal{D}$. The importance of this is apparent: any curve γ in $\mathbb{R}^k$ can be uniquely lifted to a horizontal curve of $\phi_*\mathcal{D}$ once one of its points has been lifted (due to the fundamental theorem of ODEs, and as long as the curve does not escape the adapted neighbourhood ϕ). Property (c) explains quantitatively how to move vertically in the fibration by describing a path in the base. These properties justify calling $\pi_L \circ \phi$ a Lagrangian projection. As remarked in the introduction, this notion is ubiquitous in the control theory literature []. In this context, the space $T\mathbb{R}^k$ corresponds to the space of controls.

The following Lemma proves the existence of Lagrangian projections:

Lemma 18. *Let $(M, \mathcal{D})$ be a manifold endowed with a bracket-generating distribution. Fix a point $p \in M$. Then, there exists an adapted set of coordinates in a neighbourhood of p.*

Proof. Fix a basis $\{Y_1, \cdots, Y_n\}$ of T_pM such that $\{Y_1, \cdots, Y_k\}$ is a basis of $\mathcal{D}(p)$. We construct a set of coordinates ψ by applying the exponential map to this basis, effectively taking Y_i to ∂_i. These coordinates already satisfy Property (b) if the neighbourhood of p considered is small enough.

Using these coordinates and π_L we define a framing $\{X_1, \cdots, X_k\}$ of $\mathcal{D}$ by lifting the coordinate vector fields ∂_i of $\mathbb{R}^k$. By definition $\mathcal{D}_i(p)$ is spanned by a collection of bracket expressions of length at most i whose entries are vector fields tangent to $\mathcal{D}$. We claim that $\mathcal{D}_i(p)$ is in fact spanned by bracket expressions whose entries are elements of the framing $\{X_1, \cdots, X_k\}$. Indeed, if we have any bracket expression, we can write each of its entries as a C^∞-combination of the framing. Expanding such an expression yields the claim. This argument yields a framing $\{X_1, \cdots, X_n\}$ of TM in which $\{X_1(p), \cdots, X_{k_i}(p)\}$ spans $\mathcal{D}_i(p)$.

Now we find a linear transformation Ψ of $\mathbb{R}^n$ mapping each $X_l(p)$ to ∂_l. By construction we have $\pi_L \circ \Psi = \pi_L$, so the vector fields $\{X_1, \cdots, X_k\}$, which are the lifts to $\mathcal{D}$ of the coordinate vectors, are mapped under Ψ to the lifts to $\Psi_*\mathcal{D}$ of the coordinate vectors. The chart $\Psi \circ \psi$ is therefore adapted. □

The following proposition reflects the fact that, when working in an adapted coordinate chart, the increase in certain fiber-coordinate produced by taking a bracket of two vector fields can be reinterpreted as an area condition. This is the content of the following proposition. Let $(M, \mathcal{D})$ be a manifold endowed with a bracket-generating distribution and consider an adapted set of coordinates in a neighbourhood of $p \in M$, $\phi : (\mathbb{D}_r(p), p) \to (\mathbb{R}^n, 0)$. Consider a framing $X_1, \cdots, X_n$ associated to the adapted chart as the one described in Lemma 18.

Proposition 19. *Assume $[X_i, X_j = B(X_1, \cdots, X_{\mathrm{J}})] = X_z$. Then, any curve $\gamma : \mathbb{S}^1 \to \phi(\mathcal{D}_r(p))$ enclosing area A in the plane $\mathbb{R}^2 = \langle \partial_i, \partial_j \rangle$ lifts to the distribution as a curve $\tilde{\gamma} : [0,1] \to M$ where the difference in the values of the ∂_z-coordinate between the points $\gamma(1)$ and $\gamma(0)$ is $A\,(1 + O(r))$. More precisely, if we denote by α_z be the dual covector associated to X_z,*

$$\int_\gamma \alpha_z = A\,(1 + O(r))\,.$$

Proof. Denote by $\Gamma_{(\tilde{\gamma}(1), \tilde{\gamma}(0))}$ the segment connecting the points $\tilde{\gamma}(1)$ and $\tilde{\gamma}(0)$. Denote by $\beta := \tilde{\gamma} \# \Gamma_{(\tilde{\gamma}(1), \tilde{\gamma}(0))}$ the concatenation of the curves $\tilde{\gamma}$ and $\Gamma_{(\tilde{\gamma}(1), \tilde{\gamma}(0))}$. Note that

$$\int_\beta \alpha_z = \int_{\Gamma_{(\tilde{\gamma}(1), \tilde{\gamma}(0))}} \alpha_z$$

and, thus, this integral measures the difference of the ∂_z-coordinate values of the points $\tilde{\gamma}(1)$ and $\tilde{\gamma}(0)$. Consider a topological disk $\mathcal{D}_{\tilde{\gamma}}$ bounded by β and whose boundary gets projected to γ in the projection onto the plane $\langle \partial_i, \partial_j \rangle$. By Stokes Theorem,

$$\int_\beta \alpha_z = \int_{\mathcal{D}_{\tilde{\gamma}}} d\alpha_z.$$

By Cartan's formula we have that

$$d\alpha_z(X_i, X_j) = \alpha_z\left([X_j, X_i]\right).$$

Thus, if we particularize this equation at the point $p \in M$, we get that

$$d\alpha_z(p)(X_i(p), X_j(p)) = 1,$$

and it vanishes when evaluated at any other combination of two elements of the framing associated to the coordinate chart. Thus, the $2-$form $d\alpha$ coincides with $dx_i \wedge dx_j$ in the origin at the level of $0-$jets. As an application of Taylor's Remainder Theorem we get that

$$\int_{\mathcal{D}_{\tilde{\gamma}}} d\alpha_z = \int_{\mathcal{D}_{\tilde{\gamma}}} dx_i \wedge dx_j + O(r) = A + A \cdot O(r)$$

yielding the claim. □

Remark 20. *Two curves that are $C^\infty-$close in the base $\phi_*\mathcal{D}(p) \subset \phi(\mathbb{D}_r(p))$ lift to horizontal curves that are $C^\infty-$close upstairs.*

Remark 21. *The horizontal lifting of a curve to an Ehresmann connection does not depend on its parametrization.*

