TU
WIEN

TECHNISCHE
UNIVERSITÄT
WIEN

VIENNA
UNIVERSITY OF
TECHNOLOGY

DIPLOMARBEIT

The Poincaré Lemma
in Subriemannian Geometry

ausgeführt am Institut für

Wirtschaftsmathematik

der Technischen Universität Wien

unter Anleitung von Ao.Univ.Prof. Mag. Dr. Josef Teichmann

durch

Philipp Harms

Josefstädterstraße 14/54
1080 Wien

_____   _____
Datum   Unterschrift

Abstract

This work is a short, self-contained introduction to subriemannian geometry with special emphasis on Chow's Theorem. As an application, a regularity result for the the Poincaré Lemma is presented.

At the beginning, the definitions of a subriemannian geometry, horizontal vectorfields and horizontal curves are given. Then the question arises: Can any two points be connected by a horizontal curve? Chow's Theorem gives an affirmative answer for bracket generating distributions. (A distribution is called bracket generating if horizontal vectorfields and their iterated Lie brackets span the whole tangent space.)

We present three different proofs of Chow's Theorem; each one is interesting in its own. The first proof is based on the theory of Stefan and Sussmann regarding integrability of singular distributions. The second proof is elementary and gives some insight in the shape of subriemannian balls. The third proof is based on infinite dimensional analysis of the endpoint map.

Finally, the study of the endpoint map allows us to prove a regularity result for the Poincaré Lemma in a form suited to subriemannian geometry: If for some $r \geq 0$ all horizontal derivatives of a given function f are known to be r times continuously differentiable, then so is f.

Sections 1 to 3 are the common work of Martin Bauer and Philipp Harms.

Contents

1 The definition of a subriemannian geometry

For the basic definitions in differential geometry, see for example the book of Bröcker and Jänich [4]. A more detailed treatment, including calculus in Banach spaces and Banach manifolds, can be found in the book of Abraham et al. [1]. A good introduction to subriemannian geometry is the book of Montgomery [8]. Throughout this paper, manifolds are assumed to be paracompact and smooth. Vectorfields and mappings between manifolds are assumed to be smooth if not stated otherwise.

Definition 1.1. Let M be a manifold. Suppose that for each $x \in M$ we are given a sub vector space H_x of $T_x M$. The disjoint union $H = \coprod_{x \in M} H_x$ is called a distribution on M. We call tangent vectors in H horizontal.

Definition 1.2. A distribution H is called regular if the dimension of H_x is locally constant, and singular otherwise.

Definition 1.3. Let $\mathfrak{X}(M)$ denote the set of all vectorfields on M, and let $\mathfrak{X}_{loc}(M)$ be the set of local vectorfields, i.e.

$$\mathfrak{X}_{loc}(M) = \bigcup \{\mathfrak{X}(U) : U \text{ is an open subset of } M\} .$$

To avoid the need for remarks such as "provided the domains of the vectorfields intersect", we shall declare the vectorfield defined on the empty set to be an element of $\mathfrak{X}_{loc}(M)$.

Definition 1.4. A local vectorfield is called horizontal if it is horizontal at every point where it is defined. Let $\mathfrak{X}_H(M)$ denote the set of horizontal vectorfields and $\mathfrak{X}_{loc,H}(M)$ the set of local horizontal vectorfields on M.

Definition 1.5. We say that a set of local vectorfields $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ spans a distribution H if for all points $x \in M$, H_x is the linear hull of the set $\{X(x) : X \in \mathcal{V}\}$. We say that a distribution H is a smooth if it is spanned by $\mathfrak{X}_{loc,H}(M)$.

Remark 1.6. Every set of vectorfields spans a distribution, which is obviously smooth (the linear hull of the empty set is the vector space $\{0\}$).

Definition 1.7. A distribution H is called locally finitely generated if for every point $x \in M$ there is neighbourhood U of x and a finite set of vectorfields $X_1, \ldots, X_k \in \mathfrak{X}_H(U)$ spanning H on U.

Definition 1.8. Let H be a smooth distribution on M. A subriemannian metric $\langle \cdot, \cdot \rangle$ on H is a map that assigns to each point $x \in M$ an inner product $\langle \cdot, \cdot \rangle_x$ on H_x. Let $\|X\|_x$ designate $\sqrt{\langle X, X \rangle_x}$ for $X \in H_x$. We say that a subriemannian metric is smooth if for all local horizontal vectorfields X and Y, $x \mapsto \langle X(x), Y(x) \rangle_x$ is smooth.

The smoothness of $\langle \cdot, \cdot \rangle$ does not imply that $\|\cdot\| : H \to \mathbb{R}$ is continuous, when H is endowed with the subspace topology of TM.

Example 1.9. To give an example, let H be the distribution on $\mathbb{R}$ spanned by the vectorfield $x\frac{\partial}{\partial x}$, and let $\langle\cdot,\cdot\rangle$ be such that

$$\left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial x} \right\rangle_x = \frac{1}{x} \quad \text{for} \quad x \neq 0.$$

Then $\langle\cdot,\cdot\rangle$ is smooth, but is not continuous, where H is given the subspace topology. To see that $\langle\cdot,\cdot\rangle$ is smooth, let X_1, X_2 be horizontal vectorfields defined near zero, $X_i(x) = f_i(x) \left.\frac{\partial}{\partial x}\right|_x$ with f_i smooth, $f_i(0) = 0$. Then

$$\langle X_1(x), X_2(x)\rangle_x = \frac{1}{x}f_1(x)f_2(x) = \frac{1}{x}\int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f_1(xs)f_2(xs))\mathrm{d}s =$$
$$= \int_0^1 (f_1'(xs)f_2(xs) + f_1(xs)f_2'(xs)) \ \mathrm{d}s \in C^\infty.$$

To see that $\|\cdot\|$ is not continuous, take $X_n = \frac{1}{\sqrt{n}} \left.\frac{\partial}{\partial x}\right|_{\frac{1}{n}}$. Then X_n converges to the zero tangent vector at $x = 0$, but $\|X_n\| = 1$.

Definition 1.10. A subriemannian manifold is a manifold endowed with a smooth distribution and a smooth subriemannian metric on this distribution.

From now on, we will only consider smooth distributions and smooth subriemannian metrics.

Definition 1.11. A curve γ is called absolutely continuous if it is absolutely continuous in any chart, i.e. for any chart (U, u) of M and any interval $[t_0, t_1]$ such that $\gamma([t_0, t_1]) \subset U$, the curve $u \circ \gamma|[t_0, t_1]$ is absolutely continuous.

Definition 1.12. A curve $\gamma : [a, b] \rightarrow M$ is said to be controlled by local vectorfields $X_1, \ldots, X_k$ defined near $\gamma([a, b])$ if γ is absolutely continuous and if there are L^1-functions $u_1, \ldots, u_k$ such that the equation

$$\dot\gamma(t) = \sum_{i=1}^k u_i(t) \ X_i(\gamma(t))$$

holds almost everywhere. $u_1, \ldots, u_k$ are called the controls of γ with respect to the vectorfields $X_1, \ldots, X_k$. We also say that the controls steer the point $\gamma(a)$ to the point $\gamma(b)$.

Definition 1.13. A curve γ is called horizontal if it is a concatenation of curves that are controlled by horizontal vectorfields.

Definition 1.14. For a point $x \in M$, the accessible set $Acc(x)$ is the set of all points which can be reached by a horizontal curve starting at x. A set $S \subset M$ is called horizontally pathwise connected if any two points in S can be connected by a horizontal curve lying in S.

Lemma 1.15. *If γ is a horizontal curve in a subriemannian manifold, then the function $t \mapsto \|\dot\gamma(t)\|_{\gamma(t)}$ is integrable.*

4

Proof. A horizontal curve is a concatenation of curves controlled by horizontal vectorfields. If γ is controlled by horizontal vectorfields $X_1, \ldots, X_k$ with L^1-controls $u_1, \ldots, u_k$, we have

$$\|\dot{\gamma}\|^2_{\gamma(t)} = \sum_{i=1}^{k} \sum_{j=1}^{k} |u_i(t)|\, |u_j(t)| \left| \langle X_i(\gamma(t)), X_j(\gamma(t)) \rangle_{\gamma(t)} \right|$$

$$\leq K \left(\sum_{i=1}^{k} |u_i(t)| \right) \left(\sum_{j=1}^{k} |u_j(t)| \right) = K \left(\sum_{i=1}^{k} |u_i(t)| \right)^2,$$

where K bounds the continuous functions

$$t \mapsto \left| \langle X_i(\gamma(t)), X_j(\gamma(t)) \rangle_{\gamma(t)} \right|, \quad 1 \leq i, j \leq k. \qquad \square$$

Definition 1.16. If M is a subriemannian manifold, we define the length of a horizontal curve $\gamma : [a, b] \to M$ by

$$l(\gamma) = \int_a^b \|\dot{\gamma}(t)\|_{\gamma(t)} \, \mathrm{d}t.$$

Definition 1.17. The subriemannian distance, also called Carnot-Carathéodory distance, between two points x and y on a subriemannian manifold is given by

$$d(x, y) = \inf l(\gamma),$$

where the infimum is taken over all horizontal curves that connect x to y. The distance is infinite if there is no such curve. The subriemannian ball of radius ϵ centered at $x \in M$ is denoted by

$$B(\epsilon, x) = \{y \in M : d(x, y) < \epsilon\}.$$

Definition 1.18. A horizontal curve γ that connects x to y is called a geodesic if $l(\gamma) = d(x, y)$.

1.1 Accessibility and the set of horizontal curves

We have defined the accessible set as the set of points reachably by a certain class of curves, namely the curves as in Definition 1.13. Theorems 3.20 and 3.66 confirm that our choice of admissible curves is a good one. It can be seen in the proofs of these theorems that the important point is that we require a curve to be controlled by horizontal vectorfields. Bellaïche [3] shows that instead of L^1-controls, we can use any reasonable smaller set of controls without changing the notion of accessibility:

Theorem 1.19. *Let C be a dense subspace in $L^1([0, 1]; \mathbb{R}^k)$. Then any point accessible from x by means of controls in $L^1([0, 1]; \mathbb{R}^k)$ is also accessible from x by means of controls in C.*

C can even be the set of piecewise constant functions as will be seen in Section 3.1.2 about orbits.

It is a different question whether we can enlarge the class of horizontal curves without changing the notion of accessibility. For example it would be tempting to drop the concept of controllability and instead define a curve to be horizontal if it is absolutely continuous and if its derivative lies in H whenever it exists. This works out for regular distributions as will be shown in Lemma 1.23 below, but we will see in the following example that generally this might change the notion of accessibility.

Example 1.20. Let H be the distribution on $\mathbb{R}$ that has rank zero at points $x \le 0$ and rank one at points $x > 0$. Then H is smooth. It is easy to construct an absolutely continuous curve γ with $\gamma(0) = 0$ and $\gamma(t) > 0$ for $t > 0$ such that $\dot{\gamma}(0)$ either vanishes or does not exist. But such a curve is not controllable by any horizontal vectorfield because controllability implies that the curve is constant. To show this, let us assume that γ is controlled by a horizontal vectorfield X. Then γ satisfies the differential equation

$$\dot{\gamma}(t) = u(t)\, X(\gamma(t)), \quad \gamma(0) = 0$$

with $u \in L^1([0,1];\mathbb{R})$. By Theorem 3.57 the above differential equation has a unique solution for small t. But also the constant curve is a solution to the above equation, and therefore γ is constant.

Lemma 1.21. *A regular, smooth distribution is locally finitely generated, and it is a vector subbundle of the tangent bundle.*

Proof. Fix an arbitrary point $x \in M$, and let k be the rank of the distribution H near x. Because H is smooth, there are local horizontal vectorfields $X_1, \ldots, X_k$ such that $X_1(x), \ldots, X_k(x)$ span H_x. $X_1, \ldots, X_k$ are linearly independent also in some neighbourhood of x, so they span H in this neighbourhood, and therefore H is locally finitely generated.

Now take arbitrary vectorfields $X_{k+1}, \ldots, X_n$ defined near x such that the tangent vectors $X_1(x), \ldots, X_n(x)$ are linearly independent. Then $X_1, \ldots, X_n$ are linearly independent also in some neighbourhood of x, and so they form a local frame of the tangent space near x. To every tangent vector in the tangent space near x, we can associate its coordinates in terms of this local frame. This yields a local trivialization for H, and so H is a subbundle of TM. $\square$

Lemma 1.22. *If $\langle \cdot, \cdot \rangle$ is a smooth subriemannian metric on a regular smooth distribution H, then $\langle \cdot, \cdot \rangle$ can be extended to a riemannian metric.*

Proof. Locally, we can find a frame $X_1, \ldots, X_n$ for TM such that $X_1, \ldots X_k$ are horizontal and orthonormal with respect to $\langle \cdot, \cdot \rangle$. On the common domain of definition of the vectorfields X_i, we can define a riemannian metric g via

$$g(X_i, X_j) = \begin{cases} 0 & \text{for } i \ne j \\ 1 & \text{for } i = j. \end{cases}$$

To get a riemannian metric defined on all of M, we patch together the obtained local riemannian metrics using a partition of unity. $\square$

Lemma 1.23. *If H is a regular smooth distribution, then a curve is horizontal if and only if it is absolutely continuous and its derivative lies in H whenever it exists.*

Proof.

(1) To show the one implication, let γ be a horizontal curve. Then by definition γ is a concatenation of curves controlled by horizontal vectorfields. Therefore it is absolutely continuous and its derivative lies in H whenever it exists.

(2) It remains to show the other implication, so let $\gamma : [a, b] \to M$ be an absolutely continuous curve whose derivative is horizontal whenever it exists. Let k be the rank of H near $\gamma([a, b])$. We can find $a = t_0 < t_1 < \cdots < t_n = b$ such that $\gamma([t_i, t_{i+1}])$ is contained in some open set U_i, U_i lies in the domain of some chart, and on each set U_i there is a set of orthonormal vectorfields $X_{i,1}, \ldots, X_{i,k} \in \mathfrak{X}_H(U_i)$ spanning H. γ is absolutely continuous, so by definition there are functions $v_{i,l} \in L^1([t_i, t_{i+1}]; \mathbb{R})$ such that for $t_i \leq t \leq t_{i+1}$ we have

$$\dot{\gamma}(t) = \sum_{l=1}^{n} v_{i,l}(t) \frac{\partial}{\partial x_l}(\gamma(t)).$$

We can extend the given subriemannian metric to a riemannian metric and use the same symbol $\langle \cdot, \cdot \rangle$ for it. Then for $t_i \leq t \leq t_{i+1}$ we have

$$\dot{\gamma}(t) = \sum_{j=1}^{k} \langle \dot{\gamma}(t), X_{i,j}(\gamma(t)) \rangle \; X_{i,j}(\gamma(t))$$

$$= \sum_{j=1}^{k} \underbrace{\left(\sum_{l=1^n} v_{i,l}(t) \left\langle \frac{\partial}{\partial x_l}(\gamma(t)), X_{i,j}(\gamma(t)) \right\rangle \right)}_{\in L^1([t_i, t_{i+1}]; \mathbb{R})} X_{i,j}(\gamma(t)),$$

so on each interval $[t_i, t_{i+1}]$, γ is controlled by horizontal vectorfields. $\qquad \square$

2 Basics of differential geometry

2.1 Vectorfields and flows

The following basic definitions and Lemmas are taken from Michor [7].

Definition 2.1. A local diffeomorphism on M is a C^∞-diffeomorphism from an open subset U of M onto an open subset V of M. If $\varphi_i : U_i \to V_i (i = 1, 2)$ are local diffeomorphisms, then the composite $\varphi_1 \circ \varphi_2$ is a local diffeomorphism, with domain $\varphi_2^{-1}(U_1)$ and range $\varphi_1(V_1 \cap U_1)$. The inverse of φ_1 is denoted by φ_1^{-1}, and is a local diffeomorphism with domain V_1 and range U_1. The formal laws

$$(\varphi_1 \circ \varphi_2) \circ \varphi_3 = \varphi_1 \circ (\varphi_2 \circ \varphi_3) \quad \text{and} \quad (\varphi_1 \circ \varphi_2)^{-1} = \varphi_2^{-1} \circ \varphi_1^{-1}$$

are clearly valid.

Definition 2.2. A curve of local diffeomorphisms through the identity on M is a smooth mapping $\varphi : U \to M$ defined on some open set $U \subset \mathbb{R} \times M$ such that

(a) U contains $\{0\} \times M$ and for all points $x \in M$, the set $U \cap \mathbb{R} \times \{x\}$ is open and connected.

(b) $\varphi_t : \; U \cap (\{t\} \times M) \to M, \; x \mapsto \varphi(t, x)$ is a local diffeomorphism.

(c) $\varphi_0 = \mathrm{Id}_M$.

φ is called a local flow if in addition, it satisfies

(d) $\varphi(t, \varphi(s, x)) = \varphi(t + s, x)$ for all t, s, x such that both sides of the equation are defined.

Definition 2.3. The infinitesimal generator of a local flow $\varphi : U \to M$ is the vectorfield X_φ given by $x \mapsto \frac{\partial}{\partial t}\big|_0 \varphi(t, x)$.

The theory of flows and vectorfields says that for each vectorfield X, there is a unique local flow whose infinitesimal generator is X and whose domain of definition is maximal with respect to inclusion.

Definition 2.4. This flow is called the maximal local flow of X and will be denoted by Fl^X.

Remark 2.5. As in Definition 2.2, we will use the notations $Fl^X(t, x)$ and $Fl^X_t(x)$ to denote the same thing.

Remark 2.6. If X is only a local (and not a global) vectorfield, then its maximal local flow is a curve of local diffeomorphism through the identity on the domain of definition of X.

If $f : M \to M$ is a diffeomorphism, then for any vectorfield $X \in \mathfrak{X}(M)$ the following mappings are vectorfields as well:

$$f^* X := Tf^{-1} \circ X \circ f, \quad f_* X := Tf \circ X \circ f^{-1}.$$

Definition 2.7. In the above situation, $f^* X$ is called the pullback of X along f, and $f_* X$ the pushforward of X along f.

