Communications in **M**athematical **A**nalysis
Volume 11, Number 1, pp. 1–40 (2011) www.math-res-pub.org/cma
ISSN 1938-9787

FOLIATING METRIC SPACES: A GENERALIZATION OF FROBENIUS' THEOREM

CRAIG CALCATERRA*
Department of Mathematics
Metropolitan State University
Saint Paul, MN 55106, USA

(Communicated by Toka Diagana)

Abstract

Using families of curves to generalize vector fields, the Lie bracket is defined on a metric space, M. For M complete, versions of the local and global Frobenius theorems hold, and flows are shown to commute if and only if their bracket is zero.

An example is given showing $L^2(\mathbb{R})$ is controllable by two elementary flows.

AMS Subject Classification: 51F99; 93B29; 53C12

Keywords: metric space, Banach space, flow, nonsmooth, foliation, infinite-dimensional control theory, Nagumo-Brézis Theorem

1 Introduction

The main goal of this paper is to further the point of view that many beautiful geometrical and analytical results valid on differentiable manifolds hold on general metric spaces. The main result of this paper is the generalization of Frobenius' Foliation Theorem, Theorem 5.4. Besides the wider relevance gained by generalization, the foundations of the subject are clarified when the limits of applicability are explored. This effort has a long and often disjointed history, only one sliver of which is directly relevant here. The approach in this paper, which has been used by several others, is to use the well-known characterization of a vector in a tangent space as an equivalence class of curves which are tangent to each other. A **curve** c on a metric space (M,d) is a continuous map $c:(\alpha,\beta)\to M$ where $(\alpha,\beta)\subset\mathbb{R}$. Two curves $c_i:(\alpha_i,\beta_i)\to M$ for $i=1,2$ are **tangent at** $t\in(\alpha_1,\beta_1)\cap(\alpha_2,\beta_2)$ if

$$\lim_{h\to 0}\frac{d\left(c_1(t+h),c_2(t+h)\right)}{h}=0.$$

In this way we may generalize a vector field (a family of vectors) on a manifold as an arc field (a family of curves) on a metric space—Definition 2.1, below.

*E-mail address: Craig.Calcaterra@metrostate.edu

It has been said the three pillars of differential geometry are: (I) the Inverse Function Theorem, (II) the Existence Theorem for ordinary differential equations (ODEs) and (III) Frobenius' Theorem. All of these classical theorems may be written with vector fields on manifolds and so may also be written with arc fields on metric spaces. We expect any result on manifolds which has a sufficiently geometrically realized proof can be generalized to metric spaces using curves in place of vectors. A metric space version of (I) is contained in [2], e.g.; and versions of (II) have been proven several times independently in e.g., [12], [2], and [5]—see Theorem 2.2 below. A version of (III) is the main result of this paper, Theorem 5.4: an involutive distribution on a complete metric space is integrable. Since the result is for complete metric spaces, it generalizes the classical result on Banach manifolds (proven, e.g., in [1]). Theorem 5.4 further generalizes the classical result by assuming only Lipschitz-type regularity instead of smoothness, which is of interest in, for example, control theory.

The simplest nontrivial foliations are the 1-dimensional foliations. A local flow gives a local foliation with 1-dimensional leaves—the integral curves' paths. In this sense (III) generalizes (II). A flow without equilibria gives a (global) foliation. The existence of a nontrivial foliation is not guaranteed by Frobenius' classical theorem on general manifolds M and depends on the global topology of M. For example, there is no 1-dimensional foliation of S^n for even numbers n, nor for any compact surface except the torus and the Klein bottle. Co-dimension 1 foliations are just as rare on finite dimensional compact manifolds, as Thurston proved for the C^∞ case. Similarly, most metric spaces do not support any nontrivial foliations. Most obviously, any discrete metric space can only support the trivial foliation of dimension 0. We give some examples of nontrivial foliations on metric spaces in Section 8, and show that new foliations of classical linear spaces are possible with the metric space generalization.

As far as I have been able to determine, this particular approach to the proof of Frobenius' classical theorem has not been vetted in the literature—though it uses basic, well-known ideas. We outline the approach in this paragraph, simplified to vector fields on a manifold. The terminology (and assumptions) will be clarified in the main body of the paper, and Figures 2 and 3 from Section 5 may aid intuition. The crux of the local Frobenius result in two dimensions is as follows: Given two transverse vector fields $f,g: M \to TM$ there exists an integral surface (tangent to linear combinations of f and g) through any point $x_0 \in M$ when the Lie bracket satisfies $[f,g] = af + bg$ for some choice of functions $a,b: M \to \mathbb{R}$ (involutivity of f and g). To prove this, define

$$S := \{F_t G_s(x_0) \in M : |s|, |t| < \delta\}$$

where F and G are the local flows of f and g guaranteed to exist by (II). Since f and g are transverse, we may choose $\delta > 0$ small enough for S to be a well-defined surface. S will be shown to be the desired integral surface through x_0. Notice S is tangent to f by construction, but it is not immediately clear S is tangent to $a'f + b'g$ for arbitrarily chosen $a', b' \in \mathbb{R}$. Notice, though, that by construction S is tangent to g at any point $x = G_s(x_0)$, and also S is tangent to $a''f + b''g$ at x for functions a'' and b''. Therefore establishing

$$(F_t)^* \left(a'f + b'g\right) = a''f + b''g \quad \text{at} \quad x = G_s(x_0) \tag{1.1}$$

for some functions a'' and b'', proves S is tangent to $a'F + b'G$ at an arbitrary point $z = F_t G_s(x_0) \in S$, since the push-forward $(F_t)_*$ and the pull-back $(F_t)^*$ are inverse to each other and preserve tangency since they are local bi-Lipschitz maps. Next, since the Lie bracket equals the Lie derivative,

$$\lim_{h\to 0} \frac{F_h^*(g) - g}{h} = [f,g] = af + bg$$

by involutivity so

$$F_h^*(g) = g + h(af + bg) + o(h) = \widetilde{a}f + \widetilde{b}g + o(h).$$

Using the fact that $F_h^*(f) = f$ for any h, and the linearity of pullback for fixed t, we have for functions a_i and $b_i : M \to \mathbb{R}$

$$F_{t/n}^*(a_i f + b_i g) = (a_{i+1} f + b_{i+1} g) + o(1/n)$$

for some functions a_{i+1} and b_{i+1}. Then since

$$F_t^* = \underbrace{F_{t/n}^* F_{t/n}^* \ldots F_{t/n}^*}_{n \text{ times}} = \left(F_{t/n}^*\right)^{(n)}$$

(where the superscript in round brackets denotes composition n times) we get (1.1) as follows:

$$\begin{aligned} F_t^*(a_0 f + b_0 g) &= \lim_{n\to\infty} \left(F_{t/n}^*\right)^{(n)} (a_0 f + b_0 g) \\ &= \lim_{n\to\infty} a_n f + b_n g + n o(1/n) = a_\infty f + b_\infty g + 0 \end{aligned}$$

completing the sketch for manifolds. Convergence is carefully demonstrated in the proof of the metric space version, Theorem 5.4.

A pivotal fact on which the metric space version relies is that arc fields which satisfy certain Lipschitz-type conditions generate unique local flows (proven in [5] and reviewed in Section 2). Also a natural linear structure may be associated with a metric space (though it has no *a priori* linear structure) using compositions of flows which faithfully generalizes the linearity of vector fields; this was introduced in [7]. We present this in Section 3 along with the generalization of the Lie bracket for vector fields which uses the well-known asymptotic characterization of the Lie bracket; i.e., successively follow the flows forward and backward for time $\sqrt{t}$. This investigation further clarifies for us the surprising fact noted in [13]: smoothness is not necessary to define a geometrically meaningful Lie bracket. In Section 4, the pull-back along a flow is shown to behave naturally with linearity and the bracket, which mimics properties of the Lie derivative on manifolds. Many more such algebraic properties are valid than are contained in these sections, but in this monograph we present only the minimum machinery directly relevant to proving Frobenius' Theorem in Section 5.

Section 6 applies this local Frobenius theorem in the attempt to construct foliations on spaces which support them, yielding a global theorem on metric spaces. A metric space generalization of the Nagumo-Brézis Invariance Theorem is proven, which is used to show integrable distributions are involutive. Another facet of the classical Global Frobenius Theorem guarantees local coordinates on which there exist coordinate vector fields tangent or perpendicular to an involutive distribution. In a general metric space, lacking an inner product, we need to substitute "transverse" for "perpendicular". We cannot coordinatize general metric spaces, but we can coordinatize the n-dimensional leaves of a foliation of a general metric space.

Section 7 proves a well-known result from Hamiltonian dynamics is also valid for metric spaces: two flows commute if and only if the bracket is 0. This is not exactly a corollary of the metric space Frobenius Theorem, but the proof is a mere simplification of that from Theorem 5.4.

Finally in Section 8 an almost trivial example applying these ideas has a result which astounded me: Any Lebesgue square-integrable function may be approximated using successive compositions of two elementary flows, starting from the constant zero function. In other words, $L^2(\mathbb{R})$ is controllable by two flows—domain translation and the continuous addition of a single Gaussian. This result has interesting consequences for function approximation. You may skip straight to this Example 8.3 after perusing the following review and the definitions in Section 3. [15] is an accessible text introducing the terminology of control theory with remarks and references on infinite dimensional controllability.

We cannot in good conscience conclude without repeating the well-known history that the classical "Frobenius Theorem" was originally proven by Deahna and Clebsch before Frobenius put the result in print—a fact Frobenius readily acknowledged.