3.4. **Tangles of commutators.** Let us introduce some notation first. Given a $1-$parametric family of diffeomorphisms ϕ_t, we denote by $\gamma_{\phi_t}(x) = \phi_t(x)$ its integral curve arising from the point x. Given non-negative real numbers η, τ, we denote by $\gamma^\eta \# \beta^\tau$ the piecewise-smooth ponderated concatenation of two curves γ and β such that $\gamma(\eta) = \beta(0)$ defined as follows:

$$\gamma^\eta \# \beta^\tau = \begin{cases} \gamma(t) & t \in [0, \eta], \\ \beta(t-\eta) & t \in [\eta, \eta+\tau], \end{cases}$$

Following this notation, we define then the ponderated concatenation $\gamma^\eta_{\phi_t} \# \gamma^\tau_{\psi_t}(x)$ of the curves of two $1-$parametric families of diffeomorphisms ϕ_t, ψ_t as follows:

$$\gamma^\eta_{\phi_t} \# \gamma^\tau_{\psi_t}(x) = \begin{cases} \gamma_{\phi_t}(x) & t \in [0, \eta], \\ \gamma_{\psi_{t-\eta}}(\phi_\eta(x)) & t \in [\eta, \eta+\tau], \end{cases}$$

Given k families of diffeomorphisms $\psi_1, \cdots, \psi_k$ and $\alpha_1, \cdots, \alpha_k$ nonnegative numbers, we define the ponderated concatenation of their k curves as follows

$$\gamma^{\alpha_1}_{\psi^1_t} \# \cdots \# \gamma^{\alpha_{k-1}}_{\psi^{k-1}_t} \# \gamma^{\alpha_k}_{\psi^k_t}(x) = \left(\gamma^{\alpha_1}_{\psi^1_t} \# \cdots \# \gamma^{\alpha_{k-1}}_{\psi^{k-1}_t}\right) \# \gamma^{\alpha_k}_{\psi^k_t}(x).$$

3.4.1. *Piecewise smooth tangles.* We define piecewise smooth tangles of commutators by induction on their length.

- **Length 2.** We define first the simplest tangles (non-iterated tangles of length 2) for two $1-$parametric families of diffeomorphisms ϕ_s, ψ_s as follows

$$\mathcal{T}_{[\phi_t,\psi_t]}(x) := \gamma^t_{\phi_s} \# \gamma^t_{\psi_s} \# \gamma^t_{\phi_s^{-1}} \# \gamma^t_{\psi_s^{-1}}$$

  We define $k-$times iterated tangles of length 2 for two $1-$parametric families of diffeomorphisms ϕ_t, ψ_t as follows

$$\mathcal{T}_{[\phi_t,\psi_t]^k}(x) = \mathcal{T}_{\left[\phi_{\frac{t}{\sqrt{k}}},\psi_{\frac{t}{\sqrt{k}}}\right]} \# \overset{k}{\cdots} \# \mathcal{T}_{\left[\phi_{\frac{t}{\sqrt{k}}},\psi_{\frac{t}{\sqrt{k}}}\right]}$$

  We call *step* 1-*iteration number* to the integer k and we refer to t as the *side of the tangle.* We also use the following notation for negative iteration numbers:

$$\mathcal{T}_{[\phi_t,\psi_t]^{-k}}(x) = \mathcal{T}_{[\psi_t,\phi_t]^k}(x)$$

- **Length $n > 2$.** Given a tangle expression $A(\phi^1_t, \cdots, \phi^{n-1}_t)$ of length $n-1$, we define the associated tangle to any non-iterated tangle expression $B(\phi^1_t, \cdots, \phi^{n-1}_t, \psi_t) = [A(\phi^1_t, \cdots, \phi^{n-1}_t), \phi_t]^{\# k}$ of length n as

$$\mathcal{T}_{\left[A(\phi^1_t,\cdots,\phi^{n-1}_t),\psi_t\right]} = \gamma^t_{\psi_t^{-1}} \# \mathcal{T}_{A(\phi^1_t,\cdots,\phi^{n-1}_t)^{\#-1}} \# \gamma^t_{\psi_t} \# \mathcal{T}_{A(\phi^1_t,\cdots,\phi^{n-1}_t)}$$

  We use the following notation for negative iteration numbers:

$$\mathcal{T}_{\left[A(\phi^1_t,\cdots,\phi^{n-1}_t),\psi_t\right]^{\#-1}} = \mathcal{T}_{A(\phi^1_t,\cdots,\phi^{n-1}_t)^{\#-1}} \# \gamma^t_{\psi_t^{-1}} \# \mathcal{T}_{A(\phi^1_t,\cdots,\phi^{n-1}_t)} \# \gamma^t_{\psi_t}$$

  Finally, we extend our definition for higher iteration numbers:

$$\mathcal{T}_{\left[A(\phi^1_t,\cdots,\phi^{n-1}_t),\psi_t\right]^{\#k}} = \mathcal{T}_{\left[A(\phi^1_{\frac{t}{\sqrt{k}}},\cdots,\phi^{n-1}_{\frac{t}{\sqrt{k}}}),\psi_{\frac{t}{\sqrt{k}}}\right]} \# \overset{k}{\cdots} \# \mathcal{T}_{\left[A(\phi^1_{\frac{t}{\sqrt{k}}},\cdots,\phi^{n-1}_{\frac{t}{\sqrt{k}}}),\psi_{\frac{t}{\sqrt{k}}}\right]}$$

  and

$$\mathcal{T}_{\left[A(\phi^1_t,\cdots,\phi^{n-1}_t),\psi_t\right]^{\#-k}} = \mathcal{T}_{\left[A(\phi^1_{\frac{t}{\sqrt{k}}},\cdots,\phi^{n-1}_{\frac{t}{\sqrt{k}}}),\psi_{\frac{t}{\sqrt{k}}}\right]^{\#-1}} \# \overset{k}{\cdots} \# \mathcal{T}_{\left[A(\phi^1_{\frac{t}{\sqrt{k}}},\cdots,\phi^{n-1}_{\frac{t}{\sqrt{k}}}),\psi_{\frac{t}{\sqrt{k}}}\right]^{\#-1}}$$

  We call *step* $n-1$-*iteration number* to the integer k. We call *step* $n-1$-*iteration number* to the integer k and, as in the length 2 case, we refer to t as the *side of the tangle.*

Remark 22. *We can extend our definition of tangles by considering any reparametrization of the previously defined tangles. Therefore, we will use the term* tangle *for any curve such that its graph coincides with the one of a tangle.*

For completeness, we have introduced general tangles of commutators; i.e. tangles of general $1-$parametric families of diffeomorphisms. Nonetheless, hereafter we are going to work with a very specific type of such objects, tangles of commutators of coordinate vector field-flows (and their liftings from adapted-charts onto bracket-generating distributions). From now on, when talking about tangles we will implicitly refer to these specific type of tangles.