Definition 2.8. Let $f : M \to N$ be a smooth mapping. Two vectorsfields $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(M)$ are called f-related, if $Tf \circ X = Y \circ f$ holds, i.e. if the following diagram commutes:

$$
\begin{array}{ccc}
TM & \xrightarrow{\;Tf\;} & TN \\
\big\uparrow{\scriptstyle X} & & \big\uparrow{\scriptstyle Y} \\
M & \xrightarrow{\;f\;} & N
\end{array}
$$

Lemma 2.9. *Let $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(M)$ be f-related vectorfields for a smooth mapping $f : M \to N$. Then we have $f \circ Fl^X_t = Fl^Y_t \circ f$, whenever both sides are defined. In particular, if f is a diffeomorphism, we have $Fl^{f^* Y}_t = f^{-1} \circ Fl^Y_t \circ f$.*

Proof. We have

$$\frac{\partial}{\partial t} f \circ Fl_t^X = Tf \circ \frac{\partial}{\partial t} Fl_t^X = Tf \circ X \circ Fl_t^X = Y \circ f \circ Fl_t^X$$

and $f(Fl^X(0, x)) = f(x)$. So $t \mapsto f(Fl^X(t, x))$ is an integral curve of the vectorfield Y on N with initial value f(x), so we have

$$f(Fl^X(t, x)) = f(Fl^Y(t, f(x))) \text{ or } f \circ Fl_t^X = Fl_t^Y \circ f. \qquad \square$$

Remark 2.10. By the prescription $\mathcal{L}_X(f)(x) := X(x)(f)$, each vectorfield $X \in \mathfrak{X}(M)$ defines a derivation $\mathcal{L}_X$ on the algebra $C^\infty(M)$ of smooth functions on M, i.e. a mapping $\mathcal{L}_X : C^\infty(M) \to C^\infty(M)$ with $\mathcal{L}_X(f \cdot g) = \mathcal{L}_X(f) \cdot g + f \cdot \mathcal{L}_X(g)$. This defines a vectorspace isomorphism between $\mathfrak{X}(M)$ and the space of derivations on $C^\infty(M)$. Therefore we can identify a vectorfield X with the derivation $\mathcal{L}_X$. We will often write X instead of $\mathcal{L}_X$ when the meaning is clear from the context.

Definition 2.11. For $X, Y \in \mathfrak{X}(M)$, the Lie bracket $[X, Y] \in \mathfrak{X}(M)$ is the unique vectorfield that satisfies the formula $\mathcal{L}_{[X,Y]} = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X$. The definition of the Lie bracket can be extended to arbitrary bracket expressions $B(X_1, \ldots, X_k)$ in the obvious way.

Lemma 2.12. *Let $f : M \to N$ be a smooth map. If $X_1 \in \mathfrak{X}(M)$ is f-related to $Y_1 \in \mathfrak{X}(N)$ and $X_2 \in \mathfrak{X}(M)$ is f-related to $Y_2 \in \mathfrak{X}(N)$, then $\lambda_1 X_1 + \lambda_2 X_2$ is f-related to $\lambda_1 Y_1 + \lambda_2 Y_2$ and $[X_1, X_2]$ is f-related to $[Y_1, Y_2]$.*

Proof. The first assertion is immediate. To prove the second, we view tangent vectors as derivations on the algebra of germs of smooth functions. Let h be the germ of a smooth function on N at $f(x)$. Then $h \circ f$ is the germ of a smooth function on M at x. By assumption, we have $Tf \circ X_i = Y_i \circ f$ for $i = 1, 2$, thus:

$$\big(X_i(x)\big)(h \circ f) = \big(T_x f . X_i(x)\big)(h) = \big((Tf \circ X_i)(x)\big)(h) =$$
$$= \big((Y_i \circ f)(x)\big)(h) = \big(Y_i(f(x))\big)(h).$$

So we have $\mathcal{L}_{X_i}(h \circ f) = \big(\mathcal{L}_{Y_i}(h)\big) \circ f$, and therefore

$$\mathcal{L}_{[X_1, X_2]}(h \circ f) = \mathcal{L}_{X_1}\big(\mathcal{L}_{X_2}(h \circ f)\big) - \mathcal{L}_{X_2}\big(\mathcal{L}_{X_1}(h \circ f)\big)$$
$$= \mathcal{L}_{X_1}\Big(\big(\mathcal{L}_{Y_2}(h)\big) \circ f\Big) - \mathcal{L}_{X_2}\Big(\big(\mathcal{L}_{Y_1}(h)\big) \circ f\Big)$$
$$= \Big(\mathcal{L}_{Y_1}\big(\mathcal{L}_{Y_2}(h)\big)\Big) \circ f - \Big(\mathcal{L}_{Y_2}\big(\mathcal{L}_{Y_1}(h)\big)\Big) \circ f = \big(\mathcal{L}_{[Y_1, Y_2]}(h)\big) \circ f.$$

But this means $Tf \circ [X_1, X_2] = [Y_1, Y_2] \circ f$. $\qquad \square$

2.2 Submanifolds

Definition 2.13. A subset N of a manifold M is called a submanifold of M if for each $p \in N$ there is a chart (U, φ) in M with $p \in U$ and $n \in \mathbb{N}$ such that

$$\varphi : U \to \mathbb{R}^{n+k} = \mathbb{R}^n \times \mathbb{R}^k \quad \text{and} \quad \varphi(N \cap U) = \varphi(U) \cap (\mathbb{R}^n \times \{0\}).$$

One of the most important tools for dealing with submanifolds is the Constant Rank Theorem. Proofs of the Constant Rank Theorem can be found in the books of Bröcker and Jänich [4] and Abraham et al. [1]. In Section 3.3, we will give a proof of a generalization of this theorem to Banach spaces, but in finite dimensions, it takes the following form:

Theorem 2.14 (Constant Rank Theorem in finite dimensions). *Let $U \subset \mathbb{R}^m$ be an open set in $\mathbb{R}^m$. Let $f : U \to \mathbb{R}^n$ be of class C^r, $r \geq 1$, and assume that $f'(u)$ has constant rank k in a neighbourhood of $u_0 \in \mathbb{R}^m$. Then there exist open sets $U_1, U_2 \subset \mathbb{R}^m$ and $V_1, V_2 \subset \mathbb{R}^n$ and C^r-diffeomorphisms $\varphi : V_1 \to V_2$ and $\psi : U_1 \to U_2$ such that $(\varphi \circ f \circ \psi)(x^1, \ldots, x^m) = (x_1, \ldots, x_k, 0, \ldots, 0)$.*

Definition 2.15. An immersion is a smooth map between manifolds such that its differential is injective at every point. If $i : N \to M$ is an injective immersion, then (N, i) is called an immersed submanifold of M.

Remark 2.16. An immersion has constant (full) rank at every point. So if we look at an immersion in a chart, it is a map as required in the assumptions of Theorem 2.14. Therefore, in the right charts, an immersion looks like the map $(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_m, 0, \ldots, 0)$.

The following definitions and lemmas concerning initial submanifolds are taken from Michor [7].

Definition 2.17. For an arbitrary subset A of a manifold M and $x_0 \in A$, let $C_{x_0}(A)$ denote the set of all $x \in A$ which can be connected to x_0 by a smooth curve in M lying in A.

Definition 2.18. Let N be a subset of a manifold M. A chart (U, u) of M centered at $x \in N$ is called an initial submanifold chart for N if

$$u(C_x(U \cap N)) = u(U) \cap (R^n \times \{0\}) \text{ for some } n \in \mathbb{N}.$$

N is called an initial submanifold of M if for all $x \in N$, there is an initial submanifold chart centered at x.

Definition 2.19. We will call a mapping $i : N \to M$ initial if the following condition holds: For any manifold Z a mapping $f : Z \to N$ is smooth if and only if $i \circ f : Z \to M$ is smooth.

Remark 2.20. Let N be a submanifold of M. Then the injective immersion $i : N \to M$ is an initial mapping, and N is an initial submanifold of M.

Lemma 2.21. *Let $i : N \to M$ be an injective immersion and an initial mapping. Then $i(N)$ is an initial submanifold of M.*

Proof. Let n denote the dimension of N and m the dimension of M. Let $x \in N$. As noticed in Remark 2.16, because i is an immersion, we may choose a chart (V, v) centered at $i(x)$ on M and another chart (W, w) centered at x on N such that

$$(v \circ i \circ w^{-1})(y^1, \ldots, y^n) = (y^1, \ldots, y^n, 0, \ldots, 0).$$

Let $r > 0$ be so small that
$$\{z \in \mathbb{R}^m : |z| < 2r\} \subset v(V) \quad \text{and} \quad \{y \in \mathbb{R}^n : |y| < 2r\} \subset w(W).$$
Put
$$V_1 := v^{-1}(\{z \in \mathbb{R}^m : |z| < 2r\}) \subset N \quad \text{and} \quad W_1 := w^{-1}(\{y \in \mathbb{R}^n : |y| < 2r\}).$$
We claim that (V_1, v) is an initial submanifold chart.
$$v^{-1}(v(V_1) \cap (\mathbb{R}^n \times \{0\})) = v^{-1}(\{(y^1, \ldots, y^n, 0, \ldots, 0) : |y| < r\}) =$$
$$= i \circ w^{-1} \circ (v \circ i \circ w^{-1})^{-1}(\{(y^1, \ldots, y^n, 0, \ldots, 0) : |y| < r\}) =$$
$$= i \circ w^{-1}(\{y \in \mathbb{R}^n : |y| < r\}) = i(W_1) \subset C_{i(x)}(V_1 \cap i(N)).$$

The last inclusion holds since $i(W_1) \subset V_1 \cap i(N)$ and $i(W_1)$ is smoothly contractible. To complete the proof, it remains to show the other inclusion $C_{i(x)}(V_1 \cap i(N)) \subset i(W_1)$. Let $z \in C_{i(x)}(V_1 \cap i(N))$. Then by definition, there is a smooth curve $c : [0,1] \to M$ with $c(0) = i(x)$, $c(1) = z$ that lies entirely in $V_1 \cap i(N)$. Because i is an initial mapping, the unique curve $\tilde{c} : [0,1] \to N$ with $i \circ \tilde{c} = c$ is smooth. We claim that $\tilde{c}$ lies entirely in W_1. If not, then there is some $t \in [0,1]$ such that $\tilde{c}(t) \in w^{-1}(\{y \in \mathbb{R}^n : r \leq |y| < 2r\})$ since $\tilde{c}$ is smooth and thus continuous. But then we have
$$(v \circ c)(t) = (v \circ i)(\tilde{c}(t)) \in (v \circ i \circ w^{-1})(\{y \in \mathbb{R}^n : r \leq |y| < 2r\}) =$$
$$= \{(y, 0) \in (\mathbb{R}^n \times \{0\}) : r \leq |y| < 2r\} \subset \{z \in \mathbb{R}^m : r \leq |z| < 2r\}.$$
But $(v \circ c)(t) \in \{z \in \mathbb{R}^m : r \leq |z| < 2r\}$ means $c(t) \notin V_1$, a contradiction. So $\tilde{c}$ lies entirely in W_1, thus $z = c(1) = i(\tilde{c}(1)) \in i(W_1)$. $\qquad \square$

Lemma 2.22. *Let N be an initial submanifold of M. Then there is a unique smooth manifold structure on N such that the injection $i : N \to M$ is an injective immersion and an initial mapping.*

Proof. Uniqueness follows directly from the definition of an initial mapping. To prove existence, we claim that the charts $(C_x(U_x \cap N), u_x)$, where (U_x, u_x) is an initial submanifold chart for N, yield a smooth manifold structure for N such that i has the desired properties.

The sets $C_x(U_x \cap N)$ might not be open in the subspace topology of N, so in general, the subspace topology of N is not the same as the chart topology. But we know that the chart topology is finer than the subspace topology, and therefore it is Hausdorff. The chart changings are smooth since they are just restrictions of chart changings on M. Clearly, $i : N \to M$ is an injective immersion.

We need to show that i is an initial mapping. For $z \in Z$, we choose a chart (U, u) centered at $f(z)$, such that $u(C_{f(z)}(U \cap N)) = u(U) \cap (\mathbb{R}^n \times \{0\})$. Then $f^{-1}(U)$ is open in Z and contains a chart (V, v) centered at z with $v(V)$ a ball. Then $f(V)$ is smoothly contractible in $U \cap N$, so $f(V) \subset C_{f(z)}(U \cap N)$, and $(u|C_{f(z)}(U \cap N)) \circ f \circ v^{-1}$ is smooth. $\qquad \square$

Remark 2.23. Let N be an initial submanifold of M. Then the connected components of N are separable because M admits a riemannian metric which can be induced on N (see Abraham et al. [1], Section 5.5). However, N may have uncountably many connected components.

3 Chow's Theorem

Definition 3.1. A set of local vectorfields $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ is called involutive if for any $X, Y \in \mathcal{V}$, also $[X, Y] \in \mathcal{V}$. We will call a distribution H involutive if the set of local horizontal vectorfields $\mathfrak{X}_{loc,H}$ is involutive.

Definition 3.2. For a set of local vectorfields $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$, the Lie hull $\mathcal{L}(\mathcal{V})$ of $\mathcal{V}$ is the smallest involutive set of local vectorfields which contains $\mathcal{V}$. $\mathcal{V}$ is called bracket generating if $\mathcal{L}(\mathcal{V})$ spans TM. A smooth distribution H is called bracket generating (satisfies Hörmander's condition) if the set of locally defined horizontal vectorfields $\mathfrak{X}_{loc,H}(M)$ is bracket generating.

Theorem 3.3 (Chow's Theorem). *If H is a bracket generating distribution on a connected manifold, then any two points in the manifold can be connected by a horizontal path.*

The converse of Chow's Theorem fails: It is possible that any two points can be connected by a horizontal path, but that still the distribution is not bracket generating. In Section 3.1 we will see that this is the case exactly if $\mathcal{S}(\mathfrak{X}_{loc,H}(M))$ spans TM, but $\mathcal{L}(\mathfrak{X}_{loc,H}(M))$ does not.

Example 3.4. The following very easy example for this situation can be found in the paper of Nagel et al. [9] and in the book of Jurdjevic [5], page 65. For an other, more elaborate example see Montgomery [8], page 24. Let H be the distribution on $\mathbb{R}^2$ that is spanned by the two vectorfields

$$X_1 = \frac{\partial}{\partial x}, \quad X_2 = \phi \frac{\partial}{\partial y},$$

where ϕ is a C^∞-function on $\mathbb{R}^2$ that vanishes on the left-hand plane $x \leq 0$ and is positive otherwise. Then any two points in the plane can be connected by a horizontal curve. But H is not bracket generating because for $x < 0$, H_x is one-dimensional and therefore all the Lie brackets vanish.

Figure 1 shows that in the above example, one encounters irregularities in the shape of the subriemannian ball $B(\epsilon, x)$ as ϵ increases. The Ball-Box Theorem 3.23 and also Theorem 3.67 show that these irregularities are ruled out in the case of a bracket generating distribution. So non-bracket generating distributions are not desirable from a geometric point of view.
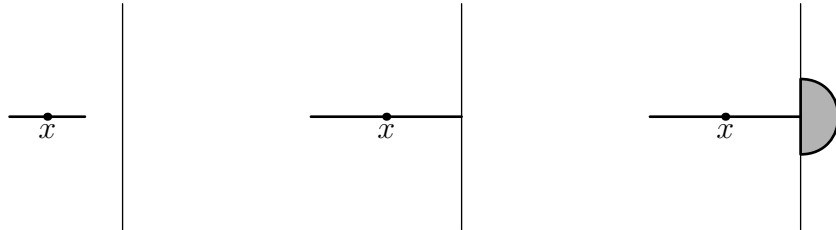

Figure 1: $B(\epsilon, x)$ for $x = (-1, 0) \in \mathbb{R}^2$ and $\epsilon = \frac{1}{2}, \epsilon = 1, \epsilon = \frac{3}{2}$.

We will give three different proofs of Chow's Theorem in Sections 3.1, 3.2 and 3.3, respectively.

3.1 First proof of Chow's theorem

This proof works under the most general assumptions compared to the proofs given in Section 3.2 and 3.3. It is based on the theory of Stefan [10] and Sussmann [11]. However, the notation and the proofs are taken from the lecture notes of Michor [7]. The outline of the proof is the following: The Stefan-Sussmann-theory asserts that the accessible set is a maximal integral manifold of a certain integrable distribution containing H. This will be shown in Theorem 3.20. The bracket generating condition implies that this distribution is equal to TM. Because M is connected, it is the maximal integral manifold of this distribution. Using Theorem 3.20, the proof of Chow's theorem is only a few lines long; it can be found at the very end of this section.

3.1.1 Integral manifolds

Definition 3.5. An integral manifold of a distribution H is a connected immersed submanifold (N, i) such that $T_x i(T_x N) = H_{i(x)}$ for all $x \in N$. An integral manifold of H is called maximal, if it is not contained in any strictly larger integral manifold of H. A distribution H is called integrable if for any $x \in M$, there is an integral manifold of H containing x.

Remark 3.6. Let H be a distribution on a connected manifold M. Then M is a maximal integral manifold of H if and only if $H = TM$.

Lemma 3.7. *Let H be a smooth distribution on M. Then we have:*

(1) If (N, i) is an integral manifold of H and $X \in \mathfrak{X}_{loc,H}(M)$, then i^*X makes sense and is an element of $\mathfrak{X}_{loc}(N)$, which is $i|i^{-1}(U_X)$-related to X, where $U_X \subset M$ is the open domain of X.

(2) If (N_1, i_1) and (N_2, i_2) are integral manifolds of H, then $i_1^{-1}(i_1(N_1) \cap i_2(N_2))$ and $i_2^{-1}(i_1(N_1) \cap i_2(N_2))$ are open subsets in N_1 and N_2, respectively; furthermore $i_2^{-1} \circ i_1$ is a diffeomorphism between them.

(3) If $x \in M$ is contained in some integral manifold of H, then it is contained in a unique maximal one.

Proof.

(1) If $i(x) \in U_X$ for $x \in N$, we have $X(i(x)) \in H_{i(x)} = T_x i(T_x N)$, so $i^*X(x) := ((T_x i)^{-1} \circ X \circ i)(x)$ makes sense. It is defined on an open subset of N and is smooth in x.