2 Review of terminology and basic results

The proofs of all of the results from this section are contained in [5] for forward flows, also called semi-flows. Minimal changes, stated here, give us the corresponding results for (bidirectional) flows. The trivial changes to the proofs are detailed in the forthcoming monograph, [4].

A **metric space** (M,d) is a set of points M with a function $d : M \times M \to \mathbb{R}$ called the **metric** which has the following properties:

(i)	$d(x,y) \geq 0$	**positivity**
(ii)	$d(x,y) = 0$ *iff* $x = y$	**nondegeneracy**
(iii)	$d(x,y) = d(y,x)$	**symmetry**
(iv)	$d(x,y) \leq d(x,z) + d(z,y)$	**triangle inequality**

for all $x,y,z \in M$. The open ball of radius r about $x \in M$ is denoted by $B(x,r) := \{y : d(x,y) < r\}$. We assume throughout this paper that (M,d) is a locally complete metric space, i.e., there exists a complete neighborhood of each point in M. Denote the open ball in M about $x_0 \in M$ with radius r by

$$B(x_0,r) := \{x \in M : d(x,x_0) < r\}.$$

A map $f : (M,d_M) \to (N,d_N)$ between metric spaces is **Lipschitz** continuous if there exists $K_f \geq 0$ such that

$$d_N(f(x_1), f(x_2)) \leq K_f d_M(x_1,x_2)$$

for all $x_1, x_2 \in M$. A **bi-Lipschitz map** is an invertible Lipschitz map whose inverse is also Lipschitz (sometimes called a lipeomorphism, i.e., stronger than a homeomorphism, weaker than a diffeomorphism).

The following definition is made in analogy with vector fields on manifolds, where vectors are represented as curves on the manifold.

Definition 2.1. An **arc field** on a metric space M is a continuous map $X : M \times [-1,1] \to M$ with locally uniformly bounded speed, such that for all $x \in M$, $X(x,0) = x$.

Saying X has **locally uniformly bounded speed** means $X(x,\cdot) : [-1,1] \to M$ is Lipschitz, locally uniformly in x. Specifically we have

$$\rho(x) := \sup_{s \neq t} \frac{d(X(x,s), X(x,t))}{|s-t|} < \infty,$$

i.e., $X(x,\cdot)$ is Lipschitz, and the function $\rho(x)$ is locally bounded, meaning there exists $r > 0$ such that

$$\rho(x,r) := \sup\{\rho(y) \,|\, y \in B(x,r)\} < \infty.$$

An **integral curve** to X is a curve σ which is **tangent** to X throughout its domain, i.e., $\sigma : (\alpha,\beta) \to M$ for some open interval $(\alpha,\beta) \subset \mathbb{R}$ such that for each $t \in (\alpha,\beta)$

$$\lim_{h\to 0} \frac{d(\sigma(t+h), X(\sigma(t),h))}{h} = 0, \tag{2.1}$$

i.e.,

$$d(\sigma(t+h), X(\sigma(t),h)) = o(h).$$

The two variables for arc fields and flows which are usually denoted by x and t are often thought of as representing space and time. In this paper x, y, and z are used for space variables, while r, s, t, and h may fill the time variable slot. An arc field X will often have its variables migrate liberally between parentheses and subscripts

$$X(x,t) = X_x(t) = X_t(x)$$

depending on which variable we wish to emphasize in a calculation. We also use this convention for flows F defined below.

The following conditions guarantee existence and uniqueness of integral curves.

Condition E1: For each $x_0 \in M$, there are positive constants r, δ and Λ_X such that for all $x, y \in B(x_0, r)$ and $t \in (-\delta, \delta)$

$$d(X_t(x), X_t(y)) \leq d(x,y)(1 + |t|\Lambda_X).$$

Condition E2:

$$d(X_{s+t}(x), X_t(X_s(x))) = O(st)$$

as $st \to 0$ locally uniformly in x; in other words, for each $x_0 \in M$, there are positive constants r, δ and Ω_X such that for all $x \in B(x_0, r)$ and $s, t \in (-\delta, \delta)$

$$d(X_{s+t}(x), X_t(X_s(x))) \leq |st|\,\Omega_X.$$

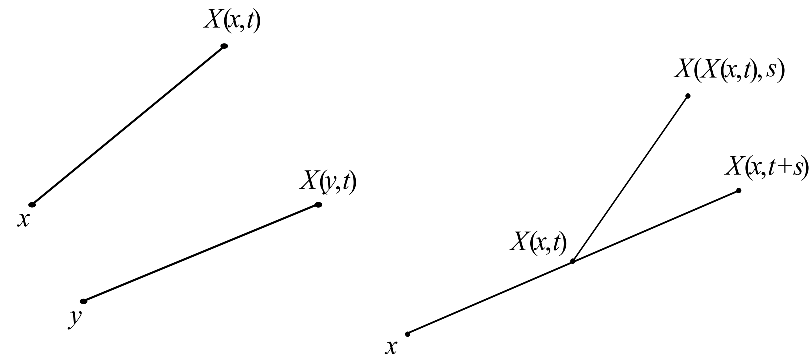


Figure 1. Conditions E1 and E2

Theorem 2.2. *Let X be an arc field satisfying E1 and E2 on a locally complete metric space M. Then given any point $x \in M$, there exists a unique integral curve $\sigma_x : (\alpha_x, \beta_x) \to M$ with $\sigma_x(0) = x$.*

Several remarks are in order. Here, x is called the **initial condition** for the integral curve σ_x in the above theorem. Uniqueness of integral curves means that for any $x \in M$, the curve $\sigma_x : (\alpha_x, \beta_x) \to M$ has maximal domain (α_x, β_x) in the sense that for any other integral curve $\widehat{\sigma}_x : \left(\widehat{\alpha}_x, \widehat{\beta}_x\right) \to M$ also having initial condition x, we have $\left(\widehat{\alpha}_x, \widehat{\beta}_x\right) \subset (\alpha_x, \beta_x)$ and $\widehat{\sigma}_x = \sigma_x|_{\left(\widehat{\alpha}_x, \widehat{\beta}_x\right)}$ (i.e., σ_x is the **maximal integral curve**).

The proof of Theorem 2.2 is constructive and shows the **Euler curves** $X_{t/n}^{(n)}(x)$ converge to the integral curve. Here we are using $f^{(n)}$ to denote the composition of a map $f : M \to M$ with itself n times so

$$X_{\frac{t}{n}}^{(n)}(x) = \underbrace{X_{\frac{t}{n}} \circ X_{\frac{t}{n}} \circ \ldots \circ X_{\frac{t}{n}}}_{n \text{ times}}(x)$$

and we have

$$\lim_{n\to\infty} X_{\frac{t}{n}}^{(n)}(x) = \sigma_x(t).$$

for suitably small $|t|$. Other, slightly different formulations of Euler curves also lead to the same result, σ, under Conditions E1 and E2, e.g.,

$$\xi_n(t) := X_{t-i\cdot 2^{-n}} X_{2^{-n}}^{(i)}(x) \quad \text{for} \quad i \cdot 2^{-n} \le t \le (i+1)\, 2^{-n}$$

also has

$$\lim_{n\to\infty} \xi_n(t) = \sigma_x(t)$$

for suitably small $|t|$.

Theorem 2.2 and those that follow are true under more general conditions outlined in [5], [12], and [8]; see also [2]. But throughout this paper, E1 and E2 are satisfied and are easier to use.

Example 2.3. A **Banach space** $(M, \|\cdot\|)$ is a normed vector space, complete in its norm (e.g., $\mathbb{R}^n$ with Euclidean norm). A Banach space is an example of a metric space with $d(u,v) := \|u - v\|$. A **vector field** on a Banach space M is a map $f : M \to M$. An **integral curve** to a vector field f with **initial condition** x is a curve $\sigma_x : (\alpha, \beta) \to M$ defined on an open interval $(\alpha, \beta) \subset \mathbb{R}$ containing 0 such that $\sigma_x(0) = x$ and $\sigma_x'(t) = f(\sigma_x(t))$ for all $t \in (\alpha, \beta)$. The classical Picard-Lindelöf Theorem guarantees unique integral curves for any locally Lipschitz f. With a few tricks, most differential equations can be represented as vector fields on a suitably abstract space.

Every Lipschitz vector field $f : M \to M$ gives rise to an arc field $X(x,t) := x + tf(x)$ and it is easy to check X satisfies E1 and E2 (cf. [5]). Further the integral curves to the arc field are exactly the integral curves to the vector field. Therefore Theorem 2.2 is a generalization of the classical Picard-Lindelöf Theorem.

Remark 2.4. Of prime import for this paper, the proof of Theorem 2.2 actually shows integral curves are **locally uniformly tangent** to X:

$$d(X_x(t), \sigma_x(t)) = o(t)$$

locally uniformly for $x \in M$, i.e., for each $x_0 \in M$ there exists a constant $r > 0$ such that for any $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in B(x_0, r)$

$$\frac{d(X_x(t), \sigma_x(t))}{|t|} < \varepsilon$$

whenever $0 < |t| < \delta$.

More than this, the proof also shows integral curves are tangent uniformly for all arc fields X which satisfy $E1$ and $E2$ for specified Λ and Ω—though this result is not used in this paper.

Similarly, two arc fields X and Y are **locally uniformly tangent**, denoted $X \sim Y$, if

$$d(X_x(t), Y_x(t)) = o(t)$$

locally uniformly for $x \in M$. It is easy to check $\sim$ is an equivalence relation. E.g., transitivity follows from the triangle inequality:

$$\frac{d(X_t(x), Z_t(x))}{|t|} \leq \frac{d(X_t(x), Y_t(x))}{|t|} + \frac{d(Y_t(x), Z_t(x))}{|t|}.$$

We use the symbol $\sim$ in many contexts in this paper (particularly Section 6), but there is always a local-uniform-tangency property associated with it.