Remark 23. *It is a well known fact that curves with corners (concatenation of two straight segments) admit parametrizations that make them smooth. Nonetheless in that case they fail to be immersed curves. As a consequence, any tangle can be reparametrized in order to make it a (non-immersed) smooth curve.*

3.4.2. *Smoothing the tangles.* We can describe a way of smoothing the corners where the previously defined tangles failed to be smooth. Note that we just have to describe a way of replacing by some local model all the corners where the curves $\gamma_{\phi^{X_i}}, \gamma_{\phi^{X_j}}$ associated to different flows $\phi_t^{X_i}, \phi_t^{X_j}$ of coordinate vector fields X_i, X_j meet. For such purpose, define first the following time-dependent vector field:

$$Y_t = \begin{cases} X_i & t \in \left[0, \frac{1}{2} - \varepsilon\right], \\ \left(1 - \left(t - \varepsilon + \frac{1}{2}\right)\right) \cdot X_i + \left(t - \varepsilon + \frac{1}{2}\right) \cdot X_j & t \in \left[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right], \\ X_j & t \in \left[\frac{1}{2} + \varepsilon, 1\right]. \end{cases}$$

Consider now the integral curve $\gamma_{\phi^{Y_t}}$ with same starting point as $\gamma_{\phi^{X_i}}$. Note that this family of curves parametrized by ε is $C^\infty-$close to $\gamma_{\phi^{X_i}} \# \gamma_{\phi^{X_j}}$; i.e.

$$\gamma_{\phi^{Y_t}} \xrightarrow[\varepsilon \to 0]{||\cdot||_{C^\infty}} \gamma_{\phi^{X_i}} \# \gamma_{\phi^{X_j}}.$$

Thus, replacing all the corners by this local model choosing ε small enough yields the desired smoothness of the tangles while locally (and thus globally) remaining $C^\infty-$close to the original ones. Note also that they remain immersed after performing this local replacement.

As noted, in the context of bracket-generating distributions, Property (c) in the definition of *adapted charts* explains quantitatively how to move vertically in the fibration by describing a path in the base. Tangles can be understood as local sets of controllers that can be attached to base curves (or families of curves) in order to quantitatively control the vertical coordinates of the corresponding liftings. Therefore, given a family of curves, our goal is twofold. One one hand, we want to introduce these tangles parametrically and, thus, obtain a homotopy upstairs by integrating the homotopy in the base. On the other hand, we wish to carry out this process through embedded curves upstairs. We explain in **??** how we do this.

3.4.3. *Parametric tangles and embeddedness.* Let M be a smooth manifold endowed with a bracket-generating distribution $\mathcal{D}$ and let $\left(\gamma^k\right)_{\{k \in K\}}$ be a family of horizontal curves parametrized by a compact set K. Take adapted-coordinates $\phi : (\mathbb{D}_r(p), p) \to (\mathbb{R}^n, 0)$ and call $\left(\gamma^k\right)_\pi$ the projected family of curves onto the base of the fibration $d_p\pi_L(\phi_*\mathcal{D}) = \mathbb{R}^k$. The following Proposition gives a geometric recipee for how to parametrically introduce a curve that is $C^\infty-$close to a *non-iterated tangle of length 2* in a local neighborhood of $\gamma^k(t_0)$ by manipulating its projection γ_π^K. We show that this can be done without introducing self-intersections in the lifted homotopy.

Proposition 24. *Let X_i, X_j be two elements in the adapted framing such that they generate X_ℓ by Lie-Bracket; i.e. $[X_i(p), X_j(p)] = X_\ell(p)$. Fix a natural number $h \in \mathbb{N}$. For small enough $r > 0$ in the choice of adapted coordinates $\phi : (\mathbb{D}_r(p), p) \to (\mathbb{R}^n, 0)$, given $\gamma_\pi : [t_0 - \delta, t_0 + \delta] \to d_p\pi_L(\phi_*\mathcal{D}) = \mathbb{R}^k$ an integral curve for ϕ^{X_i}, there exists a homotopy $\gamma_\pi^u, u \in [0,1]$ of immersed curves in the base such that*

i) $\gamma_\pi^0 = \gamma_\pi$.
ii) γ_π^1 *is $C^\infty-$close to the tangle* $\mathcal{T}_{\left[\phi^{X_i}, \phi^{X_j}\right]^{\#h}}$.
iii) $\gamma_\pi^u|_{[t_0-\delta, t_0-\delta+\epsilon]} = \gamma_\pi|_{[t_0-\delta, t_0-\delta+\epsilon]}$ *and* $= \gamma_\pi^u|_{[t_0+\delta-\varepsilon, t_0+\delta]} = \gamma_\pi|_{[t_0+\delta-\varepsilon, t_0+\delta]}$ *for small enough* ε *and every* $u \in [0,1]$.
iv) *The lifting γ^u of γ_π^u onto the connection is embedded for every $u \in [0,1]$.*

Properties i) and ii) just state that the homotopy connects the original curve with a new curve at $u = 1$ which is C^∞−close to the tangle $\mathcal{T}_{[\phi^{X_i},\phi^{X_j}]^{\#h}}$ and to another curve at $u = 2$ which is C^∞−close to $\mathcal{T}_{[\phi^{X_j},\phi^{X_i}]^{\#h}}$. Property iii) says that this homotopy is relative to both endpoints. This, in particular, implies that the lifted homotopy through horizontal curves is also relative to the starting point. Finally, Property iv) guarantees that the lifting of the homotopy to the connection takes place through embedded horizontal curves.