(2) Let $X \in \mathfrak{X}_{loc,H}(M)$. Then $i_j^*X \in \mathfrak{X}_{loc}(N_j)$ and is i_j-related to X for $j = 1, 2$. By Lemma 2.9 we have

$$i_j \circ Fl_t^{i_j^*X} = Fl_t^X \circ i_j.$$

Now choose $x_j \in N_j$ such that $i_1(x_1) = i_2(x_2) = x \in M$ and choose vectorfields $X_1, \ldots, X_k \in \mathfrak{X}_{loc,H}$ such that $\{X_1(x), \ldots, X_k(x)\}$ form a basis of H_x. Then

$$f_j(t_1, \ldots, t_k) := (Fl_{t_1}^{i_j^*X_1} \circ \cdots \circ Fl_{t_k}^{i_j^*X_k})(x_j)$$

13

is a smooth mapping defined near zero $\mathbb{R}^k \to N_j$. Since $\frac{\partial}{\partial t_i}\big|_0 f_j = i_j^* X_i$ for $j = 1, 2$, we see that f_j is a diffeomorphism near zero. Finally we have

$$
\begin{aligned}
(i_2^{-1} \circ i_1 \circ f_1)(t_1, \ldots, t_k) &= (i_2^{-1} \circ i_1 \circ Fl_{t_1}^{i_1^* X_1} \circ \cdots \circ Fl_{t_k}^{i_1^* X_k})(x_1) \\
&= (i_2^{-1} \circ Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_k}^{X_k} \circ i_1)(x_1) \\
&= (Fl_{t_1}^{i_2^* X_1} \circ \cdots \circ Fl_{t_k}^{i_2^* X_k} \circ i_2^{-1} \circ i_1)(x_1) \\
&= f_2(t_1, \ldots, t_k).
\end{aligned}
$$

So $i_2^{-1} \circ i_1$ is a diffeomorphism, as required.

(3) Let N be the union of all integral manifolds containing x. Choose the union of all the atlases of these integral manifolds as atlas for N, which is a smooth atlas by (2). Note M admits a riemannian metric which can be induced on N, so N is separable (see Abraham et al. [1], section 5.5). $\qquad\square$

Remark 3.8. If H is an integrable distribution on a manifold M, then by Lemma 3.7 each point is contained in a unique maximal integral manifold. These maximal integral manifolds form a partition of M. This partition is called the foliation of M induced by the integrable distribution H, and each maximal integral manifold is called a leaf of the foliation.

Definition 3.9. A set of local vectorfields $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ is called stable if for all $X, Y \in \mathcal{V}$ and for all t for which it is defined, the local vectorfield $(Fl_t^X)^* Y$ is again an element of $\mathcal{V}$. Let $\mathcal{S}(\mathcal{V})$ denote the minimal stable set of local vectorfields containing $\mathcal{V}$.

Lemma 3.10. *Let $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ be a set of local vectorfields. Then $\mathcal{S}(\mathcal{V})$ consists of all local vectorfields of the form $(Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_k}^{X_k})^* Y$ for $X_i, Y \in \mathcal{V}$.*

Proof. We have to show that the set of vectorfields of the above form is stable. By Lemma 2.9, the flow of such a vectorfield is

$$
Fl_t^{(Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_k}^{X_k})^* Y} = Fl_{-t_k}^{X_k} \circ \ldots Fl_{-t_1}^{X_1} \circ Fl_t^Y \circ Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_k}^{X_k}. \qquad\square
$$

Definition 3.11. A local vectorfield $X \in \mathfrak{X}_{loc}(M)$ is called an infinitesimal automorphism of a distribution H if $T_x(Fl_t^X)(H_x) \subset H_{Fl^X(t,x)}$ whenever defined. Let $\mathrm{aut}(H)$ denote the set of all infinitesimal automorphisms of H.

Lemma 3.12. $\mathrm{aut}(H)$ *is stable.*

Proof. Let $X, Y \in \mathrm{aut}(H)$, and let $Z := (Fl_{t_1}^X)^* Y$. Then

$$
T_x(Fl_t^Z)(H_x) = T_x(Fl_{-t_1}^X \circ Fl_t^Y \circ Fl_{t_1}^X)(H_x) \subset H_{Fl^Z(t,x)}. \qquad\square
$$

Theorem 3.13 (Stefan-Sussmann)**.** *Let H be a smooth distribution on a manifold M. Then the following conditions are equivalent:*

(1) H is integrable.

(2) $\mathfrak{X}_{loc,H}(M)$ *is stable.*

(3) There exists a subset $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ *such that* $\mathcal{S}(\mathcal{V})$ *spans* H.

(4) $\mathrm{aut}(H) \cap \mathfrak{X}_{loc,H}(M)$ *spans* H.

Proof.

$(1) \Rightarrow (2)$ Let $X \in \mathfrak{X}_{loc,H}(M)$ and let L be the leaf through $x \in M$, with $i : L \to M$ the inclusion. Then $Fl^X_{-t} \circ i = i \circ Fl^{i^*X}_{-t}$ by Lemma 2.9, so we have

$$
\begin{aligned}
T_x(Fl^X_{-t})(H_x) = T(Fl^X_{-t}).T_x i.T_x L &= T(Fl^X_{-t} \circ i).T_x L \\
&= Ti.T_x(Fl^{i^*X}_{-t}).T_x L \\
&= Ti.T_{Fl^{i^*X}(-t,x)}L = H_{Fl^X(-t,x)}.
\end{aligned}
$$

This implies that $(Fl^X_t)^*Y \in \mathfrak{X}_{loc,H}(M)$ for any $Y \in \mathfrak{X}_{loc,H}(M)$.

$(2) \Rightarrow (4)$ In fact, (2) says that $\mathfrak{X}_{loc,H} \subset \mathrm{aut}(H)$.

$(4) \Rightarrow (3)$ We can choose $\mathcal{V} = \mathrm{aut}(H) \cap \mathfrak{X}_{loc,H}(M)$: for $X, Y \in \mathcal{V}$, we have $(Fl^X_t)^*Y \in \mathfrak{X}_{loc,H}(M)$; so $\mathcal{V} \subset \mathcal{S}(\mathcal{V}) \subset \mathfrak{X}_{loc,H}(M)$ and H is spanned by $\mathcal{V}$.

$(3) \Rightarrow (1)$ We have to show that each point $x \in M$ is contained in some integral submanifold for the distribution H. Let $\dim H_x = n$. There are $X_1, \dots, X_n \in \mathcal{S}(\mathcal{V})$ such that $X_1(x), \dots, X_n(x)$ span H_x. As in the proof of Lemma 3.7 (2), we consider the mapping

$$
f(t_1, \dots, t_n) := (Fl^{X_1}_{t_1} \circ \cdots \circ Fl^{X_n}_{t_n})(x)
$$

defined near zero in $\mathbb{R}^n$. Since the rank of f at zero is n, the image under f of a small open neighbourhood of zero is a submanifold N of M. We claim that N is an integral manifold of H. The tangent space $T_{f(t_1,\dots,t_n)}N$ is linearly generated by the the tangent vectors

$$
\begin{aligned}
\frac{\partial}{\partial t_k}(Fl^{X_1}_{t_1} &\circ \cdots \circ Fl^{X_n}_{t_n})(x) = \\
&= T(Fl^{X_1}_{t_1} \circ \cdots \circ Fl^{X_{k-1}}_{t_{k-1}})X_k((Fl^{X_{k+1}}_{t_{k+1}} \circ \cdots \circ Fl^{X_n}_{t_n})(x)) \\
&= ((Fl^{X_1}_{-t_1})^* \dots (Fl^{X_{k-1}}_{-t_{k-1}})^*X_k)(f(t_1, \dots, t_n)).
\end{aligned}
$$

Since $\mathcal{S}(\mathcal{V})$ is stable, these vectors lie in $H_{f(t_1,\dots,t_n)}$. They even span $H_{f(t_1,\dots,t_n)}$ because $\dim H_{f(t)} = \dim H_x = n$, as we will prove now. Indeed, since $\mathcal{S}(\mathcal{V})$ spans H and is stable, $T_x(Fl^X_t)$ is a vector space isomorphism from H_x onto $H_{Fl^X(t,x)}$ for all $X \in \mathcal{S}(\mathcal{V})$, and the conclusion follows from the form of f. $\quad\square$

Theorem 3.14. *Let H be an integrable distribution of a manifold M. Then for each $x \in M$ there exists a chart (U, u) around x with $u(U) = \{y \in \mathbb{R}^m : |y^i| < \epsilon$ for all $i\}$ for some $\epsilon > 0$, and a countable subset $A \subset \mathbb{R}^{m-n}$, such that for the leaf L through x we have*

$$
u(U \cap L) = \{y \in u(U) : (y^{n+1}, \dots, y^m) \in A\}.
$$

Each leaf is an initial submanifold. If furthermore the distribution has locally constant rank, the theorem holds with the same n for each leaf L.

Definition 3.15. The chart (u, U) of the above theorem is called a *distinguished chart* for the distribution or the foliation. A connected component of $U \cap L$ is called a *plaque*.

Proof of Theorem 3.14. Let L be the leaf through x, $\dim L = n$. Let $X_1, \ldots, X_n \in \mathfrak{X}_{loc,H}(M)$ be local vectorfields such that $X_1(x), \ldots, X_n(x)$ is a basis of H_x. We choose a chart (V, v) centered at $x \in M$ such that the vectors

$$X_1(x), \ldots, X_n(x), \left.\frac{\partial}{\partial v^{n+1}}\right|_x, \ldots, \left.\frac{\partial}{\partial v^m}\right|_x$$

form a basis of $T_x M$. Then

$$f(t_1, \ldots, t_n) := (Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_n}^{X_n})(v^{-1}(0, \ldots, 0, t^{n+1}, \ldots, t^m))$$

is a diffeomorphism from a neighbourhood of zero in $\mathbb{R}^n$ onto a neighbourhood of x in M. Let (U, u) be the chart given by f^{-1}, suitably restricted. We have

$$y \in L \iff (Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_n}^{X_n})(y) \in L$$

for all $t_1, \ldots, t_n$ and all y for which both expressions make sense. So we have

$$f(t_1, \ldots, t_n) \in L \iff f(0, \ldots, 0, t^{n+1}, \ldots, t^m) \in L,$$

and consequently $L \cap U$ is the disjoint union of connected sets of the form $\{y \in U : (u^{n+1}(y), \ldots, u^m(y)) = \text{ constant }\}$. Since L is a connected immersed submanifold of M, it is second countable and only a countable set of constants can appear in the description of $U \cap L$ given above. We see that $u(C_x(L \cap U)) = u(U) \cap (\mathbb{R}^n \times \{0\})$, so we have found initial submanifold charts for L, and therefore L is an initial submanifold. According to the proof of Lemma 2.22, the charts that make (L, i) an immersed submanifold with the property that $i : L \to M$ is an initial mapping are given by constraining u to $C_x(L \cap U)$. However, in Lemma 3.7, L as an integral manifold has already been given a smooth manifold structure. We will see immediately that the two structures are the same: $u|C_x(L \cap U)$ is the inverse of the map $(t_1, \ldots t_n) \mapsto (Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_n}^{X_n})(x)$, and therefore $u|C_x(L \cap U)$ is just one of the charts given to L in Lemma 3.7.

The arguments given above are valid for any leaf of dimension n meeting U, so also the assertion for an integrable distribution of locally constant rank follows. $\square$

3.1.2 Orbits and integral manifolds

Definition 3.16. A *pseudo-group of local diffeormorphisms* is a set G of local diffeormorphisms such that for each $\varphi, \psi \in G$, we have

$$\varphi \circ \psi \in G \text{ and } \varphi^{-1} \in G,$$

where $\varphi \circ \psi$ and φ^{-1} are as in Definition 2.1. For $\mathcal{V} \subset \mathfrak{X}_{loc,H}(M)$, let $G_{\mathcal{V}}$ denote the smallest pseudo-group of local diffeomorphisms which contains the flows of all vectorfields in $\mathcal{V}$.

Remark 3.17. The elements of $G_\mathcal{V}$ are precisely the maps of the form

$$Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_n}^{X_n}, \text{ with } X_i \in \mathcal{V}.$$

Definition 3.18. Let G be a pseudo-group of local diffeomorphisms on M. Two points $x, y \in M$ are called G-equivalent if there is a $\varphi \in G$ such that $\varphi(x) = y$. This defines an equivalence relation on M. Its equivalence classes are called the orbits of G or G-orbits.

Theorem 3.19. *Let* $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ *and* $x \in M$. *Then the* $G_\mathcal{V}$-*orbit through* x *is the maximal integral manifold through* x *of the distribution spanned by* $\mathcal{S}(\mathcal{V})$.

Proof. The distribution spanned by $\mathcal{S}(\mathcal{V})$ is integrable because of the Stefan-Sussmann Theorem 3.13. Let L be the leaf through x of the distribution spanned by $\mathcal{S}(\mathcal{V})$, and let O be the $G_\mathcal{V}$-orbit containing x.

(1) To show $L \subset O$, we will prove that any two points in L are $G_\mathcal{V}$-equivalent. By the part (3) $\Rightarrow$ (1) of the proof of the Stefan-Sussmann Theorem 3.13 we know that the map
$$f(t_1, \ldots, t_n) := (Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_n}^{X_n})(y),$$
where $X_i \in \mathcal{S}(\mathcal{V})$ and $y \in L$, is a diffeomorphism from an open connected neighbourhood of zero in $\mathbb{R}^n$ onto an open subset U of L. Because of Lemma 3.10 and Lemma 2.9, $Fl_{t_i}^{X_i}$ is the concatenation of flows of vectorfields in $\mathcal{V}$. Therefore by definition any point in U is $G_\mathcal{V}$-equivalent to y. Because y was chosen arbitrarily, every point $y \in L$ has a neighbourhood whose points are all $G_\mathcal{V}$-equivalent. So the $G_\mathcal{V}$-equivalence classes in L are open in L and form a partition of L. Because L is connected, all but one of these equivalence classes are empty, and therefore all points in L are $G_\mathcal{V}$-equivalent.

(2) It remains to show the other inclusion $O \subset L$. Let $y \in O$, i.e.

$$y = Fl_{t_1}^{X_1} \circ \cdots \circ Fl_{t_k}^{X_k}(x) \quad \text{with} \quad X_i \in \mathcal{V} \subset \mathcal{S}(\mathcal{V}).$$

Then $y \in L$ because L is stable under the flows of vectorfields in $\mathcal{S}(\mathcal{V})$. This follows from Lemma 2.9 applied to the pullback i^*X of a vectorfield, where $i : L \to M$ is the inclusion. $\square$

3.1.3 Accessible set and integral manifolds

Theorem 3.20. *Let* H *be a regular distribution on* M *and* $x \in M$. *Then the accessible set* $Acc(x)$ *is the leaf through* x *of the distribution spanned by* $\mathcal{S}(\mathfrak{X}_{loc,H}(M))$.

Proof. Let E be the distribution spanned by $\mathcal{S}(\mathfrak{X}_{loc,H}(M))$. Then E is integrable because the Stefan-Sussmann Theorem 3.13. Let L be the leaf through x of the distribution E.

(1) To show $L \subset Acc(x)$, we have to show that L is horizontally pathwise connected. This is a consequence of Theorem 3.19 with $\mathcal{V} = \mathfrak{X}_{loc,H}(M)$ because any $\mathfrak{X}_{loc,H}(M)$-orbit is of course horizontally connected.

(2) To show the other inclusion $Acc(x) \subset L$, we have to show that a horizontal path starting in L remains in L. In fact, it suffices to show that a path controlled by horizontal vectorfields remains in L. So let $\gamma : [0,1] \to M$ be an absolutely continuous path such that

$$\dot{\gamma}(t) = \sum_{j=1}^{k} u_j(t) \, X_j(\gamma(t))$$

with L^1-functions u_j. We can pull the vectorfields X_j back to L along the inclusion $i : L \to M$ and look at the differential equation

$$\dot{\gamma}(t) = \sum_{j=1}^{k} u_j(t) \, (i^*X_j)(\gamma(t)), \quad \gamma(0) = i^{-1}(x)$$

on L. By restricting the controls to some smaller interval $[0, \epsilon]$, we can make their L^1-norm so small that we can apply Theorem 3.57, which gives us a solution $\tilde{\gamma} : [0, \epsilon] \to L$. By uniqueness, γ and $\tilde{\gamma}$ coincide on $[0, \epsilon]$. From these arguments we see that γ locally remains in L. Therefore the set $\{t \in [0,1] : \gamma(t) \in L\}$ is open and non-empty. But also the set $\{t \in [0,1] : \gamma(t) \notin L\}$ is open for the same reasons. Because $[0,1]$ is connected, one of these sets must be empty, and therefore γ remains in L. $\square$

3.1.4 Proof of Chow's Theorem

Lemma 3.21. *An integrable distribution is involutive.*

Proof. Let H be an integrable distribution; we have to show that $\mathfrak{X}_{loc,H}$ is involutive. Take vectorfields $X, Y \in \mathfrak{X}_{loc,H}$ defined near $x \in M$, and let (N, i) be an integral manifold of H near x. By Lemma 3.7.1, i^*X and i^*Y are local vectorfields on N, and they are i-related to X and Y, respectively. By Lemma 2.12, also $[i^*X, i^*Y]$ is i-related to $[X, Y]$. Therefore $[X, Y]$ is horizontal. $\square$

Theorem 3.22 (Chow's Theorem)**.** *If H is a bracket generating distribution on a connected manifold, then any two points in the manifold can be connected by a horizontal path.*

Proof. Let E denote the integrable distribution spanned by $\mathcal{S}(\mathfrak{X}_{loc,H}(M))$, and let $x \in M$. By the bracket generating condition and Lemma 3.21, $E = TM$. M is connected, so M is the leaf through x of E. By Theorem 3.20 $Acc(x)$ also is the leaf through x of E, so $Acc(x) = M$. $\square$

3.2 Second proof of Chow's Theorem

This proof works only for regular distributions. The idea of the proof comes from Montgomery [8]. He derives Chow's theorem from the Ball-Box Theorem, which is interesting in itself.

Theorem 3.23 (Ball-box Theorem)**.** *Let H be a regular bracket generating distribution on an n-dimensional manifold M, and let $x_0 \in M$. Then there exist coordinates $y_1, \ldots, y_n$ centered at x_0 and positive constants $c < C$ and $\epsilon_0 > 0$ such that for all $\epsilon < \epsilon_0$,*

$$\mathrm{Box}^w(c\epsilon) \subset B(\epsilon, x_0) \subset \mathrm{Box}^w(C\epsilon).$$

Here, w is some vector in $\mathbb{N}^n$ called the weighting associated to the growth vector at x_0 and $\mathrm{Box}^w(\epsilon)$ is the w-weighted box of size ϵ given in coordinates by the expression

$$\mathrm{Box}^w(\epsilon) = \left\{ y \in \mathbb{R} : |y_i| \le \epsilon^{w_i}, i = 1, \ldots, n \right\}.$$

To prove Chow's Theorem it is sufficient to show that the subriemannian ball $B(\epsilon, x_0)$ is a neighbourhood of x_0, and this is all we will prove here. The reason is that we were confronted with technical difficulties in the proof given in the book of Montgomery [8]. We refer to Nagel et al. [9] for the full proof and a much longer, but very nice and complete treatment of the subject.