Corollary 2.5. *Assume the conditions of Theorem 2.2 and let $s \in (\alpha_x, \beta_x)$ and $y = \sigma_x(s)$. Then $\alpha_y = \alpha_x - s$ and $\beta_y = \beta_x - s$ so*

$$(\alpha_y, \beta_y) = \left(\alpha_{\sigma_x(s)}, \beta_{\sigma_x(s)}\right) = \{t : \alpha_x - s < t < \beta_x - s\}.$$

Thus $t \in (\alpha_y, \beta_y)$ if and only if $t + s \in (\alpha_x, \beta_x)$, and then we have

$$\sigma_{\sigma_x(s)}(t) = \sigma_x(s+t).$$

Defining $W \subset M \times \mathbb{R}$ by

$$W := \{(x,t) \in M \times \mathbb{R} : t \in (\alpha_x, \beta_x)\}$$

and $F : W \to M$ by $F(x,t) := \sigma_x(t)$ we have:

(i) *$M \times \{0\} \subset W$ and $F(x,0) = x$ for all $x \in M$.*
(ii) *For each (fixed) $x \in M$, $F(x,\cdot) : (\alpha_x, \beta_x) \to M$ is the maximal integral curve σ_x to X.*
(iii) *$F(F(x,s),t) = F(x,s+t)$.*

F is called the **local flow** generated by the arc field X. Compare Condition E2 with property (iii) above to see why an arc field might be thought of as a "pre-flow".

Theorem 2.6. *Let $\sigma_x : (\alpha_x, \beta_x) \to M$ and $\sigma_y : (\alpha_y, \beta_y) \to M$ be two integral curves to an arc field X. Assume $(\alpha_x, \beta_x) \cap (\alpha_y, \beta_y) \supset I$ for some interval I, and assume the inequality of Condition E1 holds for a single value Λ_X on a set containing*

$$\{\sigma_x(t) : t \in I\} \cup \{\sigma_y(t) : t \in I\}.$$

Then

$$d(\sigma_x(t), \sigma_y(t)) \leq e^{\Lambda_X |t|} d(x,y) \text{ for all } t \in I,$$

i.e.,

$$d(F_t(x), F_t(y)) \leq e^{\Lambda_X |t|} d(x,y). \tag{2.2}$$

Theorem 2.7. *For F and W as above, W is open in $M \times \mathbb{R}$ and F is continuous on W.*

For fixed t it is clear F_t is a local bi-Lipschitz map, when defined, by Theorem 2.6. Compare Condition E1 with line (2.2) to see why E1 may be thought of as a local linearity property for X, needed for the continuity of F.

Definition 2.8. An arc field X on a metric space M is said to have **linear speed growth** if there is a point $x \in M$ and positive constants c_1 and c_2 such that for all $r > 0$

$$\rho(x,r) \leq c_1 r + c_2, \tag{2.3}$$

where $\rho(x,r)$ is the local bound on speed given in Definition 2.1.

Theorem 2.9. *Let X be an arc field on a complete metric space M, which satisfies E1 and E2 and has linear speed growth. Then F is a (full)* ***flow*** *with domain $W = M \times \mathbb{R}$.*

Example 2.10. Every local flow on a metric space is generated by an arc field. Any local flow F gives rise to an arc field $\overline{F} : M \times [-1,1] \to M$ defined by

$$\overline{F}(x,t) := \begin{cases} F(x,t) & \text{if } t \in \left(\frac{\alpha_x}{2}, \frac{\beta_x}{2}\right) \\ F\left(x, \frac{\alpha_x}{2}\right) & \text{if } t \in \left[-1, \frac{\alpha_x}{2}\right] \\ F\left(x, \frac{\beta_x}{2}\right) & \text{if } t \in \left[\frac{\beta_x}{2}, 1\right]. \end{cases}$$

(The issue here is that F, being a *local* flow, may have α_x or $\beta_x < 1$.) Clearly the local flow generated by $\overline{F}$ is F. Since all our concerns with arc fields are local, we will never focus on $t \notin \left(\frac{\alpha_x}{2}, \frac{\beta_x}{2}\right)$ and henceforth we will not notationally distinguish between $\overline{F}$ and F as arc fields.

With this identification of flows being arc fields (but not necessarily *vice-versa*) we may simplify Remark 2.4 to: $X \sim F$ if X satisfies E1 and E2.

3 The bracket and linearity

To simplify notation we drop parentheses for expressions such as $Y_t \circ X_s(x) = Y_t(X_s(x))$ and write $Y_t X_s(x)$ since the composition of arbitrary maps is associative.

Definition 3.1. The **bracket** of two arc fields X and Y is the map $[X,Y] : M \times [-1,1] \to M$ with

$$[X,Y](x,t) := \begin{cases} Y_{-\sqrt{t}} X_{-\sqrt{t}} Y_{\sqrt{t}} X_{\sqrt{t}}(x) & \text{for } t \geq 0 \\ X_{-\sqrt{|t|}} Y_{-\sqrt{|t|}} X_{\sqrt{|t|}} Y_{\sqrt{|t|}}(x) & \text{for } t < 0. \end{cases} \tag{3.1}$$

There are many different equivalent characterizations of the Lie bracket on a manifold. (3.1) uses the obvious choice of the **asymptotic** characterization to generalize the concept to metric spaces. $[X,Y](x,t)$ traces out a small "parallelogram" in M starting at x, which hopefully almost returns to x. The bracket measures the failure of X and Y to commute as will be made clear in Theorems 7.1 and 5.4.

Definition 3.2. We say X **&** Y **verge** if

$$d(Y_s X_t(x), X_t Y_s(x)) = O(|st|)$$

locally uniformly in x, i.e., for each $x_0 \in M$ there exists a ball $B(x_0, r)$ for some $r > 0$ for which there exist positive constants C_{XY} and δ such that

$$d(Y_s X_t(x), X_t Y_s(x)) \leq C_{XY} |st|$$

for all $|s|, |t| < \delta$ for all $x \in B(x_0, r)$.

Lemma 3.3. *If X & Y verge and satisfy E1 and E2 then*

$$d(Y_{-t} X_{-t} Y_t X_t(x), x) = O(t^2)$$

locally uniformly for $x \in M$.

Proof.

$$\begin{aligned}
& d(Y_{-s} X_{-t} Y_s X_t(x), x) \\
& \leq d(Y_{-s} X_{-t} Y_s X_t(x), Y_{-s} X_{-t} X_t Y_s(x)) + d(Y_{-s} X_{-t} X_t Y_s(x), Y_{-s} Y_s(x)) + d(Y_{-s} Y_s(x), x) \\
& \leq d(Y_s X_t(x), X_t Y_s(x))(1 + |s|\Lambda_Y)(1 + |t|\Lambda_X) + t^2 \Omega_X (1 + |s|\Lambda_Y) + s^2 \Omega_Y \\
& \leq C_{XY} |st| (1 + |s|\Lambda_Y)(1 + |t|\Lambda_X) + t^2 \Omega_X (1 + |s|\Lambda_Y) + s^2 \Omega_Y \leq C\left(|st| + t^2 + s^2\right)
\end{aligned}$$

where

$$C := \max\{C_{XY}(1 + \Lambda_Y)(1 + \Lambda_X), \Omega_X(1 + \Lambda_Y), \Omega_Y\}.$$

Letting $s = t$ gives the result. ☐

Proposition 3.4. *If X & Y verge and satisfy E1 and E2 then $[X,Y]$ is an arc field.*

Proof. We establish the local bound on speed. The purpose of Lemma 3.3 is to give $d([X,Y](x,t),x) = O(t)$ for $t \geq 0$. Similarly, for $t < 0$

$$\begin{aligned} & d(X_tY_tX_{-t}Y_{-t}(x),x) \\ & \leq d(X_tY_tX_{-t}Y_{-t}(x),X_tX_{-t}(x)) + d(X_tX_{-t}(x),x) \\ & \leq d(Y_tX_{-t}Y_{-t}(x),X_{-t}(x))(1+|t|\Lambda_X) + t^2\Omega_X \end{aligned}$$

which, using this trick again, gives

$$\begin{aligned} & \leq d(X_{-t}Y_{-t}(x),Y_{-t}X_{-t}(x))(1+|t|\Lambda_X)(1+|t|\Lambda_Y) \\ & + t^2\Omega_Y(1+|t|\Lambda_Y) + t^2\Omega_X = O(t^2) \text{ since } X \,\&\, Y \text{ verge.} \end{aligned}$$

Therefore

$$d([X,Y]_t(x),x) = O(t)$$

for both positive and negative t. Then since $\sqrt{|t|}$ is Lipschitz except at $t = 0$ we see $[X,Y]$ has bounded speed. □

Example 3.5. As in Example 2.3 let $f,g : B \to B$ be Lipschitz vector fields on a Banach space B, and let X and Y be their corresponding arc fields

$$\begin{aligned} X(x,t) &:= x + tf(x) \\ Y(x,t) &:= x + tg(x) \end{aligned}$$

It is easy to check X & Y verge:

$$\begin{aligned} & d(Y_sX_t(x),X_tY_s(x)) \\ & = \|x+tf(x)+sg(x+tf(x)) - [x+sg(x)+tf(x+sg(x))]\| \\ & \leq |t|\,\|f(x)-f(x+sg(x))\| + |s|\,\|g(x+tf(x))-g(x)\| \\ & \leq |t|K_f\,\|x-(x+sg(x))\| + |s|K_g\,\|x+tf(x)-x\| \\ & \leq |st|(K_f\,\|g(x)\| + K_g\,\|f(x)\|) \end{aligned}$$

so $C_{XY} := (K_f\,\|g(x)\| + K_g\,\|f(x)\|)$.