Proof. We break the proof in two parts. First we deal with the base case (step 1-iteration number $h = 1$). Afterwards we will prove the general case.

Base case.

We first prove the case for non-iterated tangles; i.e. $h = 1$. Describe the homotopy through immersed cuves in the base as in Figure 2.

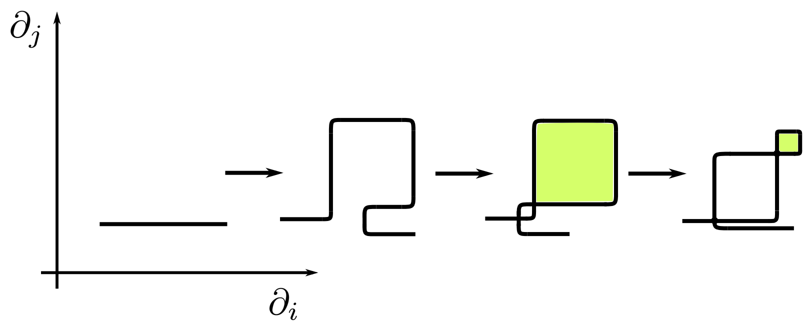


Figure 2. Parametric birth of a length 2 non-iterated tangle.

The first three depicted steps in the movie correspond to the homotopy $\gamma_\pi^u, u \in [0, 1]$, while these three steps together with the fourth one describe the whole homotopy $\gamma_\pi^u, u \in [0, 2]$. Points $i), ii)$ and $iii)$ follow automatically from the description of the homotopy taking place in the projected plane $\langle \partial_i, \partial_j \rangle$. Therefore all we have to check is that any pair of intersection points taking place in the base (at most two pairs, depending on the value of $u \in [0, 2]$) lift to different points upstairs. This fact can be achieved trivially if the rank of the distribution $\mathcal{D}$ is greater than 2, since we can use an additional coordinate ∂_z in order to perform the homotopy while remaining embedded already in the base. This fact is depicted in Figure 3, where it is shown how to avoid a crossing in the 3−plane $\langle \partial_i, \partial_j, \partial_z \rangle$ during the homotopy.

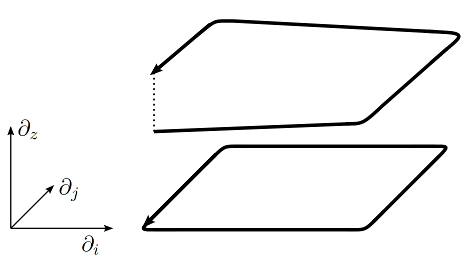


Figure 3. Increase of the additional coordinate ∂_z during the homotopy in order to achieve embeddedness in distributions $\mathcal{D}$ of rank greater than 2.

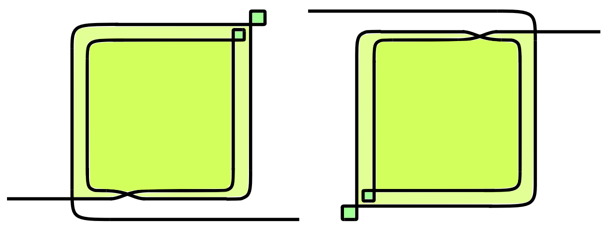
So, let us assume that the distribution is of rank 2 and, thus, because of the bracket-generating condition we have that $[X_i, X_\ell] = X_m$, where X_m is some other element in the adapted frame.

Let us denote by $(\gamma_\pi^u(t_1), \gamma_\pi^u(t_2))$ the 1−parametric family of pair of points corresponding to the upper-right autointersection in the homotopy in Figure 2. By Proposition 19 the difference in the values of the ∂_ℓ-coordinate between the liftings of the points $\gamma(t_1)$ and $\gamma(t_2)$ is $(1 + O(r))$, where A^u is the area enclosed by the curve $\gamma_\pi^u|_{[t_1,t_2]}$ in the plane $\langle \partial_i, \partial_j \rangle$. Therefore for a sufficiently small choice of $r > 0$, $A^u\,(1 + O(r))$ is a positive number.

Denote by $(\gamma^u_\pi(t'_1), \gamma^u_\pi(t'_2))$ the 1−parametric family of pair of points corresponding to the other autointersection in Figure 2. If the lifting $\gamma|_{[t'_1,t'_2]}$ of the curve $\gamma_\pi|_{[t'_1,t'_2]}$ projects onto the plane $\langle \partial_i, \partial_\ell\rangle$ as an opened curve then we are done, since this means that the ∂_ℓ-coordinates of the points $\gamma(t'_1)$ and $\gamma(t'_2)$ are different. Otherwise, we get a closed loop in the plane $\langle \partial_i, \partial_\ell\rangle$ that, again by Proposition 19, encloses area $B^u\,(1 + O(r))$. We conclude then that the ∂_m-coordinates of the liftings of both points are different by the same argument as for the other autointersection points.

General case.

We now show how to perform the homotopy in order to get h−times iterated tangles of length 2. The construction is based on the previous case. We first carry out half of the homotopy $\gamma^u, u \in [0, 1]$ as in the first three steps in Figure 2. We perform this same construction in a small segment from the lower branch of the tangle (see Figure 4a) while remaining inside the outermost tangle.



(A) Realization of $\mathcal{T}_{\left[\phi^{X_i},\phi^{X_j}\right]^{\#2}}$.

(B) Realization of $\mathcal{T}_{\left[\phi^{-X_j},\phi^{-X_i}\right]^{\#2}}$.

FIGURE 4. Length 2 tangles with step 1-iteration number h=2.