The outline of the proof presented in this work is the following: Locally, some horizontal vectorfields $X_1, \ldots, X_k$ together with a selection of their brackets span the tangent space. To these vectorfields correspond the flows $Fl_t^{X_1}, \ldots, Fl_t^{X_k}$ and a selection of brackets of these flows. It will take some time to calculate the first non-vanishing derivative at $t = 0$ of the brackets of flows, but this is done in Theorem 3.31: The first non-vanishing derivative at $t = 0$ of a bracket of $Fl_t^{X_1}, \ldots, Fl_t^{X_k}$ is the same bracket applied to $X_1, \ldots, X_k$. In Theorem 3.38 this result will allow us to construct a C^1-diffeomorphism ψ from the concatenation of the above flows and their brackets. Finally the openness of $B(\epsilon, x_0)$ follows from the openness of ψ.

Definition 3.24. For sets of local vectorfields $\mathcal{V}, \mathcal{W} \subset \mathfrak{X}_{loc}(M)$, we define

$$[\mathcal{V}, \mathcal{W}] = \{ [X, Y] : X \in \mathcal{V}, Y \in \mathcal{W} \} \quad \text{and}$$
$$\mathcal{V} + \mathcal{W} = \{ X + Y : X \in \mathcal{V}, Y \in \mathcal{W} \}.$$

For a set of local vectorfields $\mathcal{V}$, we inductively define the subsets

$$\mathcal{V}_1 = \mathcal{V}, \quad \mathcal{V}_{k+1} = \mathcal{V}_k + [\mathcal{V}, \mathcal{V}_k].$$

Definition 3.25. Let H be a smooth distribution. Then the distribution spanned by $\mathfrak{X}_{loc,H}(M)_k$ will be denoted by H_k.

Definition 3.26. Let $C^\infty(M)$ denote the set of smooth functions from M to $\mathbb{R}$. For a set of locally defined vectorfields $\mathcal{V}$, let

$$\mathrm{span}_{C^\infty(M)} \mathcal{V} = \left\{ \sum u_i X_i : u_i \in C^\infty(M), X_i \in \mathcal{V} \right\},$$

where the number of summands in the above sum has to be finite.

Lemma 3.27. *Let $\mathcal{V}$ be a set of locally defined vectorfields. Then*

$$(\mathrm{span}_{C^\infty(M)} \mathcal{V})_k = \mathrm{span}_{C^\infty(M)}(\mathcal{V}_k).$$

Particularly, if $\mathcal{V}$ spans H, then $\mathcal{V}_k$ spans H_k.

Proof. Let $\mathcal{W} = \mathrm{span}_{C^\infty(M)}(\mathcal{V})$.

(1) $\mathrm{span}_{C^\infty(M)}(\mathcal{V}_k) \subset \mathcal{W}_k$: Clearly, $\mathcal{V} \subset \mathcal{W}$, so we have $\mathcal{V}_k \subset \mathcal{W}_k$, and

$$\mathrm{span}_{C^\infty(M)}(\mathcal{V}_k) \subset \mathrm{span}_{C^\infty(M)}(\mathcal{W}_k) = \mathcal{W}_k.$$

The last equality holds because $\mathcal{W}_k$ is a $C^\infty(M)$-module, as will be shown right now. Indeed we claim that if $\mathcal{W}$ is a $C^\infty(M)$-module, then $\mathcal{W}_k$ is a $C^\infty(M)$-module, too. This is trivial for $k = 1$. For $k \geq 1$ and $X \in \mathcal{W}_{k+1}$, we can write

$$X = Y + [Z, W] \in \mathcal{W}_k + [\mathcal{W}, \mathcal{W}_k].$$

Let $f \in C^\infty(M)$. Then

$$fX = fY + (Wf)Z + [fZ, W] \in \mathcal{W}_k + [\mathcal{W}, \mathcal{W}_k] = \mathcal{W}_{k+1},$$

and this proves our claim.

(2) $\mathcal{W}_k \subset \mathrm{span}_{C^\infty(M)}(\mathcal{V}_k)$: This can be proven by induction. For $k = 1$, the statement is trivial. For the inductive step, let $X \in \mathcal{W}_{k+1} = \mathcal{W}_k + [\mathcal{W}, \mathcal{W}_k]$. By the inductive assumption, $\mathcal{W}_k \subset \mathrm{span}_{C^\infty(M)}(\mathcal{V}_k)$. Therefore, with $X_i \in \mathcal{V}_k$, $Y_i \in \mathcal{V}, Z_j \in \mathcal{V}_k$, we can write

$$\begin{aligned}
X &= \sum u_i X_i + \left[\sum v_i Y_i, \sum w_j Z_j\right] \\
&= \sum u_i X_i + \sum \sum v_i w_j [Y_i, Z_j] + v_i (Y_i w_j) Z_j - w_j (Z_j v_i) Y_i \\
&\in \mathrm{span}_{C^\infty(M)}(\mathcal{V}_{k+1}).
\end{aligned}$$
$\square$

Lemma 3.28. *If H is a bracket generating distribution, then locally, there is an $r \in \mathbb{N}$ such that $H_r = TM$.*

Proof. Let n be the dimension of M. For $x \in M$, we find vectorfields $X_1, \ldots, X_n \in \mathcal{L}(\mathfrak{X}_{loc,H}(M))$ defined near x such that $X_1(x), \ldots, X_n(x)$ form a basis of $T_x M$. Because $X_1, \ldots, X_n$ are linearly independent at x, they are also linearly independent in some neighbourhood of x. Furthermore, because

$$\mathcal{L}(\mathfrak{X}_{loc,H}(M)) = \sum_{k \in \mathbb{N}} \mathfrak{X}_{loc,H}(M)_k,$$

every X_i lies in some $\mathfrak{X}_{loc,H}(M)_{k_i}$. Therefore, they all lie in $\mathfrak{X}_{loc,H}(M)_r$, where $r = \max\{k_i : i = 1, \ldots, n\}$. Consequently, $H_r = TM$ near x. $\square$

Definition 3.29. For a bracket generating distribution such that $H_r = TM$, we define the growth vector $n(x) \in \mathbb{R}^r$ at $x \in M$ to be the vector whose components $n_i(x)$ are the rank of the distribution H_i at x. (Obviously $n(x)$ may vary from point to point.)

The following theorem and its proof is taken from a paper of Mauhart and Michor [6]; it describes the correspondence between brackets of flows and brackets of vectorfields.

Definition 3.30. For curves of local diffeomorphisms φ and ψ, we define the bracket of local diffeomorphisms

$$[\varphi, \psi](t, x) = \left(\psi_t^{-1} \circ \varphi_t^{-1} \circ \psi_t \circ \phi_t\right)(x)$$

for all (t, x) where the equation is well-defined. As for Lie brackets of vectorfields, we extend this definition in the obvious way to arbitrary formal bracket expressions.

Theorem 3.31. *For $i = 1 \ldots k$, let φ^i be curves of local diffeomorphisms through the identity. Let X_i be their infinitesimal generators. Then for each formal bracket expression B of length k we have*

$$0 = \left.\frac{\partial^l}{\partial t^l}\right|_0 B(\varphi_t^1, \ldots, \varphi_t^k) \quad \text{for } 1 \leq l < k,$$

$$B(X_1, \ldots, X_k) = \frac{1}{k!} \left.\frac{\partial^k}{\partial t^k}\right|_0 B(\varphi_t^1, \ldots, \varphi_t^k).$$

In the above theorem, $B(X_1, \ldots, X_k)$ clearly is a vectorfield, but what about the k-th derivative of $B(\varphi_t^1, \ldots, \varphi_t^k)$? To explain this, we need the following lemma:

Lemma 3.32. *Let $c : \mathbb{R} \to M$ be a smooth curve. If $c(0) = x \in M$, $c'(0) = 0, \ldots, c^{(k-1)}(0) = 0$, then $c^{(k)}(0)$ is a well-defined tangent vector in $T_x M$ which is given by the derivation $f \mapsto (f \circ c)^{(k)}(0)$.*

Proof. We have to show that the Leibnitz rule holds for the map $f \mapsto (f \circ c)^{(k)}(0)$.

$$((f \cdot g) \circ c)^{(k)}(0) = ((f \circ c) \cdot (g \circ c))^{(k)} =$$

$$= \sum_{j=0}^{k} \binom{k}{j} (f \circ c)^{(j)}(0) \cdot (g \circ c)^{(k-j)}(0) =$$

$$= (f \circ c)^{(k)}(0)\, g(q) + f(q)\, (g \circ c)^{(k)}(0)$$

since all other summands vanish: $(f \circ c)^{(j)}(0) = 0$ for $j = 1, \ldots, k-1$. $\qquad\square$

Lemmas 3.33, 3.34 and 3.35 constitute the proof of Theorem 3.31. In the following, we will use the notation $\partial_t^k = \frac{\partial^k}{\partial t^k}$.

Lemma 3.33. *Let φ, ψ be curves of local diffeomorphisms through Id_M and let $f \in C^\infty(M)$. Then we have*

$$\left.\partial_t^k\right|_0 (\varphi_t \circ \psi_t)^* f = \sum_{j=0}^{k} \binom{k}{j} (\left.\partial_t^j\right|_0 \psi_t^*)(\left.\partial_t^{k-j}\right|_0 \varphi_t^*) f.$$

For l curves of local diffeormphisms $\varphi^1, \ldots \varphi^l$ through Id_M the multinomial version of this formula also holds:

$$\left.\partial_t^k\right|_0 (\varphi_t^1 \circ \cdots \circ \varphi_t^l)^* f = \sum_{j_1 + \cdots + j_l = k} \frac{k!}{j_1! \ldots j_l!} (\left.\partial_t^{j_l}\right|_0 (\varphi_t^l)^*) \ldots (\left.\partial_t^{j_1}\right|_0 (\varphi_t^1)^*) f.$$

Proof. We will only show the binomial version. For a function $h(t, s)$ of two variables we have

$$\partial_t^k h(t, t) = \sum_{j=0}^{k} \binom{k}{j} \partial_t^j \partial_s^{k-j} h(t, s) \Big|_{t=s}$$

since for $h(t, s) = f(t) g(s)$ this is just a consequence of the Leibnitz rule, and linear combinations of such decomposable tensors are dense in the space of all functions of two variables in the compact C^∞-topology, so that by continuity the formula holds for all functions. In the following form it implies the lemma:

$$\partial_t^k\big|_0 f(\varphi(t, \psi(t, x))) = \sum_{j=0}^{k} \binom{k}{j} \partial_s^{k-j} \partial_t^j f(\varphi(t, \psi(s, x)))\big|_{t=s=0} . \qquad \square$$

Lemma 3.34. *Let φ be a curve of local diffeormphisms through Id_M with first non-vanishing derivative*

$$k! \, X = \partial_t^k\big|_0 \varphi_t.$$

Then the inverse curve of local diffeomorphims φ^{-1} has first non-vanishing derivative

$$-k! \, X = \partial_t^k\big|_0 \varphi_t^{-1}.$$

Proof. We have $\varphi_t^{-1} \circ \varphi_t = \mathrm{Id}_M$, so by Lemma 3.33 we get for $1 \le j \le k$

$$\begin{aligned}
0 &= \partial_t^j\big|_0 (\varphi_t^{-1} \circ \varphi_t)^* f = \sum_{i=0}^{j} \binom{i}{j} (\partial_t^i\big|_0 (\varphi_t)^*)(\partial_t^{j-i}\big|_0 (\varphi_t^{-1})^*) f \\
&= \partial_t^j\big|_0 (\varphi_t)^* (\varphi_0^{-1})^* f + (\varphi_0)^* \, \partial_t^j\big|_0 (\varphi_t^{-1})^* f \\
&= \partial_t^j\big|_0 (\varphi_t)^* f + \partial_t^j\big|_0 (\varphi_t^{-1})^* f
\end{aligned}$$

as required. $\qquad \square$

Lemma 3.35. *Let φ be a curve of local diffeormphisms through Id_M with first non-vanishing derivative $m! \, X = \partial_t^m\big|_0 \varphi_t$, and let ψ be a curve of local diffeormphisms through Id_M with first non-vanishing derivative $n! \, Y = \partial_t^n\big|_0 \psi_t$. Then the curve of local diffeomorphisms $[\varphi, \psi]$ has first non-vanishing derivative*

$$(m + n)! \, [X, Y] = \partial_t^{m+n}\big|_0 [\varphi, \psi]_t.$$

Proof. By the multinomial version of Lemma 3.33 we have

$$\begin{aligned}
A_N f &:= \partial_t^N\big|_0 (\psi_t^{-1} \circ \varphi_t^{-1} \circ \psi_t \circ \varphi_t)^* f \\
&= \sum_{i+j+k+l=N} \frac{N!}{i!j!k!l!} (\partial_t^i\big|_0 (\varphi_t^{-1})^*)(\partial_t^j\big|_0 (\psi_t^{-1})^*)(\partial_t^k\big|_0 (\varphi_t)^*)(\partial_t^l\big|_0 (\psi_t)^*) f.
\end{aligned}$$

Let us suppose that $1 \le n \le m$, the case $m \le n$ is similar. If $N < n$ all summands vanish. If $N = n$ we have by Lemma 3.34

$$A_N f = (\partial_t^n\big|_0 (\varphi_t)^*) f + (\partial_t^n\big|_0 (\psi_t)^*) f + (\partial_t^k\big|_0 (\varphi_t^{-1})^*) f + (\partial_t^k\big|_0 (\psi_t^{-1})^*) f = 0.$$

If $n < N \le m$ we have, using again Lemma 3.34

$$
\begin{aligned}
A_N f &= \sum_{j+l=N} \frac{N!}{j!l!} (\partial_t^j\big|_0 (\psi_t)^*)(\partial_t^k\big|_0 (\psi_t^{-1})^*) f + \delta_N^m (\partial_t^m\big|_0 (\varphi_t)^*) f + \partial_t^m\big|_0 (\varphi_t^{-1})^*) f) \\
&= (\partial_t^N\big|_0 (\psi_t^{-1} \circ \psi_t)^*) f + 0 = 0.
\end{aligned}
$$

Now we come to the difficult case $m, n < N \le m+n$. By Lemma 3.33 and since all other terms vanish we get

$$
\begin{aligned}
A_N f = \;& \partial_t^N\big|_0 (\psi_t^{-1} \circ \varphi_t^{-1} \circ \psi_t)^* f \\
&+ \binom{N}{m} (\partial_t^m\big|_0 \varphi_t^*)(\partial_t^{N-m}\big|_0 (\psi_t^{-1} \circ \varphi_t^{-1} \circ \psi_t)^*) f + (\partial_t^N\big|_0 \varphi_t^*) f.
\end{aligned}
$$

By Lemma 3.33 again we get

$$
\begin{aligned}
\partial_t^N\big|_0 (\psi_t^{-1} \circ \varphi_t^{-1} \circ \psi_t)^* f =& \\
= \;& \sum_{j+k+l=N} \frac{N!}{j!k!l!} (\partial_t^j\big|_0 (\psi_t)^*)(\partial_t^k\big|_0 (\varphi_t^{-1})^*)(\partial_t^l\big|_0 (\psi_t^{-1})^*) f \\
= \;& \sum_{j+l=N} \binom{N}{j} (\partial_t^j\big|_0 (\psi_t)^*)(\partial_t^l\big|_0 (\psi_t^{-1})^*) f + \binom{N}{m} (\partial_t^{N-m}\big|_0 (\psi_t)^*)(\partial_t^m\big|_0 (\varphi_t^{-1})^*) f+ \\
&+ \binom{N}{m} (\partial_t^m\big|_0 (\varphi_t^{-1})^*)(\partial_t^m\big|_0 (\psi_t^{-1})^*) f + \partial_t^N\big|_0 (\varphi_t^{-1})^*) f \\
= \;& \binom{N}{m} (\partial_t^{N-m}\big|_0 (\psi_t)^*) m! \, (-Xf) + \binom{N}{m} (\partial_t^m\big|_0 (\varphi_t^{-1})^*) n! \, (-Yf) + \partial_t^N\big|_0 (\varphi_t^{-1})^* f \\
= \;& \delta_{m+n}^N (m+n)! \, (XY - YX) f + \partial_t^N\big|_0 (\varphi_t^{-1})^* f \\
= \;& \delta_{m+n}^N (m+n)! \, [X,Y] f + \partial_t^N\big|_0 (\varphi_t^{-1})^* f.
\end{aligned}
$$

From the second expression above one can also read off that

$$
\partial_t^{N-m}\big|_0 (\psi_t^{-1} \circ \varphi_t^{-1} \circ \psi_t)^* f = \partial_t^{N-m}\big|_0 ((\varphi_t^{-1})^* f.
$$

If we put the last conclusions together we get, using Lemma 3.33 and Lemma 3.34 again:

$$
\begin{aligned}
A_N f &= \delta_{m+n}^N (m+n)! \, [X,Y] f + \partial_t^N\big|_0 (\varphi_t^{-1})^* f \\
&\quad + \binom{N}{m} (\partial_t^m\big|_0 \varphi_t^*)(\partial_t^{N-m}\big|_0 (\varphi_t^{-1})^*) f + (\partial_t^N\big|_0 \varphi_t^*) f \\
&= \delta_{m+n}^N (m+n)! \, [X,Y] f + (\partial_t^N\big|_0 (\varphi_t^{-1} \circ \varphi_t)^*) f \\
&= \delta_{m+n}^N (m+n)! \, [X,Y] f. \qquad \square
\end{aligned}
$$

We will use the following two lemmas in our proof of Chow's Theorem.