Therefore, even though the vector fields may not be smooth, so their Lie bracket is undefined, their metric space bracket is meaningful and will give us geometric information as we shall see in Theorem 5.4.

Definition 3.6. If X and Y are arc fields on M then define $X+Y$ to be the arc field on M given by

$$(X+Y)_t(x) := Y_tX_t(x).$$

For any function $a : M \to \mathbb{R}$ define the arc field aX by

$$aX(x,t) := X(x,a(x)t). \tag{3.2}$$

If a is Lipschitz, then aX is an arc field.

To be fastidiously precise we need to define $aX_x(t)$ for all $t \in [-1,1]$ so technically we must specify

$$aX(x,t) := \begin{cases} \left.\begin{array}{cc} X(x, a(x)t) & -\frac{1}{|a(x)|} \le t \le \frac{1}{|a(x)|} \\ X(x,1) & t > 1/|a(x)| \\ X(x,-1) & t < -1/|a(x)| \end{array}\right\} & \text{when } a(x) \neq 0 \\ \quad x \qquad\quad \text{for } -1 \le t \le 1 & \text{if } a(x) = 0 \end{cases} \tag{3.3}$$

using the trick from Example 2.10. Again, we will not burden ourselves with this detail; in all cases our concern with the properties of an arc field $X_x(t)$ is only near $t = 0$.

It is a simple definition check to prove aX is an arc field when a is Lipschitz, since $aX_x(t) = X_x(a(x)t)$ is Lipschitz in t if $X_x(t)$ is: assuming $a(x) \neq 0$,

$$\begin{aligned} \rho_{aX}(x) &:= \sup_{s \neq t} \frac{d(X_x(a(x)s), X_x(a(x)t))}{|s-t|} = \sup_{s \neq t} \frac{d(X_x(s), X_x(t))}{\left|\frac{s}{a(x)} - \frac{t}{a(x)}\right|} \\ &= a(x) \sup_{s \neq t} \frac{d(X_x(s), X_x(t))}{|s-t|} = a(x)\rho_X(x) \end{aligned}$$

so

$$\begin{aligned} \rho_{aX}(x,r) &:= \sup_{y \in B(x,r)} \{\rho_{aX}(y)\} = \sup_{y \in B(x,r)} \{a(y)\rho(y)\} \\ &\le (a(x) + rK_a)\rho_X(x,r) < \infty. \end{aligned}$$

Now we have the beginnings of a linear structure associated with M. For instance, expressions such as $X - Y$ make sense:

$$X - Y := X + (-1)Y$$

where -1 is a constant function on M. Further, 0 is an arc field defined as the constant map

$$0(x,t) := x$$

and satisfies $0 + X = X = X + 0$ for any X. Notice from the definition, we have $[X,Y] = -[Y,X]$. Another trivial definition check shows this multiplication is associative and commutative:

$$(a \cdot b)X = a(bX) \qquad \text{and} \qquad (a \cdot b)X = (b \cdot a)X$$

where $\cdot$ denotes multiplication of functions.

Proposition 3.7. *Assume X & Y verge and satisfy E1 and E2. Then their sum $X + Y$ satisfies E1 and E2.*

Proof. Checking Condition E1:

$$\begin{aligned} & d((X+Y)_t(x), (X+Y)_t(y)) \\ & = d(Y_t X_t(x), Y_t X_t(y)) \le d(X_t(x), X_t(y))(1 + |t|\Lambda_Y) \\ & \le d(x,y)(1 + |t|\Lambda_X)(1 + |t|\Lambda_Y) \le d(x,y)\left(1 + |t|(\Lambda_X + \Lambda_Y) + t^2\Lambda_X\Lambda_Y\right) \\ & \le d(x,y)(1 + |t|\Lambda_{X+Y}) \end{aligned}$$

where $\Lambda_{X+Y} := \Lambda_X + \Lambda_Y + \Lambda_X \Lambda_Y < \infty$.

Condition E2:

$$\begin{aligned}
&d\left((X+Y)_{s+t}(x), (X+Y)_t (X+Y)_s (x)\right) \\
&= d\left(Y_{s+t}X_{s+t}(x), Y_t X_t Y_s X_s (x)\right) \\
&\leq d\left(Y_{s+t}X_{s+t}(x), Y_t Y_s X_{s+t}(x)\right) + d\left(Y_t Y_s X_{s+t}(x), Y_t X_t Y_s X_s (x)\right) \\
&\leq |st|\,\Omega_Y + d\left(Y_s X_{s+t}(x), X_t Y_s X_s (x)\right)(1+|t|\,\Lambda_Y) \\
&\leq |st|\,\Omega_X + \left[d\left(Y_s X_{s+t}(x), Y_s X_t X_s (x)\right) + d\left(Y_s X_t (y), X_t Y_s (y)\right)\right](1+t\Lambda_X)
\end{aligned} \tag{3.4}$$

where $y := X_s(x)$. Notice

$$\begin{aligned}
d\left(Y_s X_{s+t}(x), Y_s X_t X_s (x)\right) &\leq d\left(X_{s+t}(x), X_t X_s (x)\right)(1+|s|\,\Lambda_Y) \\
&\leq |st|\,\Omega_X (1+|s|\,\Lambda_Y) = O(|st|)
\end{aligned}$$

and the last summand of (3.4)is also $O(|st|)$ since X & Y verge, so E2 is satisfied. □

So in this case, the flow H generated by $X+Y$ is computable with Euler curves as

$$H(x,t) = \lim_{n\to\infty} (X+Y)^{(n)}_{t/n}(x) = \lim_{n\to\infty} \left(Y_{t/n}X_{t/n}\right)^{(n)}(x). \tag{3.5}$$

Therefore, this definition of $X+Y$ using compositions is a direct generalization of the concept of adding vector fields on a differentiable manifold (see [1, Section 4.1A]). One of the inspirations for this paper, [7] introduced the sum of semigroups on a metric space in the same spirit as defined here, with commensurable conditions.

When X & Y verge and satisfy E1 and E2, we also have $(X+Y) \sim (Y+X)$ since

$$\left(Y_{t/n}X_{t/n}\right)^{(n)} = Y_{t/n}\left(X_{t/n}Y_{t/n}\right)^{(n-1)} X_{t/n}$$

whence both arc fields $X+Y$ and $Y+X$ are (locally uniformly) tangent to the flow H using (3.5).

Proposition 3.8. *If X satisfies E1 and E2 and $a : M \to \mathbb{R}$ is a Lipschitz function, then aX satisfies E1 and E2.*

Proof. E1:

$$\begin{aligned}
&d\left(aX_x(t), aX_y(t)\right) \\
&= d\left(X_x(a(x)t), X_y(a(y)t)\right) \\
&\leq d\left(X_x(a(x)t), X_x(a(y)t)\right) + d\left(X_x(a(y)t), X_y(a(y)t)\right) \\
&\leq |a(x)t - a(y)t|\,\rho(x) + d(x,y)(1+a(y)\,|t|\,\Lambda_X) \\
&\leq d(x,y)\left(K_a\,|t|\,\rho(x) + 1 + a(y)\,|t|\,\Lambda_X\right) = d(x,y)(1+|t|\,\Lambda_{aX})
\end{aligned}$$

where $\Lambda_{aX} := K_a \rho(x) + a(y)\Lambda_X < \infty$.

E2: For all $x_0 \in M$ and $\delta > 0$ we know a is bounded by some $A > 0$ on $B(x_0, \delta)$ since a is Lipschitz.

$$
\begin{aligned}
& d\left(aX_x(s+t), aX_{aX_x(s)}(t)\right) \\
& = d\left(X_x(a(x)(s+t)), X_{X_x(a(x)s)}(a(X_x(a(x)s))t)\right) \\
& \leq d\left(X_x(a(x)(s+t)), X_{X_x(a(x)s)}(a(x)t)\right) \\
& + d\left(X_{X_x(a(x)s)}(a(x)t), X_{X_x(a(x)s)}(a(X_x(a(x)s))t)\right) \\
& \leq a(x)|s| \cdot a(x)|t|\Omega_X + \rho \cdot |a(x)t - a(X_x(a(x)s))t| \\
& \leq |st|[a(x)]^2\Omega_X + |t|\rho K_a d(x, X_x(a(x)s)) \\
& \leq |st|[a(x)]^2\Omega_X + |st|\rho^2 K_a a(x) \leq |st|\Omega_{aX}
\end{aligned}
$$

where $\Omega_{aX} := A^2\Omega_X + \rho^2 K_a A$. □

Combining these results gives

Theorem 3.9. *If a and b are locally Lipschitz functions and X & Y verge and satisfy E1 and E2, then $aX + bY$ is an arc field which satisfies E1 and E2 and so has a unique local flow.*

If in addition a and b are globally Lipschitz and X and Y have linear speed growth, then $aX + bY$ generates a unique flow.

Proof. We haven't proven aX and bY verge, but this is a straightforward definition check, as is the fact that $aX + bY$ has linear speed growth. □

Local flows have the following useful linearity property:

Proposition 3.10. *If F is a local flow then interpreting F as an arc field we can perform the following operations:*

1. *if a and b are constant then* $aF + bF = (a+b)F$
2. *if a and b are real functions then* $(aF + bF)_t(x) = (a + b \circ (aF)_t)F_t(x)$.