We repeat this process h times. The fact that all created pairs of intersection points lift as different points follows from the same argument as before; i.e. if $rank(\mathcal{D}) > 2$ then it is trivial. If $rank(\mathcal{D}) = 2$ the area argument applies again by noting that the big squares in each tangle contribute with positive area while the little ones contribute with negative area. Therefore, for a sufficiently small choice of $r > 0$ and if we perform the construction assuring that the big squares are sufficiently close one to each other, then the claim follows. For the other half of the homotopy we proceed inductively from the innermost tangle onto the outermost one. We shrink the innermost tangle arbitrarily until it gets very small comparing with the immediate one in which it is contained. Then we perform the homotopy $\gamma^u, u \in [1, 2]$ (movie from third to fourth diagrams in Figure 2) to this little tangle as described in the base case. As we already proved this does not create autointersection points in the lifting of this little tangle. Note also that since it is arbitrarily small, by the area argument, other autointersection points in the base do not get substantially affected by this homotopy and, therefore, the lifting of the whole curve remains embedded. We repeat this process inductively in the depth of the tangles until we reach the outermost one. Finally, we enlarge all the previously shrinked tangles starting from the innermost one and proceeding inductively on the depth of the tangles. By the same argument of the first half of the homotopy, this process of enlarging does not create new autointersection points upstairs since big squares contribute with very negative area in this case and the little ones with relatively small positive area. Properties $i), ii)$ and $iii)$ are satisfied by construction.

□

Proposition 25. *Let* $A(\phi^{X_1}_t, \cdots, \phi^{X_a}_t) = C((\phi^{X_1}_t, \cdots, \phi^{X_{a-1}}_t), \phi^{X_a})$ *be a Tangle expression such that* $C(X_1(p), \cdots, X_{a-1}(p)) = X_j(p)$, $A(X_1(p), \cdots, X_a(p)) = X_\ell(p)$. *For small enough* $r > 0$ *in the choice of adapted coordinates* $\phi : (\mathbb{D}_r(p), p) \to (\mathbb{R}^n, 0)$*, given* $\gamma_\pi : [t_0 - \delta, t_0 + \delta] \to d_p\pi_L(\phi_*\mathcal{D}) = \mathbb{R}^k$ *an integral curve for* ϕ^{X_i}*, there exist large enough natural numbers* $M_1, \cdots, M_a \in \mathbb{N}$ *such that if we call* $B(\phi^{X_1}_t, \cdots, \phi^{X_a}_t)$ *the formal tangle expression resulting from replacing the* $i-th$ *step iteration number* $(1 \le i \le n)$ *from* $A(\phi^{X_1}_t, \cdots, \phi^{X_a}_t)$ *by* M_i*, then*

There exists a homotopy $\gamma^u_\pi, u \in [0, 1]$ *of immersed curves in the base such that*

i) $\gamma^0_\pi = \gamma_\pi$.

ii) γ_π^1 *is* $C^\infty-$*close to a reparametrization of the tangle* $\mathcal{T}_{B(\phi_t^{X_1},\cdots,\phi_t^{X_a})}$.
iii) $\gamma_\pi^u|_{[t_0-\delta,t_0-\delta+\epsilon]} = \gamma_\pi|_{[t_0-\delta,t_0-\delta+\epsilon]}$ *and* $= \gamma_\pi^u|_{[t_0+\delta-\varepsilon,t_0+\delta]} = \gamma_\pi|_{[t_0+\delta-\varepsilon,t_0+\delta]}$ *for small enough* ε *and every* $u \in [0,1]$.
iv) *The lifting* γ^u *of* γ_π^u *onto the connection is embedded for every* $u \in [0,1]$.

The integers M_i *act as a scaling factor at each step of the tangle. It is a necessary condition* M *to be large enough in order to both get the* $C^\infty-$*closeness from Property* ii) *and also not to get certain branches of the curve interacting with some others during the homotopy.*

Proof. Denote by $\mathcal{T}_{A_\lambda(\phi_t^{X_1},\cdots,\phi_t^{X_\lambda})}$ or, in short, $\mathcal{T}_{A_\lambda}$, the length λ tangle of which $A(\phi_t^{X_1},\cdots,\phi_t^{X_a})$ is made up, omitting its iteration number; i.e. we take it as a non-iterated tangle. We can also assume that

$$T_{A_{\lambda+1}} = \left[\phi_t^{X_j}, \mathcal{T}_{A_\lambda}\right]$$

and

$$\left[X_j, A_\lambda(X_1,\cdots,X_\lambda)\right](p) = X_{\lambda+1}(p).$$

We will construct the homotopy inductively on the length of each subtangle expression involved in $A(\phi_t^{X_1},\cdots,\phi_t^{X_a})$, starting from the length 2-tangle non-iterated tangle $\mathcal{T}^2_{\left[\phi^{X_1},\phi^{X_2}\right]}$ of which $A(\phi_t^{X_1},\cdots,\phi_t^{X_a})$ is made up. Therefore, we parametrically introduce the tangle $\mathcal{T}_{\left[\phi^{X_1},\phi^{X_2}\right]^{\#M_1}}$ as in Proposition 24, followed by another tangle $\mathcal{T}_{\left[\phi^{X_1},\phi^{X_2}\right]^{\#-M_1}}$. This constitutes the first part of the homotopy $\gamma^u_{\pi,u\in[t_0-\delta,t_0-\delta+\frac{2\delta}{N}]}$, where M_1 will be specified at the end of the proof.

We show now to perform the homotopy $\gamma^u_{\pi,u\in[t_0-\delta+(\lambda-1)\frac{2\delta}{N},t_0-\delta+\lambda\frac{2\delta}{N}]}$ asumming we have already constructed the homotopy $\gamma^u_{\pi,u\in[t_0-\delta,t_0-\delta+(\lambda-1)\frac{2\delta}{N}]}$.

Now we perform a further homotopy by manipulating the projection of the curve in the plane $\langle\partial_j,\partial_\lambda\rangle$. First, we make the curve tangent to the coordinate vector field ∂_j in a small neighborhood. There, we carry out the homotopy $\gamma^u_{\pi,u\in[t_0-\delta+(\lambda-1)\frac{2\delta}{N},t_0-\delta+\lambda\frac{2\delta}{N}]}$ $M_\lambda - 1$ more times in $M_\lambda - 1$ separated points of the curve not touching the tangles constructed in previous steps (i.e. each of the $M_{\lambda-1}$ homotopies taking place one after the other both in the parameters u and t). Note that if M_λ is big enough, this can be achieved since the side of each tangle decreases with M_λ at a factor of $\frac{1}{\sqrt{M_\lambda}}$.