Lemma 3.36. *Let $\varphi : (-\epsilon, \epsilon) \to M$ be a smooth curve whose derivatives at zero of order up to and including $n-1$ are zero, i.e. $\varphi'(0) = 0, \dots, \varphi^{(n-1)}(0) = 0$. Then $t \mapsto \varphi(\sqrt[n]{t})$ is a C^1-function on $[0, \epsilon)$ whose first derivative at zero equals $\frac{1}{n!} \varphi^{(n)}(0)$. If n is odd, it is a C^1-function on $(-\epsilon, \epsilon)$.*

Proof. This is a local question, so we can work in a chart and assume without loss of generality that $M = \mathbb{R}^n$ and $\varphi(0) = 0$. First, we will look at $t \geq 0$ only and show differentiability from the right at $t = 0$ using the Taylor formula $\varphi(t) = \frac{t^n}{n!}\varphi^{(n)}(0) + o(t^n)$.

$$\lim_{t \searrow 0} \frac{1}{t}\left(\varphi(\sqrt[n]{t}) - \varphi(0)\right) = \lim_{t \searrow 0} \frac{1}{t}\left(\frac{t}{n!}\varphi^{(n)}(0) + o(t)\right) = \frac{\varphi^{(n)}(0)}{n!}.$$

Now, we show that the derivative is continuous from the right at $t = 0$. We use the Taylor formula for φ' and get $\varphi'(t) = \varphi^{(n)}(0)\frac{t^{n-1}}{(n-1)!} + o(t^{n-1})$, and therefore

$$
\begin{aligned}
\lim_{t \searrow 0} \frac{d}{dt}\varphi(\sqrt[n]{t}) &= \lim_{t \searrow 0} \varphi'(\sqrt[n]{t})\frac{t^{(1/n)-1}}{n} = \\
&= \lim_{t \searrow 0} \varphi^{(n)}(0)\frac{\sqrt[n]{t}^{n-1}}{(n-1)!}\frac{t^{(1/n)-1}}{n} + o(\sqrt[n]{t}^{n-1})\frac{t^{(1/n)-1}}{n} = \frac{\varphi^{(n)}(0)}{n!}.
\end{aligned}
$$

If n is odd, all the above calculations can be done for $t < 0$ as well. $\qquad\square$

Lemma 3.37. *Let $\varphi_1, \varphi_2 : (-\epsilon, \epsilon) \to M$ be smooth functions whose derivatives at zero of order up to and including $n - 1$ are zero and $\varphi_1^{(n)}(0) = -\varphi_2^{(n)}(0)$. Then the function*

$$\psi : (-\epsilon, \epsilon) \to M, \quad \psi(t) = \begin{cases} \varphi_1(\sqrt[n]{t}) & \text{for } t \geq 0, \\ \varphi_2(-\sqrt[n]{-t}) & \text{for } t < 0 \end{cases}$$

is a C^1 function.

Proof. Again, this is a local question, so we can work in a chart and assume without loss of generality that $M = \mathbb{R}^n$ and $\varphi_1(0) = \varphi_2(0) = 0$. We have to show that the left and right derivatives of ψ at $t = 0$ are equal. According to Lemma 3.36, the right derivative of ψ at 0 is $\frac{1}{n!}\varphi_1^{(n)}(0)$. To calculate the left derivative, we look at the function $t \mapsto \varphi_2(-t)$. Its derivatives at zero of order up to and including $n - 1$ are zero, and its n-th derivative at zero is $(-1)^n\varphi_2^{(n)}(0)$. So we can apply Lemma 3.36 to it and calculate

$$
\begin{aligned}
\lim_{t \nearrow 0} \frac{1}{t}\left(\psi(t) - \psi(0)\right) &= \lim_{t \nearrow 0} \frac{1}{t}\varphi_2(-\sqrt[n]{-t}) = -\lim_{t \searrow 0} \frac{1}{t}\varphi_2(-\sqrt[n]{t}) = \\
&= -\frac{1}{n!}(-1)^n\varphi_2^{(n)}(0) = \frac{1}{n!}(-1)^n\varphi_1^{(n)}(0) = \frac{1}{n!}\varphi_1^{(n)}(0).
\end{aligned}
$$

$\qquad\square$

Theorem 3.38. *If H is a regular bracket generating distribution on a manifold M, then for all $\epsilon > 0$, the subriemannian ball $B(\epsilon, x_0)$ is a neighbourhood of x_0.*

Proof. (1) Let n be the dimension of M. We may work locally, so Lemma 3.28 asserts that there exists an $r \in \mathbb{N}$ such that $H_r = TM$. Let $(n_1, \ldots, n_r)$ be the growth vector at x_0. Because H is regular, we can take orthonormal horizontal vectorfields $X_1, \ldots, X_{n_1}$ defined near x_0 spanning H near x_0. $\mathcal{V} = \{X_1, \ldots, X_{n_1}\}$ spans H, so Lemma 3.27 asserts that $\mathcal{V}_j$ spans H_j for $j = 1, \ldots, r$. So there are

bracket expressions $B_1, \ldots, B_n$ such that the vectorfields $X_i = B_i(X_1, \ldots, X_{n_1})$ span TM, and such that for $1 \le i \le n_1$, B_i is simply a vectorfield, for $n_1 < i \le n_2$, B_i consists of one bracket etc. For $i = 1, \ldots, n$, let w_i denote the number of brackets in B_i increased by one. w_i is called the weighting of the vectorfield X_i.

(2) Let $\varphi_1, \ldots, \varphi_{n_1}$ be the local flows of the vectorfields $X_1, \ldots, X_{n_1}$. So $\varphi_1, \ldots, \varphi_{n_1}$ are curves of local diffeormorphisms through the identity. Now, for $i = 1, \ldots, n$, let $\varphi_i = B_i(\varphi_1, \ldots, \varphi_{n_1})$. Note that this does not change the definition of $\varphi_1, \ldots, \varphi_{n_1}$ because $B_1, \ldots, B_{n_1}$ are bracket expressions of length zero. According to Theorem 3.31, in any coordinate system centered at x we have $\varphi_i(t, x) = t^{w_i} X_i(x) + o(t^{w_i})$. If w_i is even, t^{w_i} is never negative, so we can not move in the negative X_i direction. This motivates the approach of adapting the commutators of flows.

(3) For $i = 1, \ldots, n$, we build the formal bracket expression C_i by swapping the arguments of the outermost bracket in B_i. So, if E_i and F_i are bracket expressions such that $B_i = [E_i, F_i]$, then $C_i = [F_i, E_i]$. Then we define

$$\chi_i(t, x) = \begin{cases} B_i(\varphi_1, \ldots, \varphi_{n_1})(t, x) & \text{for } w_i \text{ even, } t \ge 0, \\ C_i(\varphi_1, \ldots, \varphi_{n_1})(t, x) & \text{for } w_i \text{ even, } t \ge 0, \\ B_i(\varphi_1, \ldots, \varphi_{n_1})(t, x) & \text{for } w_i \text{ odd.} \end{cases}$$

(4) For $i = 1, \ldots, n$, we write

$$\sigma_i : \mathbb{R} \to \mathbb{R}, \quad \sigma_i(t) := \begin{cases} t^{w_i} & \text{for } w_i \text{ even, } t \ge 0, \\ -t^{w_i} & \text{for } w_i \text{ even, } t < 0, \\ t^{w_i} & \text{for } w_i \text{ odd.} \end{cases}$$

Then σ_i is a non-smooth homeomorphism. Let s_i be its inverse, i.e.

$$s_i : \mathbb{R} \to \mathbb{R}, \quad s_i(t) = \begin{cases} \sqrt[w_i]{t} & \text{for } w_i \text{ even, } t \ge 0, \\ -\sqrt[w_i]{-t} & \text{for } w_i \text{ even, } t < 0, \\ \sqrt[w_i]{t} & \text{for } w_i \text{ odd.} \end{cases}$$

Let $\psi_i(t, x) := \chi_i(s_i(t), x)$ and let

$$\begin{aligned} \chi(t_1, \ldots, t_n) &:= \chi_1(t_1) \circ \cdots \circ \chi_n(t_n)(x_0) \\ \psi(t_1, \ldots, t_n) &:= \psi_n(t_n) \circ \cdots \circ \psi_1(t_1)(x_0). \end{aligned}$$

According to Lemmas 3.36 and 3.37, the maps $\psi_1, \ldots, \psi_n$ are C^1 in t for each fixed x. For each fixed t, they are C^∞ in x. Therefore ψ is a C^1-map defined on a neighbourhood of $0 \in \mathbb{R}^n$. We calculate its partial derivatives at 0:

$$\left. \frac{\partial}{\partial t_i} \right|_0 \psi = \frac{X_i(x_0)}{w_i!} \quad \text{because} \quad \left. \frac{\partial}{\partial x} \right|_{(0,0)} \psi_i = \mathrm{Id} .$$

$X_1(x_0), \ldots, X_n(x_0)$ are linearly independent tangent vectors, so we may apply the inverse function theorem and invert ψ in a neighbourhood of 0. The inverse of ψ is C^1, so it is continuous, and therefore ψ is an open mapping.

(5) The map χ is a composition of $\chi_1(t_1), \ldots, \chi_n(t_n)$, and each map χ_i is a bracket of the flows $\varphi_1, \ldots, \varphi_{n_1}$ which correspond to the horizontal vectorfields $X_1, \ldots, X_{n_1}$. If we expand the map χ in terms of the flows $\varphi_1, \ldots, \varphi_{n_1}$, let K denote the number of flows that appear in this expansion. Now, we claim that $\chi(\mathrm{Box}(\epsilon/K))$ is contained in the subriemannian ball $B(\epsilon, x_0)$, where

$$\mathrm{Box}(\epsilon) = \{y \in \mathbb{R}^n : |y_i| \leq \epsilon, i = 1, \ldots, n\}.$$

To prove the claim, we note that for $1 \leq i \leq n_1$ and any point x in the domain of definition of $\varphi_i(t_i)$, the subriemannian distance $d(x, \varphi_i(t_i)(x))$ is less than ϵ/K if $|t_i| < \epsilon/K$. The reason is that x and $\varphi_i(t_i)(x)$ can be connected by the horizontal path $t \mapsto \varphi_i(t, x), 0 \leq t \leq t_i$. The length of this path is $|t_i|$ because φ_i is the flow of the vectorfield X_i, and $X_1, \ldots, X_{n_1}$ were taken to be orthonormal. Now the claim follows because x_0 and $\chi(t_1, \ldots, t_n)$ can be connected by a concatenation of K horizontal paths, each of length less than ϵ/K.

(6) Define the weighted box $Box^w(\epsilon) = \{y \in \mathbb{R} : |y_i| \leq \epsilon^{w_i}, i = 1, \ldots, n\}$ as in the Ball-Box Theorem 3.23. $Box^w(\epsilon/K)$ is a neighbourhood of $0 \in \mathbb{R}^n$ and ψ is an open mapping, therefore $\psi(Box^w(\epsilon/K))$ is a neighbourhood of x_0. Because $B(\epsilon, x_0) \supset \chi(\mathrm{Box}(\epsilon/K) = \psi(Box^w(\epsilon/K))$, we have shown that $B(\epsilon, x_0)$ is a neighbourhood of x_0 as well. $\square$

Lemma 3.39. *If M is connected and every point has a horizontally pathwise connected neighbourhood, then M is horizontally pathwise connected.*

Proof. Any connected and locally pathwise connected topological space is pathwise connected. Here, we only allow horizontal paths, but the proof remains the same: Take an arbitrary point $x_0 \in M$. Then $Acc(x_0)$ is open. Let $Acc(x_0)^C$ denote the set of all points not lying in $Acc(x_0)$. $Acc(x_0)^C$ is open because it is the union $\bigcup \{Acc(x) : x \notin Acc(x_0)\}$. Because M is connected, either $Acc(x_0)$ or $Acc(x_0)^C$ is empty. Because $x_0 \in Acc(x_0)$, $Acc(x_0)^C$ must be empty, and $Acc(x_0) = M$. $\square$

Theorem 3.40 (Chow's Theorem). *If H is a regular bracket generating distribution on a connected manifold, then any two points in the manifold can be connected by a horizontal path.*

Proof. Theorem 3.38 asserts that every point x in the manifold has a horizontally pathwise connected neighbourhood, namely $B(\epsilon, x)$, so we can apply Lemma 3.39. $\square$

3.3 Third proof of Chow's Theorem

For this proof we need the assumption that H is locally finitely generated. The outline of the proof is the following: By Lemma 3.39 it is sufficient to prove that every point has a horizontally pathwise connected neighbourhood, so we can work locally and assume that $X_1, \ldots, X_k$ are defined everywhere and that M is an open

connected subset of $\mathbb{R}^n$. We will see in Theorem 3.57 that under suitable conditions, the differential equation

$$\dot{\gamma}(t) = \sum_{i=1}^{k} u_i(t) X_i(\gamma(t)), \quad \gamma(0) = x$$

has a unique solution γ, and γ depends smoothly on u and x. Then also the endpoint map, that is the map assigning to u the endpoint of γ is smooth. Using some calculus in Banach spaces we will show in Theorem 3.65 that the image of the endpoint map is an initial submanifold of M. If H is bracket generating, this initial submanifold has the same dimension as M, and therefore it is an open subset of M. In fact, it is all of M, and this will conlude the proof of Chow's Theorem.

Remark 3.41. The most important part in the proof is the fact that the endpoint map is a smooth map from a Banach space to M. We can ask if the set of horizontal curves can be endowed with the structure of a Banach manifold such that the endpoint map is a smooth map from this Banach manifold to M. The idea is to expand the derivative of γ in terms of a local horizontal frame to get coordinates for γ. For regular distributions, this is done in Montgomery [8], Appendix D.

3.3.1 Calculus in Banach spaces

A good treatment of calculus in Banach spaces, including the following theorems and definitions, can be found in the book of Abraham et al. [1].

Definition 3.42. If $X_1, \ldots, X_k$ and Y are Banach spaces, let $L(X_1, \ldots, X_k; Y)$ denote the Banach space of k-multilinear bounded operators of $X_1, \ldots, X_k$ to Y. If $X_i = X, 1 \leq i \leq k$, this space is denoted by $L^k(X, Y)$.

Remark 3.43. There are (natural) norm-preserving isomorphisms

$$L(X_1, L(X_2, \ldots, X_k)) \cong L(X_1, \ldots, X_k; Y) \cong L(X_1, \ldots, X_{k-1}; L(X_k; Y))$$
$$\cong L(X_{i_1}, \ldots, X_{i_k}; Y)$$

where $i_1, \ldots, i_k$ is a permutation of $(1, \ldots, k)$.

Definition 3.44. Let X and Y be Banach spaces, and U an open subset of X. A map $f : U \to Y$ is called (Fréchet-)differentiable at $x \in U$ if there exists a linear bounded map $L \in L(X, Y)$ such that

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - Lh\|_Y}{\|h\|_X} = 0.$$

If it exists, a map L with the above property is unique. L will be denoted $f'(x)$, $\frac{\mathrm{d}f}{\mathrm{d}x}(x)$ or $\frac{\partial f}{\partial x}(x)$. Moreover, if f is differentiable at every point in U and if the map

$$f' : U \to L(X, Y), \quad x \mapsto f'(x)$$

is continuous, we say f is of class C^1 or continuously differentiable. Proceeding inductively, we define

$$f^{(k)} = (f^{(k-1)})' : U \to L^k(X, Y)$$

27

if it exists. Here we have identified $L(X, L^{k-1}(X, Y))$ with $L^k(X, Y)$. If $f^{(k)}$ exists and is continuous, we say f is of class C^k. f is said to be of class C^∞ or smooth if it is of class C^k for all $k \geq 1$.

Definition 3.45. Let U be a subset of a Banach space. Then a map $F : U \to U$ is called a contraction if there is $L < 1$ such that for all $x, \widetilde{x} \in U$

$$\|F(x) - F(\widetilde{x})\| \leq L \|x - \widetilde{x}\|.$$

Theorem 3.46 (Banach Fixed Point Theorem). *Let U be a closed subset of a Banach space X and $F : U \to U$ a contraction. Then F has a unique fixed point.*

Proof. Choose $x_0 \in U$. We define a sequence x_n inductively by $x_{n+1} := F(x_n)$. First we will show by induction that for all $n \in \mathbb{N}$

$$\|x_{n+1} - x_n\| \leq L^n \|x_1 - x_0\|.$$

For $n = 0$ the statement is trivial. For the inductive step we get

$$\|x_{n+2} - x_{n+1}\| = \|F(x_{n+1}) - F(x_n)\| \leq L \|x_{n+1} - x_n\| \leq L^{n+1} \|x_1 - x_0\|.$$

We use this result to show that x_n is a Cauchy sequence. Because $0 \leq L < 1$ we can find for all $\epsilon > 0$ an $N \in \mathbb{N}$ with

$$L^N < \epsilon \frac{(1 - L)}{\|x_1 - x_0\|}.$$

Using the triangle inequality we get for all $m, n \in \mathbb{N}$ with $m > n \geq N$

$$
\begin{aligned}
\|x_m - x_n\| & \leq & \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \cdots + \|x_{n+1} - x_n\| \\
& \leq & L^{m-1} \|x_1 - x_0\| + L^{m-2} \|x_1 - x_0\| + \cdots + L^n \|x_1 - x_0\| \\
& = & \|x_1 - x_0\| L^n \sum_{k=0}^{m-n-1} L^k < \|x_1 - x_0\| L^n \sum_{k=0}^{\infty} L^k \\
& = & \|x_1 - x_0\| L^n \frac{1}{1 - L} < \epsilon.
\end{aligned}
$$

Now x_n is a Cauchy sequence and converges to a point $x^* \in U$. x^* is a fixed point of F because

$$\|x^* - F(x^*)\| = \lim_{n \to \infty} \|x_n - F(x_n)\| = \lim_{n \to \infty} \|x_n - x_{n+1}\| = 0.$$

To prove uniqueness of the fixed point, assume that x_1^* and x_2^* are fixed points of F. Then

$$0 \leq \|x_1^* - x_2^*\| = \|F(x_1^*) - F(x_2^*)\| \leq L \|x_1^* - x_2^*\|.$$

$0 \leq L < 1$ implies that $\|x_1^* - x_2^*\| = 0$ and thus $x_1^* = x_2^*$. $\qquad\square$

Theorem 3.47. *Let $F : U \times V \subset X \times Y \to U$ with the following properties:*

(a) X and Y are Banach spaces

(b) $U \subset X$ *is closed*

(c) F *is of class* C^r *for* $r \geq 0$

(d) *There is* $L < 1$ *such that for all* $y \in V$ *we have* $\|F(x, y) - F(\widetilde{x}, y)\| \leq L \|x - \widetilde{x}\|$.

Then for all $y \in V$ *there exists a fixpoint* $x^*(y)$ *and the function*

$$x^* : V \to U, \quad y \mapsto x^*(y)$$

is of class C^r.

Proof. (1) We will show that the function x^* is continuous if F is continuous.

$$\|x^*(y + h) - x^*(y)\| =$$
$$= \|F(x^*(y + h), y + h) - F(x^*(y), y + h) + F(x^*(y), y + h) - F(x^*(y), y)\|$$
$$\leq L \|x^*(y + h) - x^*(y)\| + \|F(x^*(y), y + h) - F(x^*(y), y)\|$$

It follows:

$$\|x^*(y + h) - x^*(y)\| \leq \frac{1}{(1 - L)} \|F(x^*(y), y + h) - F(x^*(y), y)\| \qquad (*)$$

Since F is continuous and $(1 - L) > 0$ the continuity of x^* is clear.