Proof. This is another obvious definition check:

$$
\begin{aligned}
2. \quad (aF + bF)_t(x) &= (bF)_t(aF)_t(x) = F_{b((aF)_t(x))t}F_{a(x)t}(x) \\
&= F_{(a(x)+(b\circ(aF)_t)(x))t}(x) = (a + b \circ (aF)_t)F_t(x)
\end{aligned}
$$

and 1. follows from 2. □

4 Contravariance

If $\phi : M_1 \to M_2$ is a bi-Lipschitz map, then the pull-back of an arc field X on M_2 is the arc field ϕ^*X on M_1 given by

$$\phi^*X(x,t) := \phi^{-1}(X(\phi(x), t))$$

or in other notation,

$$(\phi^*X)_t(x) = \phi^{-1}X_t\phi(x)$$

which is a direct analog of the pull-back of a vector field on a manifold using curves to represent vectors. The definition for flows is identical, replacing X with F. The pull-back to M_1 of an integral curve σ to an arc field on M_2 is analogous:

$$(\phi^*\sigma)_x(t) := \phi^{-1}\left(\sigma_{\phi(x)}(t)\right).$$

The pull-back of a function $a : M_2 \to \mathbb{R}$ is the function $\phi^* a : M_1 \to \mathbb{R}$ defined as $\phi^* a(x) := a(\phi(x))$.

Proposition 4.1. *If $\phi : M_1 \to M_2$ is a bi-Lipschitz map and the arc field X on M_2 has unique integral curves then ϕ^*X has unique integral curves. The integral curves to ϕ^*X are the pull-backs of integral curves to X.*

Proof. This is obvious: if F is a local flow for X then

$$\begin{aligned}
&d\left(\phi^*X\left(\phi^*F(x,t),s\right),\phi^*F(x,t+s)\right)\\
&=d\left(\phi^{-1}X\left[\phi\phi^{-1}F(\phi(x),t),s\right],\phi^{-1}F(\phi(x),t+s)\right)\\
&=d\left(\phi^{-1}X\left[F(\phi(x),t),s\right],\phi^{-1}F(\phi(x),t+s)\right)\\
&\le K_\phi d\left(X\left[F(\phi(x),t),s\right],F(\phi(x),t+s)\right)=K_\phi o(s)=o(s)
\end{aligned}$$

so ϕ^*F is a local flow for ϕ^*X.

Similarly if σ is an integral curve to ϕ^*X then $\left(\phi^{-1}\right)^*\sigma$ is an integral curve to X so by uniqueness there can be only one such σ. □

The push-forward of any function, curve or flow is defined similarly, e.g.,

$$\phi_*F(x,t) := \phi\left(F\left(\phi^{-1}(x),t\right)\right).$$

It is easy to check push-forward is covariant (i.e., $(\phi\circ\psi)_* = \phi_*\circ\psi_*$) and pull-back is contravariant (i.e., $(\phi\circ\psi)^* = \psi^*\circ\phi^*$). It is also clear that push-forward and pull-back are inverse operations and Proposition 4.1 holds *mutatis mutandis* for push-forward in place of pull-back.

Proposition 4.2 (Linearity of Pull-back)**.** *If X and Y are arc fields on M and $\phi : M_1 \to M_2$ is a bi-Lipschitz map, then*
(i) $\phi^*(X+Y) = \phi^*(X) + \phi^*(Y)$
(ii) $\phi^*(aX) = (a\circ\phi)\,\phi^*(X) = \phi^*(a)\,\phi^*(X)$.

Proof. Trivial definition check. □

Since the pull-back and linearity are established for arc fields, we can now explore another characterization of the bracket. In the context of M being a smooth manifold, let F and G be local flows generated by smooth vector fields f and g. There it is well known the following "dynamic" characterization of the Lie bracket is equivalent to the asymptotic characterization

$$[f,g] = \frac{d}{dt}(F_t)^*g\bigg|_{t=0}. \tag{4.1}$$

Using

$$\left.\frac{d}{dt}(F_t)^* g\right|_{t=0} = \lim_{t\to 0}\frac{(F_t)^* g - g}{t} = [f,g]$$

for inspiration, we return to the context of metric spaces where, with F and G viewed as arc fields, their bracket $[F,G]$ is defined. Formally we have

$$F_t^* G_t(x) = (t[F,G]+G)_t(x) \qquad \text{for } t \geq 0 \text{ and} \tag{4.2}$$

$$F_s^* G_s(x) = (-s[-F,-G]-G)_{-s}(x) \qquad \text{for } s < 0 \tag{4.3}$$

which hold because

$$\begin{aligned}(t[F,G]+G)_t(x) &= G_t[F,G]_{t^2}(x)\\ = G_t G_{-t} F_{-t} G_t F_t(x) &= F_{-t} G_t F_t(x) = F_t^* G_t(x)\end{aligned}$$

and

$$\begin{aligned}&(-s[-F,-G]-G)_{-s}(x)\\ &= G_s[-F,-G]_{s^2}(x) = G_s(-G)_{-|s|}(-F)_{-|s|}(-G)_{|s|}(-F)_{|s|}(x)\\ &= G_s G_{|s|} F_{|s|} G_{-|s|} F_{-|s|}(x) = F_{-s} G_s F_s(x) = F_s^* G_s(x).\end{aligned}$$

These facts will be used in the heart of the proof of our main result, Theorem 5.4, as will the following

Proposition 4.3. $(F_s)^* X \sim X$.

Proof. Using the properties of flows $F_t = F_{-s+t+s} = F_{-s} F_t F_s$ and $F_t^{-1} = F_t$ we get

$$\begin{aligned}&d\left(\left((F_s)^* X\right)_t(x), X_t(x)\right)\\ &\leq d\left(F_{-s} X_t F_s(x), F_{-s} F_t F_s(x)\right) + d\left(F_t(x), X_t(x)\right)\\ &\leq e^{s\Lambda_X} d\left(X_t(y), F_t(y)\right) + o(t) = o(t)\end{aligned}$$

where $y := F_s(x)$ and the exponential comes from Theorem 2.6. □

5 Local Frobenius Theorem

Definition 5.1. Two arc fields X and Y are (locally uniformly) **transverse** if for each $x_0 \in M$ there exists a $\delta > 0$ such that

$$d(X_s(x), Y_t(x)) \geq \delta(|s|+|t|)$$

for $|t| < \delta$ for all $x \in B(x_0, \delta)$.

Example 5.2. On the plane $\mathbb{R}^2$ with Euclidean norm $\|\cdot\|$ any two linearly independent vectors $u, v \in \mathbb{R}^2$ give us the transverse arc fields

$$X_t(x) := x + tu \qquad \text{and} \qquad Y_t(x) := x + tv.$$

To check this, it is easiest to define a new norm on $\mathbb{R}^2$ by

$$\|x\|_{uv} := |x_1| + |x_2|$$

where $x = x_1 u + x_2 v$ and $x_1, x_2 \in \mathbb{R}$. Since all norms on $\mathbb{R}^2$ are metrically equivalent there must exist a constant $C > 0$ such that $\|x\|_{uv} \leq C \|x\|$ for all $x \in \mathbb{R}^2$. Then taking $\delta := \frac{1}{C}$

$$d(X_s(x), Y_t(x)) = \|su - tv\| \geq \delta \|su - tv\|_{uv} = \delta(|s| + |t|).$$

A localization argument shows any pair of continuous vector fields f and g on a differentiable manifold give transverse arc fields if f and g are non-colinear at each point.

A (2-dimensional) **surface** is a 2-dimensional topological manifold, i.e., locally homeomorphic to $\mathbb{R}^2$.

For any subset $N \subset M$ and element $x \in M$ the **distance** from x to N is defined as

$$d(x, N) := \inf \{d(x, y) : y \in N\}.$$

This function d is not a metric, obviously, but it does satisfy a sort of triangle inequality:

$$d(x, N) \leq d(x, y) + d(y, N)$$

for all $x, y \in M$.

Definition 5.3. A surface $S \subset M$ is an **integral surface** of two arc fields X and Y if for any Lipschitz functions $a, b : M \to \mathbb{R}$ then S is **locally uniformly tangent** to $aX + bY$ for $x \in S$, i.e.,

$$d((aX + bY)_t(x), S) = o(t)$$

locally uniformly in x. Locally uniform tangency is denoted $S \sim aX + bY$.

Theorem 5.4. *Assume X & Y verge, are transverse, and satisfy E1 and E2 on a locally complete metric space M. Let F and G be the local flows of X and Y. If $[F, G] \sim aX + bY$ (locally uniform tangency) for some Lipschitz functions $a, b : M \to \mathbb{R}$, then for each $x_0 \in M$ there exists an integral surface S through x_0.*

Proof. It may be beneficial to review the outline of this proof from the third paragraph of the introduction. The metric space constructs of the previous sections will now be inserted into the manifold outline. A rigorous verification of the analytic estimates requires some tedious, but straightforward, calculations detailed here.