We bend the curve M_λ times as in Figure 5 in order to realize the M_λ-times iterated tangle $\mathcal{T}^{M_\lambda}_{\lambda+1}$ as in Figure 5, where each pair of boxes represents, following the recipee given in the previous step, one of the constructed M_λ homotopies realizing the tangle $\mathcal{T}^{\#M_{\lambda-1}}_{\lambda-1}$ by the first box and $\mathcal{T}^{\#-M_{\lambda-1}}_{\lambda-1}$ by the second one.

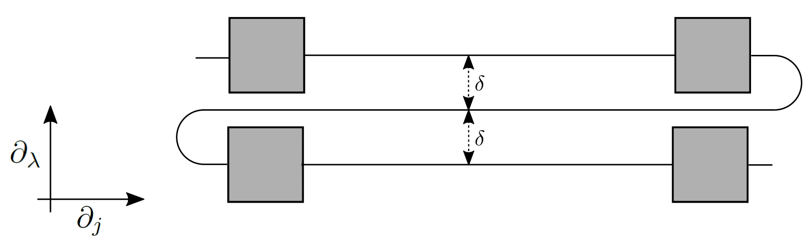


FIGURE 5. Schematic description of the inductive step in the construction of the homotopy ($M = 2$).

Note that as M_λ increases, the side of each tangle tends to zero and, thus, the boxes can be made arbitrarily small. The separation width between the bended branches δ can be made arbitrarily small too. Therefore, the resulting curve is clearly $C^\infty-$close to certain reparametrization of the tangle $\mathcal{T}^{\#M}_{\lambda+1}$.

Properties $i), ii)$ and $iii)$ follow from the construction. Property $iv)$ follows from two facts. First, note that the introduction of the tangles of length 2 do not create autointersections. Indeed, by

Proposition 24 introducing one tangle of length 2 does not create autointersection points and, since we introduced all these tangles one after the other in the parameters t, u, additional autointersection points are not created. Finally, for higher length tangles we do not introduce autointersection points even in the base. So, the claim follows.

□

4. The proof

To be done.

5. Appendix: Technical lemmas on commutators

Given a vector field Z on M we write ϕ_t^Z for its flow at time t. Given a pair of vector fields X and Y on M, we want to compare, in a quantitative manner, the flow of their Lie bracket $\phi_t^{[X,Y]}$ with the commutator of their flows ϕ_t^X and ϕ_t^Y. The contents of this appendix will be an important technical ingredient in the proof of our main theorems.

Given 1–parameter families (not necessarily subgroups) $(\varphi_t)_{t\in\mathbb{R}}$ and $(\psi_t)_{t\in\mathbb{R}}$ in the diffeomorphism group $\mathrm{Diff}(M)$, we can define in a given local chart the map $[\psi(t), \varphi(s)] = \varphi_s \circ \psi_t \circ \varphi_s^{-1} \circ \psi_t^{-1}(x)$. Note that if we take $s = t$ then this map is the commutator of the families taken for each time t, and we denote it by $[\psi_t, \varphi_t] := [\psi(t), \varphi(s)]$.

Lemma 26. *Write $X = \frac{\partial}{\partial t}\Big|_{t=0}\phi_t$ and $Y = \frac{\partial}{\partial t}\Big|_{t=0}\psi_t$ and assume $\phi_0 = \psi_0 = Id$. Then the following statements hold:*

i) There exists a 2−parametric family of diffeomorphisms $\varepsilon_{ts} = o(ts)$ such that

$$\psi_t \circ \phi_s(x) = \varepsilon_{ts} \circ \varphi_{ts}^{X+Y}(x)$$

ii) $\frac{\partial^k}{\partial t^k}\Big|_{t,s=0}[\psi(t), \phi(s)](x) = 0$ *for any* $k \in \mathbb{N}$,

iii) $\frac{\partial^k}{\partial s^k}\Big|_{t,s=0}[\psi(t), \phi(s)](x) = 0$ *for any* $k \in \mathbb{N}$,

iv) $\frac{1}{2}\frac{\partial^2}{\partial t \partial s}\Big|_{t=0}[\psi(t), \phi(s)](x) = [X, Y]$.

v) There exists a 2−parametric family of diffeomorphisms $\varepsilon(ts) = o(ts)$ such that

$$[\phi_t, \psi_s] = \varepsilon_{ts} \circ \varphi_{ts}^{[X,Y]}.$$

Proof. Part $i)$ follows from Taylor's Remainder Theorem applied to the composition map $\psi_t \circ \phi_s(x)$, $ii)$ and $iii)$ are obvious, $iv)$ follows from the definition of Lie Bracket and $v)$ follows from an application of Taylor's Remainder Theorem together with $ii), iii)$ and $iv)$. □

Remark 27. *A trivial but rather useful observation that we will eventually make use of is the following one. If ϕ_t, ψ_t, are two flows such that for some 1−parametric family of diffeomorphisms ε_t*

$$\phi_t = \varepsilon_t \circ \psi_t$$

then ε_t is also a flow since it is the composition of two flows $\varepsilon_t = \psi_t^{-1} \circ \phi_t$. Moreover, if $\varepsilon = o(t)$, then there exists another flow $\tilde{\varepsilon} = o(t)$ such that

$$\psi_t = \tilde{\varepsilon}_t \circ \phi_t.$$

For this last statement just note that

$$\phi_t = \varepsilon_t \circ \psi_t \Longrightarrow \phi_t \circ \psi_t^{-1} = o(t) \Longrightarrow \psi_t \circ \phi_t^{-1} = o(t)$$

and thus there exists some flow $\tilde{\varepsilon}_t = o(t)$ such that $\psi_t = \tilde{\varepsilon}_t \circ \phi_t$.

The following Lemma formalizes how errors inside a commutator of 1−parametric families of diffeomorphisms can be taken out the bracket expressions when comparing to the flow of the respective brackets.