(2) We will show that x^* is C^1 if F is C^1. If x^* is differentiable, it must satisfy

$$\frac{\mathrm{d}x^*}{\mathrm{d}y}(y) = \frac{\partial F}{\partial x}(x^*(y), y)\frac{\mathrm{d}x^*}{\mathrm{d}y}(y) + \frac{\partial F}{\partial y}(x^*(y), y).$$

This follows from differentiation of the equality $x^*(y) = F(x^*(y), y)$. To prove that x^* is differentiable, we will apply the Banach Fixed Point Theorem to the operator

$$T(., y) : L(Y, X) \to L(Y, X), \quad A \mapsto \frac{\partial F}{\partial x}(x^*(y), y)A + \frac{\partial F}{\partial y}(x^*(y), y).$$

Because F is a contraction, we have $\left\|\frac{\partial F}{\partial x}\right\| \leq L$, and therefore

$$\left\|T(A, y) - T(\widetilde{A}, y)\right\| = \left\|\frac{\partial F}{\partial x}(x^*(y), y)A - \frac{\partial F}{\partial x}(x^*(y), y)\widetilde{A}\right\| \leq L \left\|A - \widetilde{A}\right\|,$$

so T is a contraction. By what we have shown in step (1) there is a unique fixed point $A^*(y)$ of T and A^* is continuous in y. We want to show that A^* is the derivative of x^*. With $u := x^*(y+h) - x^*(y)$ this is equivalent to $\|u - A^*(y)h\| = o(\|h\|)$. Because

$$\|u - A^*(y)h\| \leq \left\|\left(I - \frac{\partial F}{\partial x}(x^*(y), y)\right)^{-1}\right\| \left\|\left(I - \frac{\partial F}{\partial x}(x^*(y), y)\right)(u - A^*(y)h)\right\|,$$

it is sufficient to show that

$$\left\|\left(I - \frac{\partial F}{\partial x}(x^*(y), y)\right)(u - A^*(y)h)\right\| = o(\|h\|).$$

From the fixed point equation $A^*(y) = T(A^*(y), y)$ we have

$$\left(I - \frac{\partial F}{\partial x}(x^*(y), y)\right)A^*(y) = \frac{\partial F}{\partial y}(x^*(y), y),$$

and therefore

$$\left(I - \frac{\partial F}{\partial x}(x^*(y), y)\right)(u - A^*(y)h) = u - \frac{\partial F}{\partial x}(x^*(y), y)u - \frac{\partial F}{\partial y}(x^*(y), y)h$$

$$= F(x^*(y + h), y + h) - F(x^*(y), y)) - \frac{\partial F}{\partial x}(x^*(y), y)u - \frac{\partial F}{\partial y}(x^*(y), y)h.$$

The above expression is $o(\|u\| + \|h\|)$. Since $F \in C^1$, it satisfies a Lipschitz condition in y, and one can read off equation $(*)$ that $\|u(h)\| = O(\|h\|)$. Therefore the above expression is also $o(\|h\|)$, and this concludes the proof.

(3) We will show by induction that $x^* \in C^r$ if $F \in C^r$. The case $r = 0$ has been treated in step (1). For the inductive step, assume that $F \in C^{r+1}$. Then the map T defined in step (2) is C^r, and the derivative of x^* is the fixed point of T. By the inductive assumption, the derivative of x^* is C^r, and therefore x^* is C^{r+1}. $\qquad\square$

Theorem 3.48 (Inverse Function Theorem). *Let X and Y be Banach spaces, and let $U \subset X$ be an open set in X. Let $f : U \to Y$ be of class C^r, $r \geq 1$, $x_0 \in U$ and suppose that the Fréchet derivative $f'(x_0)$ is a Banach space isomorphism. Then f is a C^r-diffeomorphism of some neighbourhood of x_0 onto some neighbourhood of $f(x_0)$.*

Proof. First we note that it is enough to prove the theorem under the simplifying assumptions that

$$X = Y, \quad u_0 = f(x_0) = 0 \quad \text{and} \quad f'(x_0) = \text{Id}_X.$$

Indeed, we can replace f by the map

$$x \mapsto f'(x_0)^{-1} \circ \big(f(x + x_0) - f(x_0)\big).$$

With these assumptions we have

$$f(x) = y \iff F(x, y) := f(x) - x - y = x.$$

The differential $\frac{\partial F}{\partial x}$ is independent of y and vanishes at $(0, 0)$, so there is $\epsilon > 0$ such that

$$\left\|\frac{\partial F}{\partial x}(x, y)\right\| < \frac{1}{2} \quad \text{whenever} \quad \|x\| < \epsilon.$$

Then

$$F : B(\epsilon, 0) \times B(\epsilon/2, 0) \to B(\epsilon, 0)$$

is well-defined and satisfies the hypothesis of the Banach Fixed Point Theorem 3.47 with $L = 1/2$. The function $x^*(y)$ is of class C^r and is the inverse of f. $\qquad\square$

We will need a generalization of the Constant Rank Theorem 2.14 to Banach spaces.

Definition 3.49. A closed subspace X_1 of a Banach space X is said to be split (or complemented) if there is a closed subspace $X_2 \subset X$ such that $X = X_1 \oplus X_2$.

Remark 3.50. If X and Y are Banach spaces, Y is finite-dimensional, and $f : X \to Y$, then the kernel and the image of f are split. This situation will occur in the analysis of the endpoint map in Section 3.3.3.

Theorem 3.51 (Constant Rank Theorem in Banach spaces). *Let X and Y be Banach spaces, and let $U \subset X$ be an open set in X. Let $f : U \to Y$ be of class C^r, $r \geq 1$, $u_0 \in U$ and suppose that the Fréchet derivative $f'(u_0)$ has closed split image Y_1 with closed complement Y_2 and split kernel X_2 with closed complement X_1. In addition, assume that for all u in a neighbourhood of $u_0 \in U$, $f'(u)(X)$ is a closed subspace of Y and $f'(u)|X_1 : X_1 \to f'(u)(X)$ is a Banach space isomorphism. Then there exist open sets*

$$U_1 \subset Y_1 \oplus X_2, \quad U_2 \subset X, \quad V_1 \subset Y, \quad V_2 \subset Y$$

and there are C^r-diffeomorphisms

$$\varphi : V_1 \to V_2 \quad and \quad \psi : U_1 \to U_2$$

such that $(\varphi \circ f \circ \psi)(x, e) = (x, 0)$.

Proof.

(1) Write $f = f_1 \times f_2$ with $f_i : U \to Y_i$ for $i = 1, 2$. Then

$$\frac{\partial f_1}{\partial x_1}(u_0) : X_1 \to Y_1$$

is a Banach space isomorphism. We define

$$g : U \to Y_1 \oplus X_2, \quad g(x_1, x_2) = (f_1(x_1, x_2), x_2).$$

Then for $(h_{X_1}, h_{X_2}) \in X_1 \oplus X_2$ we have

$$g'(u)(h_{X_1}, h_{X_2}) = \begin{pmatrix} \frac{\partial f_1}{x_1}(u) & \frac{\partial f_1}{x_2}(u) \\ 0 & \mathrm{Id}_{X_2} \end{pmatrix} \begin{pmatrix} h_{X_1} \\ h_{X_2} \end{pmatrix}, \qquad (*)$$

and therefore $g'(u_0) : X = X_1 \oplus X_2 \to Y_1 \oplus X_2$ is a Banach space isomorphism. So by the Inverse Function Theorem 3.48 g has a local inverse $\psi : U_1 \to U_2 \subset U$ of class C^r, $\psi^{-1} = g|U_2$. Then $(f_1 \circ \psi)(y_1, x_2) = y_1$ for $(y_1, x_2) \in U_1$.

(2) We claim that

$$(f \circ \psi)'(y_1, x_2)|Y_1 \times \{0\} : Y_1 \times \{0\} \to (f \circ \psi)'(y_1, x_2)(Y_1 \oplus X_2)$$

31

is a Banach space isomorphism for $(y_1, x_2) \in U_1$. To prove the claim, we first notice by looking at $(*)$ that

$$g'(u)|X_1 \times \{0\} : X_1 \times \{0\} \to Y_1 \times \{0\}$$

is a Banach space isomorphism. Therefore its inverse maps $Y_1 \times \{0\}$ bijectively to $X_1 \times \{0\}$. Furthermore, the hypothesis of the theorem says that $f'(\psi(y_1, x_2))$ is a Banach space isomorphism when restricted to $X_1 \times \{0\}$, and the claim follows by the chain rule

$$(f \circ \psi)'(y_1, x_2) = f'(\psi(y_1, x_2)) \circ \psi'(y_1, x_2).$$

(3) Let $(h_{Y_1}, h_{X_2}) \in Y_1 \oplus X_2$. Then because of (2) there is $\tilde{h}_{Y_1}$ such that

$$(f \circ \psi)'(y_1, x_2)(h_{Y_1}, h_{X_2}) = (f \circ \psi)'(y_1, x_2)(\tilde{h}_{Y_1}, 0). \qquad (**)$$

We see from (1) that

$$(f_1 \circ \psi)'(y_1, x_2)(h_{Y_1}, h_{X_2}) = h_{Y_1},$$

so $\tilde{h}_{Y_1}$ equals h_{Y_1}. But then $(**)$ implies that

$$\frac{\partial f \circ \psi}{\partial x_2}(y_1, x_2) = 0,$$

so $f \circ \psi$ does not depend on the variable x_2, and we can write

$$\tilde{f} : P_{Y_1}(V) \subset Y_1 \to Y, \quad \tilde{f}(y_1) = f \circ \psi(y_1, x_2)$$

where $P_{Y_1} : Y_1 \oplus X_2 \to Y_1$ is the projection.

(4) Let $y_0 = P_{Y_1}(\psi^{-1}(u_0))$. The derivative of $\tilde{f}$ at y_0 is injective and has closed split image Y_1 with closed complement Y_2. Define

$$\tilde{g} : P_{Y_1}(V) \times Y_2 \subset Y_1 \oplus Y_2 \to Y_1 \oplus Y_2, \quad \tilde{g}(y_1, y_2) = \tilde{f}(y_1) + (0, y_2).$$

Then the derivative of $\tilde{g}$ at $(y_0, 0)$ is a Banach space isomorphism, and therefore by the Inverse Function Theorem 3.48 there is a C^r-diffeomorphism $\varphi : V_1 \to V_2$ with $V_1, V_2 \subset Y$ and $\varphi^{-1} = \tilde{g}|V_2$. Because $\tilde{f}(y_1) = \tilde{g}(y_1, 0)$, we have

$$(\varphi \circ \tilde{f})(y_1) = \varphi(\tilde{g}(y_1, 0)) = (y_1, 0) \in Y_1 \oplus Y_2$$

for all y_1 such that $(y_1, 0) \in V_2$. Therefore $(\varphi \circ f \circ \psi)(y_1, x_2) = (y_1, 0)$. $\qquad \square$

3.3.2 Superposition operators

Definition 3.52. Let $f : \Omega \times X_1 \to X_2$ be a function. Then the superposition operator associated to f is a mapping F that assigns to each function $x : \Omega \to X_1$ a function $Fx : \Omega \to X_2$ defined as

$$Fx(\omega) = f(\omega, x(\omega)).$$

F is also called composition operator or Nemytskij operator.

The book of Appell and Zabrejko [2] gives a very good overview of superposition operators acting on different spaces of functions. Our interest lies in the situation where F acts on continuous functions. The results presented here are generalizations of Appell and Zabrejko [2], chap. 6, where only the case $X_1 = X_2 = \mathbb{R}$ is treated.

Definition 3.53. If Ω is a compact metric space and X is a Banach space, then the set $C(\Omega; X)$ of continuous functions from Ω to X is a Banach space when equipped with the norm

$$\|x\|_{C(\Omega; X)} = \sup_{s \in \Omega} \|x(s)\|_X.$$

Lemma 3.54. *Let X_1 and X_2 be Banach spaces, Ω a compact metric space, U an open subset of X_1, and $f : \Omega \times U \to X_2$. If f is continuous, then its associated superposition operator $F : C(\Omega; U) \to C(\Omega; X_2)$ is continuous.*

Proof. If f is continuous, then clearly F maps continuous functions to continuous functions. It remains to show that F is continuous at every $x \in C(\Omega; U)$. Because Ω is compact and x is continuous, the graph of x is a compact subspace of $\Omega \times U$ equipped with the product metric

$$d((\omega_1, x_1), (\omega_2, x_2)) = \max\left(d(\omega_1, \omega_2), \|x_1 - x_2\|_{X_1}\right).$$

f is continuous, so around every point p in the graph of x we find a metric ball $B_{\delta(p)}(p)$ with radius $\delta(p)$ such that whenever q lies in this ball, we have

$$\|f(p) - f(q)\|_{X_2} < \epsilon/2.$$

Finitely many balls $B_{\delta(p)/2}(p)$ cover the graph of x, say the balls $B_{\delta(p_i)/2}(p_i)$ for $i = 1, \ldots, n$. Let δ be the smallest radius of these balls, i.e.

$$\delta = \min\left\{\delta(p_i)/2 : i = 1, \ldots, n\right\}.$$

Now we claim that if y is a function in $C(\Omega; U)$, we have

$$\|x - y\|_{C(\Omega; X_1)} < \delta \implies \|Fx - Fy\|_{C(\Omega; X_2)} < \epsilon.$$

To prove this claim, take a point $p_y = (\omega, y(\omega))$ in the graph of y. Because $p_x = (\omega, x(\omega))$ lies in the graph of x and the graph of x is covered by the balls $B_{\delta(p_i)/2}(p_i)$, there is a point p_{i_0} in the graph of x such that $d(p_x, p_{i_0}) < \delta(p_{i_0})/2$. Now,

$$d(p_y, p_{i_0}) \leq d(p_y, p_x) + d(p_x, p_{i_0}) < \delta + \delta(p_{i_0})/2 \leq \delta(p_{i_0}).$$

So both p_y and p_x lie in the same ball $B_{\delta(p_{i_0})}(p_{i_0})$. Therefore

$$\|f(p_x) - f(p_y)\|_{X_2} \leq \|f(p_x) - f(p_{i_0})\|_{X_2} + \|f(p_{i_0}) - f(p_y)\|_{X_2} < \epsilon/2 + \epsilon/2 = \epsilon.$$

Since this is true for all points $p_x = (\omega, x(\omega))$ and $p_y = (\omega, y(\omega))$, $\omega \in \Omega$, we get the desired inequality $\|Fx - Fy\|_{C(\Omega; X_2)} < \epsilon$. $\qquad\square$

Lemma 3.55. *Let X_1 and X_2 be Banach spaces, Ω a compact metric space, U an open subset of X_1, and $f : \Omega \times U \to X_2$. If f is continuous and has a continuous partial derivative $\frac{\partial f}{\partial x}$, then its associated superposition operator $F : C(\Omega; U) \to C(\Omega; X_2)$ is of class C^1.*

Proof. We claim that the Fréchet derivative of F at $x \in C(\Omega; U)$ is given by

$$F'(x)(h)(\omega) = \frac{\partial f}{\partial x}(\omega, x(\omega))h(\omega) \text{ or equivalently } F'(x) = M(G(x)),$$

where M and G are defined as follows:

$$M : C(\Omega; L(X_1; X_2)) \to L(C(\Omega; X_1); C(\Omega; X_2)), \quad M(g)(h)(\omega) = g(\omega)(h(\omega)).$$

It is easy to see that M is a well-defined bounded linear operator. Therefore M is smooth.

$$G : C(\Omega; U) \to C(\Omega; L(X_1; X_2))$$

is the superposition operator associated to the continuous function

$$\frac{\partial f}{\partial x} : \Omega \times U \to L(X_1; X_2).$$

By Lemma 3.54 we know that G is well-defined and continuous. To prove that $M(G(x))$ really is the Fréchet derivative of F at x, we have to show that

(1) For all $x \in C(\Omega; U)$, $M(G(x))$ is a bounded linear operator from $C(\Omega; X_1)$ to $C(\Omega; X_2)$.

(2) $\|F(x + h) - F(x) - M(G(x))(h)\|_{C(\Omega; X_2)} = o\left(\|h\|_{C(\Omega; X_1)}\right).$

(1) is clear from the definition of M. (2) follows from the Taylor formula for f:

$$\|F(x+h) - F(x) - M(G(x))(h)\|_{C(\Omega;X_2)} =$$

$$= \sup_{\omega \in \Omega} \left\| f(\omega, x(\omega) + h(\omega)) - f(\omega, x(\omega)) - \frac{\partial f}{\partial x}(\omega, x(\omega))h(\omega) \right\|_{X_2}$$

$$= \sup_{\omega \in \Omega} \left\| \int_0^1 \frac{\partial f}{\partial x}(\omega, x(\omega) + th(\omega))h(\omega) \, dt - \frac{\partial f}{\partial x}(\omega, x(\omega))h(\omega) \right\|_{X_2}$$

$$= \sup_{\omega \in \Omega} \left\| \int_0^1 \left(\frac{\partial f}{\partial x}(\omega, x(\omega) + th(\omega)) - \frac{\partial f}{\partial x}(\omega, x(\omega)) \right) h(\omega) \, dt \right\|_{X_2}$$

$$\leq \sup_{\omega \in \Omega} \sup_{0 \leq t \leq 1} \left\| \frac{\partial f}{\partial x}(\omega, x(\omega) + th(\omega)) - \frac{\partial f}{\partial x}(\omega, x(\omega)) \right\|_{L(X_1;X_2)} \|h(\omega)\|_{X_1}$$

$$\leq \|h\|_{C(\Omega;X_1)} \sup_{0 \leq t \leq 1} \|G(x + th) - G(x)\|_{C(\Omega;L(X_1;X_2))}$$

$$= o\left(\|h\|_{C(\Omega;X_1)}\right), \text{ because } G \text{ is continuous.}$$

It remains to show that $x \mapsto F'(x)$ is continuous. This is clear since G is continuous and M is smooth. $\qquad\square$

Theorem 3.56. *Let X_1 and X_2 be Banach spaces, Ω a compact metric space, U a compact subspace of X_1 and $f : \Omega \times U \to X_2$. If f is continuous and has continuous partial derivatives $\frac{\partial^k f}{\partial x^k}$ for $0 \leq k \leq n$, then its associated superposition operator $F : C(\Omega; U) \to C(\Omega; X_2)$ is of class C^n.*

Proof. The case $n = 0$ has been proven in Lemma 3.54, and the case $n = 1$ in Lemma 3.55. We will prove the theorem by induction. For the inductive step, let f have continuous partial derivatives for $1 \leq k \leq n + 1$. By Lemma 3.55, F is of class C^1 and its derivative is given by $F' = M \circ G$. G is the superposition operator associated to the function $\frac{\partial f}{\partial x}$, which is continuous and has n continuous partial derivatives. By the inductive assumption, G is of class C^n. Because M is C^∞, $F' = M \circ G$ is of class C^n. So F is of class C^{n+1}. $\qquad\square$

3.3.3 The endpoint map

In all of section 3.3.3, we will use the following assumptions:

(a) M is an open connected subset of $\mathbb{R}^n$.