Define

$$S := \{F_t G_s(x_0) : |s|, |t| < \delta\}$$

where $\delta > 0$ is chosen small enough for S to be a well-defined surface (Figure 2). I.e., $F_{t_1} G_{s_1}(x_0) = F_{t_2} G_{s_2}(x_0)$ implies $t_1 = t_2$ and $s_1 = s_2$ so

$$\phi : (-\delta, \delta) \times (-\delta, \delta) \subset \mathbb{R}^2 \to S \subset M$$

defined by $\phi(s, t) := F_t G_s(x_0)$ is a homeomorphism. Finding such a δ is possible since X and Y are transverse. To see this, assume the contrary. Then there are different choices of s_i

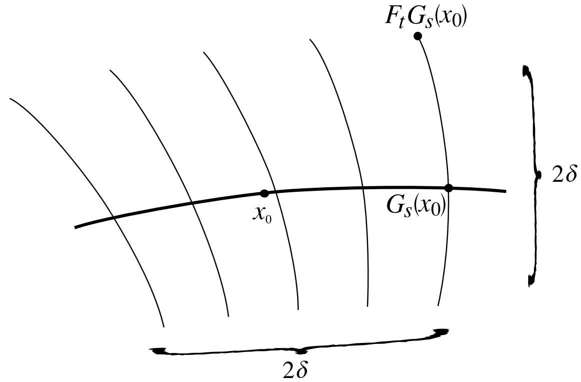


Figure 2. integral surface S

and t_i which give $F_{t_1}G_{s_1}(x_0) = F_{t_2}G_{s_2}(x_0)$ which implies $G_{s_1}(x_0) = F_{t_3}G_{s_2}(x_0)$ and letting $y := G_{s_2}(x_0)$ we must also then have

$$F_t(y) = G_s(y). \tag{5.1}$$

If this contrary assumption were true, then for all $\varepsilon > 0$ there would exist s and t with $|s|,|t| < \varepsilon$ such that (5.1) holds. Since X and Y are transverse, this cannot be so.

We will show S is the desired integral surface through x_0. Assume δ is also chosen small enough so throughout S the functions $|a|$ and $|b|$ are bounded, while the constants Λ, Ω, and ρ for X and Y hold uniformly, and that the closure of $B(x, 2\delta(\rho+1))$ is complete. This is possible because F and G have locally bounded speeds, since X and Y do.

Notice $S \sim X$ by construction, but it is not immediately clear $S \sim a'X + b'Y$ for arbitrarily chosen $a', b' \in \mathbb{R}$. Notice we can use

$$a'X + b'Y \sim a'F + b'G \sim b'G + a'F \sim b'Y + a'X$$

and so we will show $S \sim a'F + b'G$. We need to show this is true for an arbitrary point $z \in S$, so assume $z := F_tG_s(x_0)$ for some s and $t \in \mathbb{R}$. Notice by the construction of S we have $S \sim a''F + b''G$ at $x := G_s(x_0)$ for an arbitrary choice of Lipschitz functions a'' and b'' since $a''F + b''G \sim b''G + a''F$ and

$$\begin{aligned}
&\left(b''G + a''F\right)_h(x) \\
&= F_{a''\left(G_{b''(x)h}(x)\right)h} G_{b''(x)h}(x) = F_{a''\left(G_{b''(x)h}(x)\right)h} G_{b''(x)h}(x) \\
&= F_{a''\left(G_{b''(x)h}(x)\right)h} G_{b''(x)h} G_s(x_0) \in S
\end{aligned}$$

when h is small.

(x_0, x, z, s and t are now fixed for the remainder of the proof; however, we only explicitly check the case $t > 0$, indicating the changes where needed to check the $t < 0$ case.)

If we prove

$$(F_t)^*\left(a'F + b'G\right) \sim S \qquad \text{at} \qquad x = G_s(x_0) \tag{5.2}$$

this will prove $S \sim a'F + b'G$ at z, since the push-forward $(F_t)_*$ and the pull-back $(F_t)^*$ are inverse and local bi-Lipschitz maps and so preserve tangency. (See Figure 3.)

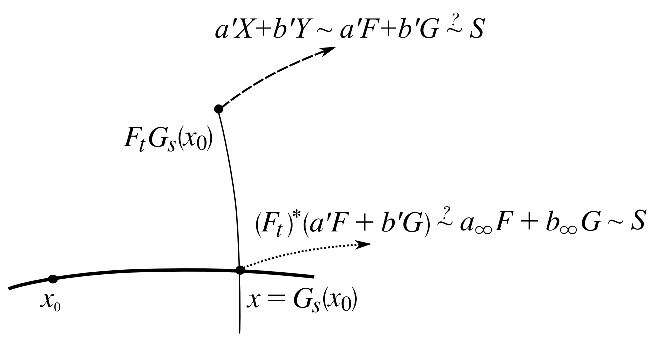


Figure 3. pull-back to $G_s(x_0)$

Restating (4.2):

$$F_t^* G_t(x) = (t[F,G]+G)_t(x)$$

so

$$F_{t/n}^* G_{t/n}(x) = \left(\tfrac{t}{n}[F,G]+G\right)_{t/n}(x) \tag{5.3}$$

for our previously fixed small $t \geq 0$ and arbitrary positive integer $n \in \mathbb{N}$. (For $t < 0$ use (4.3) instead.) For any arc fields Z and $\overline{Z}$ clearly

$$\begin{gathered} d\left(Z_t(x), \overline{Z}_t(x)\right) = o(t) \qquad \text{implies} \\ d\left((tZ)_t(x), (t\overline{Z})_t(x)\right) = d\left((Z)_{t^2}(x), (\overline{Z})_{t^2}(x)\right) = o\left(t^2\right) \end{gathered} \tag{5.4}$$

and so

$$\begin{gathered} [F,G] \sim aF+bG \qquad \text{implies} \\ d\left(\left(\tfrac{t}{n}[F,G]\right)_{t/n}(x), \left(\left(\tfrac{t}{n}(aF+bG)\right)\right)_{t/n}(x)\right) = o\left(\tfrac{1}{n^2}\right) \end{gathered} \tag{5.5}$$

since t is fixed.

We use these facts to establish (5.2), first checking

$$d\left(\left(F_t^*\left(a'F+b'G\right)\right)_{t/n}(x), S\right) = o\left(\tfrac{1}{n}\right)$$

as $n \to \infty$. At the end of the proof we will replace t/n by arbitrary $r \to 0$. Using the linearity of pull-back (Proposition 4.2) we get

$$\begin{aligned} & d\left(\left(F_t^*\left(a'F+b'G\right)\right)_{t/n}(x), S\right) \\ & = d\left(\left(\left(a'\circ F_t\right)F_t^*(F) + \left(b'\circ F_t\right)F_{t/n}^{*(n)}(G)\right)_{t/n}(x), S\right) \\ & = d\left(\left(a_0F + b_0F_{t/n}^{*(n)}(G)\right)_{t/n}(x), S\right) \end{aligned}$$

where $a_0 := a' \circ F_t$ and $b_0 := b' \circ F_t$. Using (5.3) means this last estimate is

$$
\begin{aligned}
&= d\left(\left(a_0F + b_0F^{*(n-1)}_{t/n}\left(\tfrac{t}{n}[F,G]+G\right)\right)_{t/n}(x), S\right)\\
&\le d\left(\left(a_0F + b_0F^{*(n-1)}_{t/n}\left(\tfrac{t}{n}[F,G]+G\right)\right)_{t/n}(x), \left(a_0F + b_0F^{*(n-1)}_{t/n}\left(\tfrac{t}{n}(aF+bG)+G\right)\right)_{t/n}(x)\right)\\
&+ d\left(\left(a_0F + b_0F^{*(n-1)}_{t/n}\left(\tfrac{t}{n}(aF+bG)+G\right)\right)_{t/n}(x), S\right). \qquad (5.6)
\end{aligned}
$$

We estimate the first term as

$$
\begin{aligned}
&d\left(\left(a_0F + b_0F^{*(n-1)}_{t/n}\left(\tfrac{t}{n}[F,G]+G\right)\right)_{t/n}(x), \left(a_0F + b_0F^{*(n-1)}_{t/n}\left(\tfrac{t}{n}(aF+bG)+G\right)\right)_{t/n}(x)\right)\\
&= d\left(\left(b_0F^{*}_{(n-1)t/n}\left(\tfrac{t}{n}[F,G]+G\right)\right)_{t/n}(y), \left(b_0F^{*}_{(n-1)t/n}\left(\tfrac{t}{n}(aF+bG)+G\right)\right)_{t/n}(y)\right)
\end{aligned}
$$

where $y := a_0F_{t/n}(x)$

$$
\begin{aligned}
&= d\left(\left(F^{*}_{(n-1)t/n}\left(\tfrac{t}{n}[F,G]+G\right)\right)_{b_0(y)t/n}(y), \left(F^{*}_{(n-1)t/n}\left(\tfrac{t}{n}(aF+bG)+G\right)\right)_{b_0(y)t/n}(y)\right)\\
&= d\left(\begin{array}{c}\left(F_{-(n-1)t/n}\left(\tfrac{t}{n}[F,G]+G\right)\right)_{b_0(y)t/n}\left(F_{(n-1)t/n}(y)\right)\\ ,\left(F_{-(n-1)t/n}\left(\tfrac{t}{n}(aF+bG)+G\right)\right)_{b_0(y)t/n}\left(F_{(n-1)t/n}(y)\right)\end{array}\right)\\
&= d\left(\begin{array}{c}\left(F_{-(n-1)t/n}\left(\tfrac{t}{n}[F,G]+G\right)\right)_{b_0(y)t/n}(z)\\ ,\left(F_{-(n-1)t/n}\left(\tfrac{t}{n}(aF+bG)+G\right)\right)_{b_0(y)t/n}(z)\end{array}\right) \qquad (5.7)
\end{aligned}
$$

where $z := F_{(n-1)t/n}(y)$. Then by Theorem 2.6, (5.7) is

$$
\begin{aligned}
&\le d\left(\left(\tfrac{t}{n}[F,G]+G\right)_{b_0(y)t/n}(z), \left(\tfrac{t}{n}(aF+bG)+G\right)_{b_0(y)t/n}(z)\right)e^{\Lambda_X(n-1)t/n}\\
&= d\left(G_{b_0(y)t/n}\left(\tfrac{t}{n}[F,G]\right)_{b_0(y)t/n}(z), G_{b_0(y)t/n}\left(\tfrac{t}{n}(aF+bG)\right)_{b_0(y)t/n}(z)\right)e^{\Lambda_X(n-1)t/n}\\
&\le d\left(\left(\tfrac{t}{n}[F,G]\right)_{b_0(y)t/n}(z), \left(\tfrac{t}{n}(aF+bG)\right)_{b_0(y)t/n}(z)\right)e^{\Lambda_X(n-1)t/n}e^{\Lambda_Y b_0(y)t/n}\\
&\le r\left(b_0(y)\left(\tfrac{t}{n}\right)^2\right)e^{\Lambda_X(n-1)t/n+\Lambda_Y b_0(y)t/n} =: o_1\left(\tfrac{1}{n^2}\right) \qquad (5.8)
\end{aligned}
$$

where we define

$$r(s) := d\left([F,G]_s(z), (aF+bG)_s(z)\right).$$

By the main assumption of the theorem, $r(s) = o(s)$ so notice we have $o_1\left(\frac{1}{n^2}\right) = o\left(\frac{1}{n^2}\right)$ but we need to keep a careful record of this estimate as we will be summing n terms like it; the subscript distinguishes o_1 as a specific function.