Lemma 28. *Consider a flow $\varepsilon(t) = o(t)$ and write $X = \frac{\partial}{\partial t}\Big|_{t=0}\phi_t$, $Y = \frac{\partial}{\partial t}\Big|_{t=0}\psi_t$ and assume $\phi_0 = \psi_0 = Id$. Then there exists a 2−parametric family of diffeomorphisms $\tilde{\varepsilon}(ts) = o(ts)$ such that*

$$[\varphi_s^X, \varepsilon_t \circ \varphi_t^Y] = \tilde{\varepsilon}_{ts} \circ \varphi_{ts}^{[A,B]}.$$

Proof. The result follows from a direct application of points $i)$ and $v)$ from Lemma 26. □

Remark 29. *In particular, if we take $s = t$ in this previous lemma, we get that $\tilde{\varepsilon}_{t^2}$ is an actual flow, since it is the composition of two flows as in Remark 27,*

$$\tilde{\varepsilon}_{t^2} = \left(\varphi^{[A,B]}\right)^{-1}_{t^2} \circ [\varphi_t^X, \varepsilon_t \circ \varphi_t^Y]$$

The same remark is true for the family ε_{ts} in $v)$ from Lemma 26.

Lemma 30. *Consider $X = \frac{\partial}{\partial t}\Big|_{t=0}\phi_t$, $Y = \frac{\partial}{\partial t}\Big|_{t=0}\psi_t$ and assume $\phi_0 = \psi_0 = Id$. Then, if $\phi_t = \varepsilon_t \circ \psi_t$ for certain $\varepsilon_t = o(t)$, then there exists some other $\tilde{\varepsilon}_t = o(t)$ such that*

$$\phi_t = \psi_t \circ \tilde{\varepsilon}_t$$

Proof. By Lemma 28 there exists some $h_t = o(t)$ such that $h_{t^2} \circ \varepsilon_t^{-1} \circ \psi_t^{-1} \circ \varepsilon_t \circ \psi_t = Id$. So, we have that $h_{t^2} \circ \varepsilon_t^{-1} \circ \psi_t^{-1} = \phi_t^{-1}$. The result follows from taking inverse flows at both side of the equation. □

Lemma 31. *Write $X = \frac{\partial}{\partial t}\Big|_{t=0}\phi_t$ and $Y = \frac{\partial}{\partial t}\Big|_{t=0}\psi_t$ and assume $\phi_0 = \psi_0 = Id$. Then we have the following estimation:*

$$\phi_t^{-1} \circ \psi_s \circ \phi_t = \varepsilon_{ts} \circ \varphi_s^{X + t[X,Y]}.$$

Proof. First, note that

$$\phi_t^{-1} \circ \psi_s \circ \phi_t = \psi_s \circ [\phi_t, \psi_s] = \psi_s \circ \tilde{\varepsilon}_{ts} \circ \varphi_{ts}^{[X,Y]} = \psi_s \circ \tilde{\varepsilon}_{ts} \circ \varphi_s^{t[X,Y]},$$

where the second equality follows by Lemma 26. Nevertheless, this implies the existence of a 2−parametric family of diffeomorphisms $\epsilon_{ts} = o(ts)$ such that $\psi_s \circ \varphi_s^{t[X,Y]} \circ \epsilon_{ts}$. But, from point $i)$ in Lemma 26 applied to the composition $\psi_s \circ \varphi_s^{t[X,Y]}$, there exist some other $\tilde{\epsilon}_{ts} = o(ts), \varepsilon_{ts} = o(ts)$ such that

$$\psi_s \circ \varphi_s^{t[X,Y]} \circ \epsilon_{ts} = \varphi_s^{Y + t[X,Y]} \circ \tilde{\epsilon}_{ts} = \varepsilon_{ts} \circ \varphi_s^{Y + t[X,Y]}$$

thus yielding the claim. □

Definition 32. *Given two 1−parametric families of diffeomorphisms φ_s, ψ_t, we define their k−th iterated commutator $[\psi_t, \varphi_s]^{\#k}$ as follows*

$$[\psi_t, \varphi_s]^{\#k} := \left(\left(\varphi^{-1}_{\frac{s}{\sqrt{k}}}\right) \circ \left(\psi^{-1}_{\frac{t}{\sqrt{k}}}\right) \circ \left(\varphi_{\frac{s}{\sqrt{k}}}\right) \circ \left(\psi_{\frac{t}{\sqrt{k}}}\right)\right)^k.$$

The reason for parametrizing the flows by $\frac{s}{\sqrt{k}}$ in the definition of iterated commutator is justified by the following proposition.

Proposition 33. *Write $X = \frac{\partial}{\partial t}\Big|_{t=0}\phi_t$ and $Y = \frac{\partial}{\partial t}\Big|_{t=0}\psi_t$ and assume $\phi_0 = \psi_0 = Id$. Then there exists $\varepsilon_t = o(t)$ such that*

$$[\psi_t, \varphi_s]^{\#k} = \varepsilon_{ts} \circ \varphi_{ts}^{[X,Y]}$$

Proof. By the definition of the k−th iterated commutator of flows and by Lemma 26, there exist flows $\varepsilon_t^1 = o(t), \cdots, \varepsilon_t^k = o(t)$ such that

$$[\psi_t, \varphi_t]^{\#k} = \left(\varepsilon^1_{ts/k} \circ \varphi^{[X,Y]}_{ts/k}\right) \circ \cdots \circ \left(\varepsilon^k_{ts/k} \circ \varphi^{[X,Y]}_{ts/k}\right)$$

But by Lemma 30 there exist flows $\tilde{\varepsilon}^1_t = o(t), \cdots, \tilde{\varepsilon}^k_t = o(t)$ such that

$$\left(\varepsilon^1_{ts/k} \circ \varphi^{[X,Y]}_{ts/k}\right) \circ \cdots \circ \left(\varepsilon^k_{ts/k} \circ \varphi^{[X,Y]}_{ts/k}\right) = \tilde{\varepsilon}^1_{ts/k} \circ \cdots \circ \tilde{\varepsilon}^k_{ts/k} \circ \left(\varphi^{[X,Y]}_{ts/k}\right)^k$$

and so the claim follows. □

With this battery of technical results at our disposal, we can compare how taking a given bracket expression behaves with respect to taking flows [16, Theorem 1]:

Proposition 34. *Let $X_1, X_2, \cdots, X_\lambda$ be (possibly repeated) vector fields on a manifold M. Then, for any bracket expression $A(-, \cdots, -)$ of length λ there exists a flow $\varepsilon_t = o(t)$ such that*

$$A\left(\varphi^{X_1}_t, \cdots, \varphi^{X_\lambda}_t\right) = \varepsilon_{t^\lambda} \circ \phi^{A(X_1,\cdots,X_\lambda)}_{t^\lambda}.$$

Proof. We proceed by induction on the length of the formal bracket–expression:

- For $k = 2$ the result holds by Proposition 33.
- The induction hypothesis says that the statement holds for all expressions of length $k' < k$. By definition, if $A(-, \cdots, -)$ is an expression of length k, there exists $i < k$ and an integer m such that $A(X_1, \cdots, X_k) = [B(X_1, \cdots, X_i), C(X_{i+1}, \cdots, X_k)]^{\#m}$, with $B()$ of length i and $C()$ of length $k - i$. Computing we see that there are flows $f_t = o(t)$ and $g_t = o(t)$ such that:

$$A(\phi^{X_1}_t, \cdots, \phi^{X_k}_t) = \left[B(\phi^{X_1}_t, \cdots, \phi^{X_i}_t), C(\phi^{X_{i+1}}_t, \cdots, \phi^{X_k}_t)\right]^{\#m} \stackrel{\text{IH}}{=} \left[f_{t^i} \circ \phi^{B(X_1,\cdots,X_i)}_{t^i}, g_{t^{k-i}} \circ \phi^{C(X_{i+1},\cdots,X_k)}_{t^{k-i}}\right]^{\#m} \tag{1}$$

By an application of Proposition 33 first, and by Lemma 28, there exists a flow $\varepsilon_t = o(t)$ such that

$$\left[f_t \circ \phi^{B(X_1,\cdots,X_i)}_{t^i}, g_t \circ \phi^{C(X_{i+1},\cdots,X_k)}_{t^{k-i}}\right]^{\#m} = \varepsilon_{t^k} \circ \left[\phi^{B(X_1,\cdots,X_i)}_{t^i}, \phi^{C(X_{i+1},\cdots,X_k)}_{t^{k-i}}\right] \tag{2}$$

But, since $\left[\phi^{B(X_1,\cdots,X_i)}_{t^i}, \phi^{C(X_{i+1},\cdots,X_k)}_{t^{k-i}}\right] = \phi^{A(X_1,\cdots,X_k)}_{t^k}$ the result follows from combining (1) and (2).

□

References

[1] J. Adachi. *Classification of horizontal loops in standard Engel space.* Int. Math. Res. Not. 2007 Vol. 3 (2007).
[2] V.I. Arnol'd, S.P. Novikov. *Dynamical systems VII.* Enc. of Math. Sci., vol. 16, Springer-Verlag, Berlin Heidelberg, 1994.
[3] R.L. Bryant, L. Hsu. *Rigidity of integral curves of rank 2 distributions.* Invent. Math. 114 (1993), no. 2, 435–461.
[4] E. Cartan. *Sur quelques quadratures dont l'élément différentiel contient des fonctions arbitraires.* Bull. Soc. Math. France 29 (1901), 118–130.
[5] R. Casals, J.L. Pérez, Á. del Pino, F. Presas. *Existence h-principle for Engel structures.* Invent. Math. 210 (2017), 417-451
[6] R. Casals, Á. del Pino, F. Presas. *Loose Engel structures.* arXiv:1712.09283
[7] Y Eliashberg, D. Kotschick, E. Murphy, T. Vogel. *Engel structures.* AIM Workshop report (2017).
[8] Y. Eliashberg, N. Mishachev. *Introduction to the h-principle.* Graduate Studies in Mathematics, 48. American Mathematical Society, Providence, RI, 2002.
[9] Y. Eliashberg, N. Mishachev. *Wrinkled embeddings.* Foliations, geometry, and topology, 207–232, Contemp. Math., 498, Amer. Math. Soc., Providence, RI, 2009.
[10] E. Fernández, J. Martínez-Aguinaga, F. Presas. *Fundamental groups of formal legendrian and horizontal embedding spaces.* In preparation.
[11] G. Frobenius, *Uber das Pfaffsche Problem.* J. für Reine und Agnew. Math., 82 (1877) 230–315.
[12] Z. Ge. *Horizontal path spaces and Carnot-Carathéodory metrics.* Pacific J. Math. Volume 161, Number 2 (1993), 255-286.
[13] H. Geiges. *Horizontal loops in Engel space.* Math. Ann. Volume 342, Issue 2, (2008) 291–296.
[14] M. Gromov. *Partial differential relations.* Ergebnisse Math. und ihrer Gren. , Springer–Verlag, Berlin, 1986.
[15] M. Gromov, *Carnot-Carathéodory spaces seen from within.* Sub-Riemannian geometry, Progr. Math., vol. 144, Birkhäuser, Basel, 1996.

[16] M. Mauhart, P. W. Michor. *Commutators of flows and fields.* Archivum Mathematicum (Brno) 28 (1992), 228–236.
[17] R. Montgomery. *Engel deformations and contact structures.* Northern California Symplectic Geometry Seminar, 103–117, Amer. Math. Soc. Transl. Ser. 2, 196, Amer. Math. Soc., Providence, RI, 1999.
[18] D. McDuff. *Applications of convex integration to symplectic and contact geometry.* Ann. Inst. Fourier 37 (1987), 107–133.
[19] E. Murphy. *Loose Legendrian Embeddings in high dimensional contact manifolds*, arXiv:1201.2245.
[20] Á. del Pino, F. Presas. *Flexibility for tangent and transverse immersions in Engel manifolds* arXiv:1609.09306.
[21] Á. del Pino, T. Vogel. *The Engel-Lutz twist and overtwisted Engel structures.* arXiv:1712.09286
[22] T. Vogel. *Existence of Engel structures.* Ann. of Math. (2) 169 (2009), no. 1, 79–137.

Universidad Complutense de Madrid, Facultad de Matemáticas, Plaza de Ciencias 3, 28040 Madrid, Spain

Email address: javier.martinez.aguinaga@icmat.es

Utrecht University, Department of Mathematics, Budapestlaan 6, 3584 Utrecht, The Netherlands

Email address: a.delpinogomez@uu.nl