(b) H is spanned by a finite number of vectorfields $X_1, \ldots, X_k \in \mathfrak{X}(M)$.

(c) There is $K \in \mathbb{R}$ such that $\|X_i(x)\|_{\mathbb{R}^n} \leq K$ for all $x \in M$.

(d) There is $L \in \mathbb{R}$ such that $\|X_i(x) - X_i(y)\|_{\mathbb{R}^n} \leq L$ for all $x, y \in M$.

Theorem 3.57. *Under the assumptions (a) to (d) there is an open set U,*

$$U \subset L^1([0,1]; \mathbb{R}^k) \times M, \quad \{0\} \times M \subset U$$

such that for each $(u, x) \in U$ the equation

$$\dot{\gamma}(t) = \sum_{i=1}^{k} u_i(t) X_i(\gamma(t)), \quad \gamma(0) = x \qquad (*)$$

has a unique solution $\gamma \in C([0,1]; M)$ which depends smoothly on u and x.

Proof.

(1) For the moment, let us work on an arbitrary interval $[a, b]$ contained in $[0, 1] \subset \mathbb{R}$. We need some conventions about the norms used. Define

$$\|x\|_{\mathbb{R}^n} = \sum_{i=1}^{n} |x_i|$$

$$\|u\|_{L^1([a,b];\mathbb{R}^k)} = \sum_{i=1}^{k} \int_{a}^{b} |u_i(s)| \, \mathrm{d}s$$

$$\|u\|_{C([a,b];\mathbb{R}^k)} = \sum_{i=1}^{k} \max_{s \in [a,b]} |u_i(s)|.$$

(2) Let I denote the integral operator

$$I : \begin{cases} L^1([a,b]; \mathbb{R}^k) & \to & C([a,b]; \mathbb{R}^k) \\ u & \mapsto & (t \mapsto \int_0^t u(s)\mathrm{d}s). \end{cases}$$

Then I is a linear operator with operator norm equal to one, so I is smooth. Define

$$F : \begin{cases} C([a,b]; M) \times L^1([a,b]; \mathbb{R}^k) \times M & \to & C([a,b]; \mathbb{R}^n) \\ (\gamma, u, x) & \mapsto & x + I(\sum_{i=1}^k u_i \, X_i \circ \gamma). \end{cases}$$

Note that F is defined on an open set because $C([a,b]; M)$ is open in $C([a,b]; \mathbb{R}^n)$ whenever $M \subset \mathbb{R}^n$ is open. By Theorem 3.56, F is smooth.

(3) We can rewrite the differential equation $(*)$ as an equivalent integral equation $\gamma = F(\gamma, u, x)$.

(4) Let $\gamma_0 = F(\gamma_0, u_0, x_0)$ be a solution on the interval $[a, b]$ with (γ_0, u_0, x_0) in the domain of definition of F. We claim that under the assumption

$$\|u_0\|_{L^1([a,b]; \mathbb{R}^k)} < \frac{1}{3L}$$

we can find a neighbourhood U of (u_0, x_0) in $L^1([a,b]; \mathbb{R}^k) \times M$ such that for all $(u, x) \in U$ the equation $\gamma = F(\gamma, u, x)$ has a unique solution $\gamma \in C([a,b]; M)$. To prove the claim, we choose $\epsilon > 0$ small enough such that $B(\epsilon, \gamma_0) \subset C([a,b]; M)$. Then we define

$$G : \begin{cases} \overline{B(\epsilon, \gamma_0)} \times \left(B(\frac{\epsilon}{3K}, u_0) \cap B(\frac{1}{3L}, 0) \right) \times B(\frac{\epsilon}{3}, x_0) & \to & \overline{B(\epsilon, \gamma_0)} \\ (\gamma, u, x) & \mapsto & F(x, \gamma, u). \end{cases}$$

First, we show that the image of G really is a subset of $\overline{B(\epsilon, \gamma_0)}$.

$$\|G(\gamma, u, x) - \gamma_0\|_{C([a,b]; \mathbb{R}^n)} = \|G(\gamma, u, x) - G(\gamma_0, u_0, x_0)\|_{C([a,b]; \mathbb{R}^n)}$$

$$\leq \|x - x_0\|_{\mathbb{R}^n} + \left\| \sum_{i=1}^k u_i \, X_i \circ \gamma - u_{0,i} \, X_i \circ \gamma_0 \right\|_{L^1([a,b]; \mathbb{R}^n)}$$

$$\leq \|x - x_0\|_{\mathbb{R}^n} + \sum_{i=1}^k \|u_i \, X_i \circ \gamma - u_{0,i} \, X_i \circ \gamma\|_{L^1([a,b]; \mathbb{R}^n)} +$$

$$+ \sum_{i=1}^k \|u_{0,i} \, X_i \circ \gamma - u_{0,i} \, X_i \circ \gamma_0\|_{L^1([a,b]; \mathbb{R}^n)}$$

$$\leq \|x - x_0\|_{\mathbb{R}^n} + \|u - u_0\|_{L^1([a,b]; \mathbb{R}^k)} K + \|u_0\|_{L^1([a,b]; \mathbb{R}^k)} L \|\gamma - \gamma_0\|_{C([a,b]; \mathbb{R}^n)}$$

$$\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Furthermore, G satisfies the contraction property

$$\|G(\gamma, u, x) - G(\tilde{\gamma}, u, x)\|_{C([a,b]; \mathbb{R}^n)} \leq \sum_{i=1}^k \|u_i(X_i \circ \gamma - X_i \circ \tilde{\gamma})\|_{L^1([a,b]; \mathbb{R}^n)}$$

$$\leq \|u\|_{L^1([a,b]; \mathbb{R}^k)} L \|\gamma - \tilde{\gamma}\|_{C([a,b]; \mathbb{R}^n)} \leq \frac{1}{3} \|\gamma - \tilde{\gamma}\|_{C([a,b]; \mathbb{R}^n)}$$

for all (γ, u, x) and $(\tilde{\gamma}, u, x)$ in the domain of G. Therefore we can apply the Banach Fixed Point Theorem 3.47 to G, and this proves our claim with

$$U = \left(B(\frac{\epsilon}{3K}, u_0) \cap B(\frac{1}{3L}, 0) \right) \times B(\frac{\epsilon}{3}, x_0).$$

(5) Now we will work on the interval $[0, 1]$ and drop the assumption that

$$\|u_0\|_{L^1([0,1];\mathbb{R}^k)} < \frac{1}{3L}.$$

Let $\gamma_0 = F(\gamma_0, u_0, x_0)$ be a solution on the interval $[0, 1]$ with (γ_0, u_0, x_0) in the domain of definition of F. We claim that we can find a neighbourhood U of (u_0, x_0) in $L^1([0, 1]; \mathbb{R}^k) \times M$ such that for all $(u, x) \in U$ the equation $\gamma = F(\gamma, u, x)$ has a unique solution $\gamma \in C([0, 1]; M)$. To prove the claim, we choose $\epsilon > 0$ small enough such that $B(\epsilon, \gamma_0) \subset C([a, b]; M)$. Let $0 = t_0 < \cdots < t_N = 1$ such that on each interval $[t_j, t_{j+1}]$

$$\left\| u_0|_{[t_j, t_{j+1}]} \right\|_{L^1([t_j, t_{j+1}];\mathbb{R}^n)} < \frac{1}{3L}$$

holds. Then we can apply the result obtained in (4) to each of the intervals $[t_j, t_{j+1}]$ with

$$a = t_j, b = t_{j+1}, \gamma_0 = \gamma_0|_{[t_j, t_{j+1}]}, u_0 = u_0|_{[t_j, t_{j+1}]}, x_0 = \gamma(t_j).$$

We want to concatenate the solutions on these intervals to get a solution on $[0, 1]$, so we have to make sure that the endpoint of a solution curve on $[t_j, t_{j+1}]$ lies in $B(\frac{\epsilon}{3}, \gamma(t_{j+1}))$. This is the fact for

$$U = \left(B(\frac{\epsilon}{3^N}, u_0) \cap B(\frac{1}{3L}, 0) \right) \times B(\frac{\epsilon}{3^N}, x_0).$$

The solution on each interval $[t_j, t_{j+1}]$ depends smoothly on the control and initial value. Because the operator assigning to each curve its endpoint is a linear, continuous operator $C([t_j, t_{j+1}]; M) \to M$, it is smooth, and therefore the concatenation of the solutions curves depends smoothly on x_0 and u_0. □

The remaining part of this section is taken from Bellaïche [3].

Definition 3.58. In the situation of Theorem 3.57, we define

$$\mathrm{Evol}_t^u(x) = \gamma(t) \quad \text{and} \quad \mathrm{End}_x(u) = \gamma(1)$$

for $t \in [0, 1]$, $(u, x) \in U$ and a curve $\gamma \in C([0, 1]; M)$ with initial value x controlled by u. We will call End_x the endpoint map and Evol_t^u the evolution of the controlled vectorfield $\sum_{i=1}^k u_i X_i$.

Remark 3.59. End_x is smooth for each $x \in M$. This is true because the evaluation operator

$$C([0,1]; \mathbb{R}^n) \to \mathbb{R}^n, \quad \gamma \mapsto \gamma(1)$$

is linear and bounded, and therefore smooth, and because furthermore γ depends smoothly on its control u and initial value x by Theorem 3.57.

Definition 3.60. For $u, v \in L^1([0,1]; \mathbb{R}^k)$ and $0 < s < 1$, let $u *_s v \in L^1([0,1]; \mathbb{R}^k)$ be the control function defined by

$$(u *_s v)(t) = \begin{cases} \frac{1}{s} \, u(\frac{t}{s}) & \text{if } 0 \leq t < s \\ \frac{1}{1-s} \, v(\frac{t-s}{1-s}) & \text{if } s \leq t \leq 1. \end{cases}$$

In many cases we will use $s = \frac{1}{2}$, so we define $u * v = u *_{\frac{1}{2}} v$ for the sake of a simpler notation. Furthermore, we define

$$\check{u}(t) = -u(1-t) \quad \text{for } 0 \leq t \leq 1.$$

$u *_s v$ will be called the concatenation of u and v, and $\check{u}$ the inverse control of u.

Remark 3.61. It is easy to check the following properties:

(1) $\|u *_s v\|_{L^1([0,1]; \mathbb{R}^k)} = \|u\|_{L^1([0,1]; \mathbb{R}^k)} + \|v\|_{L^1([0,1]; \mathbb{R}^k)}$.

(2) $\text{Evol}_1^{u *_s v} = \text{Evol}_1^v \circ \text{Evol}_1^u$

(3) Evol_1^u is a local diffeomorphism with inverse $\text{Evol}_1^{\check{u}}$.

Definition 3.62. Let r denote the maximal rank of End_x. A control $u \in U$ is called normal if the rank of End_x at u is equal to r.

Lemma 3.63. *If u is a normal control, so is $u *_s v$ for any $0 < s < 1$.*

Proof. The tangent space in $L^1([0,1]; \mathbb{R}^k)$ at $u *_s v$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 (u *_s v) + th, \quad h \in L^1([0,1]; \mathbb{R}^k).$$

The tangent vectors of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 (u + th) *_s v, \quad h \in L^1([0,1]; \mathbb{R}^k)$$

constitute a vector subspace of the tangent space at $u *_s v$. We have

$$T\,\text{End}_x \left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 (u + th) *_s v \right) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 \text{End}_x \left((u + th) *_s v \right) =$$

$$= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 \text{Evol}_1^v \circ \text{End}_x \, (u + th) =$$

$$= T\,\text{Evol}_1^v \circ T\,\text{End}_x \left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 u + th \right).$$

38

The vectors

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 u + th, \quad h \in L^1([0,1]; \mathbb{R}^k)$$

constitute the tangent space $T_u L^1([0,1]; \mathbb{R}^k))$. Let r denote the maximal rank of End_x. Then because we assumed that u is a normal control, the dimension of $T\,\mathrm{End}_x(T_u L^1([0,1]; \mathbb{R}^k))$ is r. Evol_1^v is a local diffeomorphism, and this concludes the proof. $\qquad\square$

Lemma 3.64. *For every point y in the image of End_x there is a normal control steering x to y.*

Proof. Suppose y is attained from x by means of a control u. Choose any normal control v. Then the control $v * \check{v} * u$ steers x to y and is normal by Lemma 3.63. $\quad\square$

Theorem 3.65. *For $x \in M$, the image of End_x is an initial submanifold, and its inclusion into M is an initial mapping.*

Proof.

(1) Let N denote the set of controls in the domain of End_x with maximal rank r. By Lemma 3.64, we can restrict End_x to N without loosing any points in the image of End_x, and we will do this without change of notation.

(2) Now, End_x is a smooth map with constant rank r. The Constant Rank Theorem 3.51 gives us for every $u \in N$ a neighbourhood U of u such that $\mathrm{End}_x(U)$ is a small r-dimensional submanifold of M. To make $\mathrm{End}_x(N)$ an immersed submanifold, we give it the coarsest topology such that End_x is an open mapping, and we use the charts of the small submanifolds $\mathrm{End}_x(U)$ as charts. We need to check that each set $\mathrm{End}_x(U)$ is open in $\mathrm{End}_x(N)$, which is trivial, and that $\mathrm{End}_x(U_1) \cap \mathrm{End}_x(U_2)$ is open. Let

$$y = \mathrm{End}_x(u_1) = \mathrm{End}_x(u_2) \in \mathrm{End}_x(U_1) \cap \mathrm{End}_x(U_2).$$

We look at the control $u_3 := u_1 * u_2^{-1} * u_2$, for which the Constant Rank Theorem gives us a neighbourhood U_3 of u_3 such that $\mathrm{End}_x(U_3)$ is a submanifold. By making U_1 and U_2 small enough, both $U_1 * u_2^{-1} * u_2$ and $u_1 * u_2^{-1} * U_2$ can be made a subset of U_3. So both $\mathrm{End}_x(U_1) = \mathrm{End}_x(U_1 * u_2^{-1} * u_2)$ and $\mathrm{End}_x(U_2) = \mathrm{End}_x(u_1 * u_2^{-1} * U_2)$ are r-dimensional submanifolds of M lying in the r-dimensional manifold $\mathrm{End}_x(U_3)$. Hence their intersection is open in $\mathrm{End}_x(U_3)$, and therefore also open in $\mathrm{End}_x(N)$. From these considerations we know that $(\mathrm{End}_x(N), i)$ is an immersed submanifold, where $i : \mathrm{End}_x(N) \to M$ is the inclusion.

(3) The rest of the proof is basically the same as the proof of Theorem 3.14. We will construct initial submanifold charts for $\mathrm{End}_x(N)$. For $y = \mathrm{End}_x(u) \in \mathrm{End}_x(N)$, by the Constant Rank Theorem 3.51, there is a chart (U, ϕ) centered at $u \in N$ and a chart (V, ψ) centered at $y \in M$ such that

$$(\psi \circ \mathrm{End}_x \circ \phi^{-1})(t^1, \ldots, t^r, e) = (t^1, \ldots, t^r, 0, \ldots, 0).$$

Then

$$f : (t^1, \ldots, t^n) \mapsto \mathrm{End}_{\psi^{-1}(0,\ldots,0,t^{r+1},\ldots,t^n)} \circ \phi^{-1}(t^1, \ldots, t^r, 0, \ldots, 0)$$

is a diffeomorphism near zero. Let (W, χ) be the chart given by f^{-1}, suitably restricted. For all $t^1, \ldots, t^n$ such that both expressions make sense, we have

$$f(t^1, \ldots, t^n) \in \mathrm{End}_x(N) \iff f(0, \ldots, 0, t^{r+1}, \ldots, t^n) \in \mathrm{End}_x(N).$$

From this we see that $\mathrm{End}_x(N) \cap W$ is the disjoint union of connected sets of the form

$$\{w \in W : \chi^{r+1}(w), \ldots, \chi^n(w) = \mathrm{constant}\}.$$

Since $\mathrm{End}_x(N)$ is a connected immersed submanifold of M, it is second countable, and only a countable set of constants can appear in the description of $\mathrm{End}_x(N) \cap W$ given above. Therefore χ is a distinguished chart as in Theorem 3.14. Furthermore, we see that

$$\chi(C_x(\mathrm{End}_x(N) \cap W)) = \chi(W) \cap \mathbb{R}^r \times \{0\},$$

so χ is an initial submanifold chart. Because we can find such a chart around every point in $\mathrm{End}_x(N)$, $\mathrm{End}_x(N)$ is an initial submanifold.

(4) According to Lemma 2.22, there is a unique smooth manifold structure that makes $\mathrm{End}_x(N)$ an immersed submanifold with the property that the inclusion is an initial mapping. It can be seen from the proof of this Lemma that charts for this structure are given by restricting the initial submanifold charts (W, χ) to $C_x(\mathrm{End}_x(N) \cap W)$. Restricting χ to $C_x(\mathrm{End}_x(N) \cap W)$ yields the inverse of the map

$$(t^1, \ldots, t^r) \mapsto \mathrm{End}_x \circ \phi^{-1}(t^1, \ldots, t^r, 0, \ldots, 0),$$

and this map already is a chart of $\mathrm{End}_x(N)$. $\qquad\square$

Theorem 3.66. *For $x \in M$, the accessible set $\mathrm{Acc}(x)$ equals the image of End_x.*

Proof. Let L denote the image of End_x. Of course, L is horizontally pathwise connected, so we have $L \subset \mathrm{Acc}(x)$. We will show that $\mathrm{Acc}(x) \subset L$ in a few steps. For the sake of a simpler notation let $\|\cdot\|$ denote $\|\cdot\|_{L^1([0,1];\mathbb{R}^k)}$.

(1) We claim that for any function $u \in L^1([0,1];\mathbb{R}^k)$ we have

$$\lim_{s \to 1} \|u *_s 0 - u\| = 0,$$

where 0 denotes the zero control. This is clear for continuous functions u by dominated convergence, and it still holds for arbitrary u by a density argument: Indeed, for $\epsilon > 0$ we can find a continuous function v with $\|u - v\| < \epsilon/3$. Then

$$\|u *_s 0 - u\| \le \|u *_s 0 - v *_s 0\| + \|v *_s 0 - v\| + \|v - u\|$$

$$= \|u - v\| + \|v *_s 0 - v\| + \|v - u\| \le \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

for s close enough to one, and this proves the claim.