Substituting (5.8) into (5.6) gives

$$\begin{aligned}
& d\left(\left(F_t^*\left(a'F+b'G\right)\right)_{t/n}(x),S\right) \\
& = d\left(\left(a_0F+b_0F_{t/n}^{*(n)}G\right)_{t/n}(x),S\right) && (5.9)\\
& \leq d\left(\left(a_0F+b_0F_{t/n}^{*(n-1)}\left(\tfrac{t}{n}\left(aF+bG\right)+G\right)\right)_{t/n}(x),S\right)+o_1\left(\tfrac{1}{n^2}\right) \\
& = d\left(\left(\begin{array}{c} a_0F+b_0\frac{t}{n}\left(a\circ F_{(n-1)t/n}\right)F \\ +b_0\cdot\left(\frac{t}{n}\left(b\circ F_{(n-1)t/n}\right)+1\right)F_{t/n}^{*(n-1)}G \end{array}\right)_{t/n}(x),S\right)+o_1\left(\tfrac{1}{n^2}\right) \\
& = d\left(\left(\begin{array}{c} \left[a_0+\left(b_0\frac{t}{n}\left(a\circ F_{(n-1)t/n}\right)\right)\circ\left(a_0F_{t/n}\right)\right]F \\ +b_0\cdot\left(\frac{t}{n}\left(b\circ F_{(n-1)t/n}\right)+1\right)F_{t/n}^{*(n-1)}G \end{array}\right)_{t/n}(x),S\right)+o_1\left(\tfrac{1}{n^2}\right) \\
& = d\left(\left(a_1F+b_1F_{t/n}^{*(n-1)}G\right)_{t/n}(x),S\right)+o_1\left(\tfrac{1}{n^2}\right) && (5.10)
\end{aligned}$$

where

$$\begin{aligned}
a_1 &:= a_0+\left(b_0\tfrac{t}{n}\left(a\circ F_{(n-1)t/n}\right)\right)\circ\left(a_0F_{t/n}\right) \qquad \text{and} \\
b_1 &:= b_0\cdot\left(\tfrac{t}{n}\left(b\circ F_{(n-1)t/n}\right)+1\right).
\end{aligned}$$

This painful calculation from the third line to the fourth line employs the linearity of pull-back (Proposition 4.2); while the fifth line is due to the linearity of F (Proposition 3.10).

After toiling through these many complicated estimates we can relax a bit, since the rest of the proof follows more mechanically by iterating the result of lines (5.9) and (5.10):

$$\begin{aligned}
& d\left(\left(a_0F+b_0F_{t/n}^{*(n)}G\right)_{t/n}(x),S\right) \\
& \leq d\left(\left(a_1F+b_1F_{t/n}^{*(n-1)}G\right)_{t/n}(x),S\right)+o_1\left(\tfrac{1}{n^2}\right) \\
& \leq d\left(\left(a_2F+b_2F_{t/n}^{*(n-2)}G\right)_{t/n}(x),S\right)+o_1\left(\tfrac{1}{n^2}\right)+o_2\left(\tfrac{1}{n^2}\right) \\
& \leq \ldots \leq d\left(\left(a_nF+b_nG\right)_{t/n}(x),S\right)+\sum_{i=1}^{n}o_i\left(\tfrac{1}{n^2}\right) && (5.11)
\end{aligned}$$

where

$$\begin{aligned}
a_2 &:= a_1+\left(b_1\tfrac{t}{n}\left(a\circ F_{(n-2)t/n}\right)\right)\circ\left(a_1F_{t/n}\right) \\
b_2 &:= b_1\cdot\left(\tfrac{t}{n}\left(b\circ F_{(n-2)t/n}\right)+1\right) \qquad \text{and in general} \\
a_i &:= a_{i-1}+\left(b_{i-1}\tfrac{t}{n}\left(a\circ F_{(n-i)t/n}\right)\right)\circ\left(a_{i-1}F_{t/n}\right) \\
b_i &:= b_{i-1}\cdot\left(\tfrac{t}{n}\left(b\circ F_{(n-i)t/n}\right)+1\right)
\end{aligned}$$

In the region of interest the $|a|$ and $|a_0|$ are bounded by some $A\in\mathbb{R}$ and $|b|$ and $|b_0|$ are bounded by some $B \in \mathbb{R}$ so

$$
\begin{aligned}
|b_1| &= \left|b_0 \cdot \left(\tfrac{t}{n}\left(b \circ F_{(n-1)t/n}\right) + 1\right)\right| \le B\left(\tfrac{t}{n}B + 1\right)\\
|b_2| &= \left|b_1 \cdot \left(\tfrac{t}{n}\left(b \circ F_{(n-1)t/n}\right) + 1\right)\right| \le B\left(\tfrac{t}{n}B + 1\right)^2\\
|b_i| &\le B\left(\tfrac{t}{n}B + 1\right)^i \qquad \text{and}
\end{aligned}
$$

$$
\begin{aligned}
|a_1| &= \left|a_0 + b_0 \tfrac{t}{n}\left(a \circ F_{(n-1)t/n}\right)\right| \le A + B\tfrac{t}{n}A\\
|a_2| &= \left|a_1 + b_1 \tfrac{t}{n}\left(a \circ F_{(n-2)t/n}\right)\right| \le \left(A + B\tfrac{t}{n}A\right) + B\left(\tfrac{t}{n}B + 1\right)\tfrac{t}{n}A\\
|a_3| &= \left|a_2 + b_2 \tfrac{t}{n}\left(a \circ F_{(n-3)t/n}\right)\right|\\
&\le A + B\tfrac{t}{n}A + B\left(\tfrac{t}{n}B + 1\right)\tfrac{t}{n}A + B\left(\tfrac{t}{n}B + 1\right)^2\tfrac{t}{n}A\\
|a_i| &\le A + \tfrac{t}{n}AB\sum_{k=0}^{i-1}\left(\tfrac{t}{n}B + 1\right)^k = A + \tfrac{t}{n}AB\frac{\left(\tfrac{t}{n}B + 1\right)^i - 1}{\tfrac{t}{n}B}\\
&= A\left(\tfrac{t}{n}B + 1\right)^i.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
|b_n| &\le B\left(\tfrac{t}{n}B + 1\right)^n \le Be^{tB} \qquad \text{and}\\
|a_n| &\le A\left(\tfrac{t}{n}B + 1\right)^n \le Ae^{tB}.
\end{aligned}
$$

Penultimately, we need to estimate the $o_i\left(\frac{1}{n^2}\right)$. Remember from line (5.8)

$$
o_1\left(\tfrac{1}{n^2}\right) := r\left(b_0(y)\left(\tfrac{t}{n}\right)^2\right)e^{\Lambda_X(n-1)t/n + \Lambda_Y b_0(y)t/n}
$$

where $r(s) = o(s)$, so

$$
\begin{aligned}
o_2\left(\tfrac{1}{n^2}\right) &= r\left(b_1(y)\left(\tfrac{t}{n}\right)^2\right)e^{\Lambda_X(n-2)t/n + \Lambda_Y b_1(y)t/n}\\
&\le B\left(\tfrac{t}{n}B + 1\right)o\left(\left(\tfrac{t}{n}\right)^2\right)e^{\Lambda_X(n-2)t/n + \Lambda_Y B\left(\frac{t}{n}B+1\right)t/n}\\
o_i\left(\tfrac{1}{n^2}\right) &= r\left(b_{i-1}(y)\left(\tfrac{t}{n}\right)^2\right)e^{\Lambda_X(n-i)t/n + \Lambda_Y b_{i-1}(y)t/n}.
\end{aligned}
$$

Consequently

$$
\begin{aligned}
\sum_{i=1}^{n} o_i\left(\tfrac{1}{n^2}\right) &\le \sum_{i=1}^{n} r\left(b_{i-1}(y)\left(\tfrac{t}{n}\right)^2\right)e^{\Lambda_X(n-i)t/n + \Lambda_Y B\left(\frac{t}{n}B+1\right)^{i-1}t/n}\\
&\le o\left(\left(\tfrac{t}{n}\right)^2\right)Be^{tB}\sum_{i=1}^{n} e^{\Lambda_X(n-i)t/n + \Lambda_Y B\left(\frac{t}{n}B+1\right)^{i-1}t/n}
\end{aligned}
$$

since $r\left(b_{i-1}(y)\left(\frac{t}{n}\right)^2\right) = o\left(\left(\frac{t}{n}\right)^2\right)Be^{tB}$ for all i. Therefore

$$
\sum_{i=1}^{n} o_i\left(\tfrac{1}{n^2}\right) \le o\left(\left(\tfrac{t}{n}\right)^2\right)Be^{tB}ne^{\Lambda_X t + \Lambda_Y Be^{tB}t/n} = o\left(\tfrac{1}{n}\right)
$$

as $n \to \infty$. Putting this into (5.11) gives

$$d\left(\left(F_t^*\left(a'F+b'G\right)\right)_{t/n}(x),S\right) \le d\left(\left(a_nF+b_nG\right)_{t/n}(x),S\right)+o\left(\tfrac{1}{n}\right)=o\left(\tfrac{1}{n}\right)$$

because of the uniform bound on $|a_n|$ and $|b_n|$. To see this notice

$$d\left(\left(a_*F+b_*G\right)_{t/n}(x),S\right)=o\left(\tfrac{1}{n}\right)$$

uniformly for bounded a_* and b_* since $a_*F+b_*G \sim b_*G+a_*F$ and as before $\left(b_*G+a_*F\right)_t(x) \in S$ using the uniform Λ and Ω derived in the proofs of Propositions 3.7 and 3.8 (cf. Remark 2.4).