(2) For $j = 1, \ldots, k$, let e_j be the unit vectors in $\mathbb{R}^k$, and let u be a normal control. From

$$\|u *_s ((1-s)\, e_j) - u\| \leq \|u *_s ((1-s)\, e_j) - u *_s 0\| + \|u *_s 0 - u\|$$
$$= \|(1-s)\, e_j\| + \|u *_s 0 - u\|$$

we see that $u *_s ((1-s)\, e_j)$ is still a normal control for s close to one because the set of normal controls is open. But $\mathrm{End}_x(u *_s ((1-s)\, e_j)) = Fl_{1-s}^{X_j}(\mathrm{End}_x(u))$. Therefore we can conclude that L is stable under the flows of X_j. The flows of X_j are smooth mappings taking values in L. Because $i : L \to M$ is an initial mapping, the flows are also smooth as mappings into L, and therefore the vectorfields X_j are tangent to L. We conclude that the tangent space of L contains H.

(3) We claim that a horizontal path never leaves L. The proof is exactly the same as the one given in Theorem 3.20. $\qquad \square$

With Theorem 3.66 we are able to prove Chow's Theorem. However, we will first prove the slightly stronger

Theorem 3.67. *If H is a bracket generating distribution on M, then the endpoint map is open and its maximal rank equals the dimension of M.*

Proof. We have to show that for an open subset U of the domain of definition of End_x and for $y = \mathrm{End}_x(u) \in \mathrm{End}_x(U)$, there is a neighbourhood of y that is contained in $\mathrm{End}_x(U)$.

(1) Because $\|u *_s v - u *_s 0\| = \|v\|$ for an arbitrary control v and by looking at step (1) in the proof of Theorem 3.66 we see that there is a neighbourhood V of 0 and some s close to one such that $u *_s V \subset U$. Then also

$$\mathrm{End}_y(V) = \mathrm{End}_x(u *_s V) \subset \mathrm{End}_x(U).$$

It is sufficient to show that $\mathrm{End}_y(V)$ is a neighbourhood of y.

(2) For $\epsilon > 0$ let $r(\epsilon)$ denote the maximal rank of End_y on the set $B(\epsilon, 0)$ in $L^1([0,1]; \mathbb{R}^k)$. r is monotone and non-zero, so there is $\epsilon_0 > 0$ such that $r(\epsilon) = r(\epsilon_0)$ for all $\epsilon < \epsilon_0$. We can choose ϵ_0 even smaller so that additionally $B(\epsilon_0, 0) \subset V$. Let V' denote the set of controls in $B(\epsilon_0, 0)$ where End_y has maximal rank $r(\epsilon_0)$. Then V' is open. $B(\epsilon/2, 0)$ contains a control v where End_y has rank $r(\epsilon_0)$. Then $v * \check{v} \in V'$, and $y = \mathrm{End}_y(v * \check{v}) \in \mathrm{End}_y(V')$. It is sufficient to show that $\mathrm{End}_y(V')$ is neighbourhood of y.

(3) End_y has constant rank on V', so by the Constant Rank Theorem 3.51 there is an even smaller neighbourhood V'' of $v * \check{v}$ such that $\mathrm{End}_y(V'')$ is a small submanifold L. It is sufficient to show that L is a neighbourhood of y.

(4) By the same arguments as in step (2) in the proof of Theorem 3.66 the vectorfields X_j are tangent to L, and so are all the brackets of X_j. Because of the bracket generating condition, the dimension of L equals the dimension of M, and therefore L is an open neighbourhood of y. $\qquad \square$

3.3.4   Proof of Chow's Theorem

Theorem 3.68 (Chow's Theorem)**.** *Let H be a locally finitely generated, bracket generating distribution on a manifold M. Then any two points in M can be joined by a horizontal path.*

Proof. First, we will work locally, and therefore we can assume that M is a connected, open subset of $\mathbb{R}^n$, and that H is spanned by $X_1, \ldots, X_k \in \mathfrak{X}(M)$. By making M small enough we can assure that assumptions (a) to (d) of Section 3.3.3 hold. So we can apply Theorem 3.66 which asserts that for any $x \in M$, $Acc(x)$ equals the image of End_x, which is an initial submanifold. Horizontal vectorfields are tangent to it, and so are all their brackets. Because of the bracket generating condition its dimension equals the dimension of M, and therefore it is open.

Now we return to the global case. By the previous considerations, every point in M has a horizontally pathwise connected neighbourhood. Then by Lemma 3.39 M is horizontally pathwise connected. $\qquad\square$

4   The Poincaré Lemma

The Poincaré Lemma in its general form states that locally, all closed differential forms are exact. For one-forms, this results to the following lemma:

Lemma 4.1. *Let U be a star domain in $\mathbb{R}^n$. Let $h_i : U \to \mathbb{R}$ be smooth functions such that $\frac{\partial h_i}{\partial x_j} = \frac{\partial h_j}{\partial x_i}$ for all $1 \leq i, j \leq n$. Then there exists a smooth function $f : U \to \mathbb{R}$ such that $h_i = \frac{\partial f}{\partial x_i}$.*

In the above situation f can be recovered from the functions h_i by integration. Indeed, if $\gamma : [0,1] \to \mathbb{R}^n$ is an absolutely continuous curve, $\gamma(0) = x_0$, $\gamma(1) = x$ and $u_i = \frac{\mathrm{d}\gamma_i}{\mathrm{d}t}$, then

$$f(x) = f(x_0) + \int_0^1 \frac{\mathrm{d}(f \circ \gamma)}{\mathrm{d}t}(t) \, \mathrm{d}t = \int_0^1 \sum_{i=1}^n u_i(t) h_i(\gamma(t)) \, \mathrm{d}t.$$

It is possible to adapt this idea to subriemannian geometry. If H is a bracket generating distribution, then by Chow's Theorem 3.3 the points x and x_0 can be connected by a horizontal path γ. Using the same procedure as above, f can be reconstructed from its horizontal derivatives alone. Indeed, if γ is controlled by horizontal vectorfields $X_1, \ldots, X_k$ with controls $u_i, \ldots, u_k \in L^1([0,1]; \mathbb{R})$, then

$$f(x) = f(x_0) + \int_0^1 \frac{\mathrm{d}(f \circ \gamma)}{\mathrm{d}t}(t) \, \mathrm{d}t = \int_0^1 \sum_{i=1}^k u_i(t)(X_i f)(\gamma(t)) \, \mathrm{d}t. \qquad (*)$$

In the following we will study how continuity and differentiability is passed on from the horizontal derivatives of f to f.

Lemma 4.2. *Let H be a regular bracket generating distribution endowed with a smooth subriemannian metric. Let $f : M \to \mathbb{R}$ be a function such that*

$$Xf : M \to \mathbb{R}, \quad (Xf)(x) := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_0 f \circ Fl_t^X(x)$$

exists and is continuous for all $X \in \mathfrak{X}_{loc,H}(M)$. Then for every point $x_0 \in M$ there is a neighbourhood U of x_0 and a constant C such that the following statement holds for all $x \in U$:

$$|f(x) - f(x_0)| \leq Cd(x, x_0).$$

Here $d(x, x_0)$ is the subriemannian distance of x and x_0 as in Definition 1.17.

Proof. This is a local statement, so we can assume that there are everywhere defined horizontal vectorfields $X_1, \ldots, X_k$ spanning H. Because H is regular, we can assume additionally that $X_1, \ldots X_k$ are orthonormal. Then for all horizontal curves γ that connect x_0 to x by means of L^1-controls u_i we have

$$|f(x) - f(x_0)| \leq \int_0^1 \sum_{i=1}^k |u_i(t)|\,|(X_if)(\gamma(t))|\ \mathrm{d}t$$

$$\leq \int_0^1 \left(\sum_{i=1}^k u_i(t)^2\right)^{1/2} \left(\sum_{i=1}^k (X_if)(\gamma(t))^2\right)^{1/2}\ \mathrm{d}t$$

$$\leq C \int_0^1 \left(\sum_{i=1}^k u_i(t)^2\right)^{1/2}\ \mathrm{d}t = C \int_0^1 \|\dot{\gamma}(t)\|\ \mathrm{d}t = C \cdot l(\gamma),$$

where C bounds $(\sum_{i=1}^k (X_if)^2)^{1/2}$. Therefore

$$|f(x) - f(x_0)| \leq C \inf_\gamma\ l(\gamma) = Cd(x, y). \qquad \square$$

It is important to note that $|f(x) - f(y)| \leq Cd(x, y)$ does not imply that f is Lipschitz continuous with respect to the Euclidean structure in a chart. This will be illustrated by the following example.

Example 4.3. We look at the function

$$f(x, y, z) = \left(|x|^{5/2} + |y|^{5/2} + |z|^{3/2}\right)^{1/2}$$

on the Heisenberg geometry, i.e. $\mathbb{R}^3$ endowed with the distribution H spanned by the vectorfields

$$X_1 = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial z} \quad \text{and} \quad X_2 = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial z}.$$

From $[X_1, X_2] = \frac{\partial}{\partial z}$ we see that H is bracket generating. We calculate

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{5}{4}\left(|x|^{5/2} + |y|^{5/2} + |z|^{3/2}\right)^{-1/2} |x|^{3/2} \operatorname{sgn}(x) & \text{for } (x, y, z) \neq (0, 0, 0) \\ 0 & \text{for } (x, y, z) = (0, 0, 0) \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} \frac{5}{4}\left(|x|^{5/2} + |y|^{5/2} + |z|^{3/2}\right)^{-1/2} |y|^{3/2} \operatorname{sgn}(y) & \text{for } (x, y, z) \neq (0, 0, 0) \\ 0 & \text{for } (x, y, z) = (0, 0, 0) \end{cases}$$

$$\frac{\partial f}{\partial z} = \begin{cases} \frac{3}{4}\left(|x|^{5/2} + |y|^{5/2} + |z|^{3/2}\right)^{-1/2} |z|^{1/2} \operatorname{sgn}(z) & \text{for } (x, y, z) \neq (0, 0, 0) \\ \pm\infty & \text{for } (x, y, z) = (0, 0, 0). \end{cases}$$

Because of the singularity of $\frac{\partial f}{\partial z}$ at zero, f is not Lipschitz continuous with respect to the Euclidean structure. However, the functions

$$\frac{\partial}{\partial x} f, \quad \frac{\partial}{\partial y} f, \quad \left(-\frac{y}{2}\frac{\partial}{\partial z}\right) f \quad \text{and} \quad \left(\frac{x}{2}\frac{\partial}{\partial z}\right) f$$

are continuous, as the following estimates show. We will denote the p-norm on $\mathbb{R}^3$ by $\|\cdot\|_p$, and we will use the fact that all norms on $\mathbb{R}^3$ are equivalent.

$$\left|\frac{\partial f}{\partial x}\right| \leq \frac{5}{4}\left(|x|^{5/2} + |y|^{5/2} + |z|^{5/2}\right)^{-1/2} |x|^{3/2}$$
$$\leq \frac{5}{4}(\|(x, y, z)\|_{5/2})^{-5/4}(\|(x, y, z)\|_1)^{3/2} \leq \operatorname{const}(\|(x, y, z)\|_1)^{1/4}.$$

$$\left|\frac{\partial f}{\partial y}\right| \leq \frac{5}{4}\left(|x|^{5/2} + |y|^{5/2} + |z|^{5/2}\right)^{-1/2} |y|^{3/2}$$
$$\leq \frac{5}{4}(\|(x, y, z)\|_{5/2})^{-5/4}(\|(x, y, z)\|_1)^{3/2} \leq \operatorname{const}(\|(x, y, z)\|_1)^{1/4}.$$

$$\left|\frac{y}{2}\frac{\partial f}{\partial z}\right| \leq \left|\frac{y}{2}\right|\frac{3}{4}\left(|x|^{5/2} + |y|^{5/2} + |z|^{5/2}\right)^{-1/2} |z|^{1/2}$$
$$\leq \frac{3}{8}\|(x, y, z)\|_1\left(\|(x, y, z)\|_{5/2}\right)^{-5/4}(\|(x, y, z)\|_1)^{1/2} \leq \operatorname{const}(\|(x, y, z)\|_1)^{1/4}.$$

$$\left|\frac{x}{2}\frac{\partial f}{\partial z}\right| \leq \left|\frac{x}{2}\right|\frac{3}{4}\left(|x|^{5/2} + |y|^{5/2} + |z|^{5/2}\right)^{-1/2} |z|^{1/2}$$
$$\leq \frac{3}{8}\|(x, y, z)\|_1\left(\|(x, y, z)\|_{5/2}\right)^{-5/4}(\|(x, y, z)\|_1)^{1/2} \leq \operatorname{const}(\|(x, y, z)\|_1)^{1/4}.$$

Therefore also $X_1 f$ and $X_2 f$ are continuous. Because any horizontal vectorfield X can be written as $X = f_1 X_1 + f_2 X_2$ with smooth functions f_1 and f_2, Xf is continuous for all horizontal vectorfields X. So the conditions of Lemma 4.2 are satisfied, and the lemma implies that f is Lipschitz continuous with respect to the subriemannian distance. But, as we have seen above, f is not Lipschitz continuous with respect to the Euclidean structure. Note that also the conditions of Theorem 4.5 are satisfied with $r = 0$, and that the theorem implies the continuity of f.

Remark 4.4. Let us suppose that M is an n-dimensional subriemannian manifold with distribution H, and that k brackets are needed to express the gradient $(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n})$ in terms of horizontal vectorfields, i.e. k is the smallest number such that $H_k = TM$. For some $r \geq k$, let $f : M \to \mathbb{R}$ be a function such that

$$Xf : M \to \mathbb{R}, \quad (Xf)(x) := \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_0 f \circ Fl_t^X(x)$$

exists and is of class C^r for all $X \in \mathfrak{X}_{loc,H}(M)$. Then the gradient $(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n})$ of f is of class C^{r-k}, and therefore f is of class C^{r-k+1}. However, the following theorem yields a stronger result, namely that f is of class C^r no matter how many brackets are needed to span the tangent space.

Theorem 4.5. *Let H be a locally finitely generated, bracket generating distribution on a manifold M, and $r \geq 0$. If $f : M \to \mathbb{R}$ is a function such that*

$$Xf : M \to \mathbb{R}, \quad (Xf)(x) := \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_0 f \circ Fl_t^X(x)$$

exists and is of class C^r for all $X \in \mathfrak{X}_{loc,H}(M)$, then f is also of class C^r.

Proof. We will show that f is of class C^r at every $x_0 \in M$. This is a local question, so we can assume that conditions (a) to (d) of Section 3.3.3 are satisfied, so that the endpoint map is well-defined.

(1) First, we will construct a smooth local section of the endpoint map. That is a smooth map u defined near x_0 such that $\mathrm{End}_{x_0}(u(x)) = x$. By Lemma 3.64 there is a normal control u_0 such that $\mathrm{End}_{x_0}(u_0) = x_0$. By Theorem 3.67 End_{x_0} is open and the rank of End_{x_0} at u_0 equals the dimension of M, say n. End_{x_0} has constant rank near u_0, so by the Constant Rank Theorem 3.51 there is a subspace $E \subset L^1([0,1]; \mathbb{R}^k)$, an open neighbourhood U of u_0, an open neighbourhood V of x and charts

$$\begin{aligned} \phi : U \subset L^1([0,1]; \mathbb{R}^k) &\to \phi(U) \subset \mathbb{R}^n \times E, \\ \psi : V \subset M &\to \psi(V) \subset \mathbb{R}^n \end{aligned}$$

such that for all $(t^1, \ldots, t^n, e) \in \phi(U)$ we have

$$\psi \circ \mathrm{End}_{x_0} \circ \phi(t^1, \ldots, t^n, e) = (t^1, \ldots, t^n).$$

Without loss of generality we can assume that $\phi(u_0) = 0$. Then the map

$$u : \begin{cases} \psi^{-1}(\{(t^1, \ldots, t^n) : (t^1, \ldots, t^n, 0) \in \phi(U)\}) \to L^1([0,1]; \mathbb{R}^k) \\ \qquad\qquad\qquad\qquad\qquad\qquad x \mapsto \phi^{-1}(\psi(x), 0) \end{cases}$$

has the desired properties.

(2) Let Γ be the operator assigning to a control in the domain of End_{x_0} the unique curve γ with this control and $\gamma(0) = x_0$. By Theorem 3.57 Γ is smooth as a map into $C([0,1]; M)$.

(3) By Theorem 3.56 about superposition operators and because $X_i f$ is of class C^r, the map
$$C([0, 1]; M) \to C([0, 1]; \mathbb{R}), \quad \gamma \mapsto (X_i f) \circ \gamma$$
is of class C^r.

(4) Let I denote the map
$$L^1([0, 1]; \mathbb{R}^k) \times C([0, 1]; \mathbb{R}^k) \to \mathbb{R}, \quad (u, h) \mapsto \int_0^1 \sum u_i(t) h_i(t) \mathrm{d}t.$$

I is smooth because it is linear and bounded.

(5) From equation $(*)$, using the notation $X = (X_1, \dots, X_k)$, we have
$$f = f(x_0) + I\left(u, (Xf) \circ \Gamma \circ u\right).$$

Therefore f is of class C^r. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

References

[1] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, tensor analysis, and applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1988. ISBN 0-387-96790-7.

[2] Jürgen Appell and Petr. P. Zabrejko. *Nonlinear superposition operators*. Cambridge University Press, 1990.

[3] A. Bellaïche. The tangent space in sub-Riemannian geometry. *J. Math. Sci. (New York)*, 83(4):461–476, 1997. ISSN 1072-3374. Dynamical systems, 3.

[4] Theodor Bröcker and Klaus Jänich. *Einführung in die Differentialtopologie*. Springer-Verlag, Berlin, 1973. Heidelberger Taschenbücher, Band 143.

[5] Velimir Jurdjevic. *Geometric control theory*, volume 52 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997. ISBN 0-521-49502-4.

[6] Markus Mauhart and Peter W. Michor. Commutators of flows and fields. *Arch. Math. (Brno)*, 28(3-4):229–236, 1992. ISSN 0044-8753.

[7] Peter W. Michor. Topics in differential geometry. 2008. URL `http://www.mat.univie.ac.at/~michor/dgbook.pdf`.

[8] Richard Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002. ISBN 0-8218-1391-9.

[9] Alexander Nagel, Elias M. Stein, and Stephen Wainger. Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.*, 155(1-2):103–147, 1985. ISSN 0001-5962.

[10] Peter Stefan. Accessibility and foliations with singularities. *Bull. Amer. Math. Soc.*, 80:1142–1145, 1974.

[11] Héctor J. Sussmann. Orbits of families of vector fields and integrability of distributions. *Trans. Amer. Math. Soc.*, 180:171–188, 1973. ISSN 0002-9947.

Index