Finally we need to check

$$d\left(\left(F_t^*\left(a'F+b'G\right)\right)_r(x),S\right)=o(r)$$

when r is not necessarily t/n. We may assume $0<t<1$ and $0<r<t$ so that $t=nr+\varepsilon$ for some $0 \le \varepsilon < r$ and integer n with $\frac{t}{r}-1<n\le\frac{t}{r}$. Therefore the above calculations give

$$\begin{aligned}&d\left(\left(F_t^*\left(a'F+b'G\right)\right)_r(x),S\right)=d\left(\left(F_\varepsilon^*F_r^{*(n)}(cF+dG)\right)_r(x),S\right)\\&\le d\left(F_\varepsilon^*\left(a_nF+b_nG\right)_r(x),S\right)+o(r)=o(r).\end{aligned}$$

□

The n-dimensional corollary of this 2-dimensional version is given in the next section.

Remark 5.5. In the assumptions of Theorem 5.4 $[F,G]$ can be replaced with $[X,Y]$ when they are tangent. Since the brackets use $\sqrt{t}$ we have $[F,G] \sim [X,Y]$ when X and Y are 2nd-order tangent to their flows, i.e.,

$$\begin{aligned}d\left(X_t(x),F_t(x)\right)&=O\left(t^2\right) \qquad \text{and}\\ d\left(Y_t(x),G_t(x)\right)&=O\left(t^2\right)\end{aligned}$$

locally uniformly. We denote 2nd-order local uniform tangency by $X \approx F$. This holds, for example, when X comes from a twice continuously differentiable vector field by Taylor's theorem. But in formulating our theorem for the nonsmooth case, the two brackets are not interchangeable. Beware: 2nd-order tangency is "big oh" of t^2, not "little oh".

We might have chosen to define the bracket $[X,Y]$ using the flows instead of the arc fields to simplify the statements of Theorem 5.4 and those below. However it is often easier to calculate the bracket and to check closure using arc fields instead of the flows.

In light of this remark, Theorem 5.4 gives

Corollary 5.6. *Assume X & Y verge, are transverse, and satisfy E1 and E2 on a locally complete metric space M. Further assume X and Y are 2nd-order tangent to their local flows F and G. If $[X,Y] \sim aX+bY$ for some Lipschitz functions $a,b : M \to \mathbb{R}$, then for each $x_0 \in M$ there exists an integral surface S through x_0.*

6 Global Frobenius Theorem

The goal of this section is to recast Theorem 5.4 in the language of distributions and foliations, and so we begin with several definitions. M is, as ever, a locally complete metric space. Just as a distribution on a manifold is a subbundle of the tangent bundle, i.e., a distribution is a set of vectors, we make the following definition on a metric space.

Definition 6.1. A **distribution** Δ on M is a set of arc fields.

The following are the archetypical examples of distributions. Using the addition and multiplication operations defined for arc fields on M (§3) we may define the **linear span** of arc fields:

$$\Delta\left(X^1,\ldots,X^n\right) := \left\{ \sum_{i=1}^n a_i X^i \middle| a_i \in Lip\,(M,\mathbb{R}) \right\}$$

where $Lip\,(M,\mathbb{R})$ is the set of Lipschitz functions on M.

The linear span of two (or more) distributions Δ^1 and Δ^2 on M is also obviously defined to give another distribution

$$\Delta^1 + \Delta^2 := \left\{X+Y | X \in \Delta^1, Y \in \Delta^1\right\}.$$

Writing

$$\Delta(X) := \{aX | a \in Lip\,(M,\mathbb{R})\}$$

we automatically have $\Delta(X,Y) = \Delta(X) + \Delta(Y)$. Associativity also holds for this formal sum:

$$\left(\Delta^1+\Delta^2\right)+\Delta^3 = \Delta^1 + \left(\Delta^2+\Delta^3\right)$$

so we may write finite summands without confusion. Then without difficulty we have

$$\Delta\left(X^1,\ldots,X^n\right) = \sum_{i=1}^n \Delta\left(X^i\right).$$

Commutativity is not generally valid, but it does hold up to tangency, defined below.

For $x \in M$ denote $\Delta_x := \{X(x,\cdot) | X \in \Delta\}$. I.e., Δ_x is a set of curves based at x. For $y \in M$ define (with another overload of d notation)

$$d\left(y,\Delta_x\right) := \inf\{d\left(y, X(x,t)\right) | X \in \Delta \text{ and } t \in [-1,1]\}.$$

Definition 6.2. An arc field X is **locally uniformly tangent** to Δ, denoted $X \sim \Delta$, if for each $x \in M$ there is an arc field $X^\Delta \in \Delta$ with $X \sim X^\Delta$ uniformly in a neighborhood of x.

Two distributions Δ and $\widetilde{\Delta}$ are **locally uniformly tangent**, denoted $\Delta \sim \widetilde{\Delta}$, if $X \sim \widetilde{\Delta}$ for each $X \in \Delta$ and $\widetilde{X} \sim \Delta$ for each $\widetilde{X} \in \widetilde{\Delta}$. Again, $\sim$ is an equivalence relation.

By definition, then, $X \sim \Delta\left(X^1,\ldots,X^n\right)$ if and only if there exist Lipschitz functions $a_i : M \to \mathbb{R}$ such that $X \sim \sum_{i=1}^n a_i X^i$. Restating, tangency between two distributions means a correspondence between tangent arc fields within the distributions.

Example 6.3. When the arc fields $\{X^i\}_{i\in I}$ satisfy E1 and E2 and mutually verge, then we have

$$\Delta\left(X^1,X^2\right)\sim\Delta\left(X^2,X^1\right)\qquad\text{and}\qquad\Delta(X)+\Delta(X)\sim\Delta(X)$$

or more generally, assuming $|I|<\infty$

$$\Delta\left(\{X^i\}_{i\in I}\right)\sim\Delta\left(\{X^j\}_{j\in J}\right)+\Delta\left(\left\{X^k\right\}_{k\in K}\right)\qquad\text{if } J\cup K=I.$$

Definition 6.4. X is (locally uniformly) **transverse** to Δ if for all $x_0\in M$ there exists a $\delta>0$ such that for all $x\in B(x_0,\delta)$ we have

$$d\left(X_x(s),Y_x(t)\right)\geq\delta\left(|s|+|t|\right)$$

for all $Y\in\Delta$ and all $|s|,|t|<\delta$. In this case we have

$$d\left(X_x(t),\Delta\right)\geq\delta|t|.$$

The arc fields $X^1,\ldots,X^n$ are **transverse** to each other if for each $i\in\{1,\ldots,n\}$ we have X^i transverse to

$$\Delta\left(X^1,\ldots,X^{i-1},X^{i+1},\ldots,X^n\right).$$

Inspecting Example 5.2 shows this definition generalizes transversality in $\mathbb{R}^n$. A set of transverse arc fields is meant to generalize linearly independent vector fields.

Definition 6.5. Let $\mathcal{F}:=\left\{X^1,\ldots,X^n\right\}$ be a set of n transverse arc fields which satisfy E1 and E2 on a neighborhood $U\subset M$ and whose flows mutually verge. $\mathcal{F}$ is a **local frame** for a distribution Δ if $\Delta\sim\Delta\left(X^1,\ldots,X^n\right)$ on U. $\mathcal{F}$ is a **global frame** for Δ if every point in M has a neighborhood on which $\mathcal{F}$ is a local frame..

A distribution is n-**dimensional** if each point in M has a neighborhood with a local frame with cardinality n.

Whether global frames of a particular dimension even exist on a space M may be difficult to answer—even when M is a manifold, where the question falls under the purview of topology and global analysis.

Definition 6.6. An n-dimensional distribution Δ is **involutive** if there exists a local frame at each point in M and if each local frame $\left\{X^1,\ldots,X^n\right\}$ has

$$\left[X^i,X^j\right]\sim\Delta$$

for all $i,j\in\{1,\ldots,n\}$.

A **surface** (or n-**surface**) is a topological manifold S (of dimension n). A surface $S\subset M$ is **locally uniformly tangent** to an arc field X, denoted $X\sim S$, if $d\left(X_t(x),S\right)=o(t)$ locally uniformly for $x\in S$.

An n-dimensional surface S is an **integral surface** for an n-dimensional distribution if for any local frame $\left\{X^1,\ldots,X^n\right\}$ we have $\sum_{k=1}^{n}a_kX^k\sim S$ for any choice of Lipschitz functions $a_k:M\to\mathbb{R}$.

An n-dimensional distribution Δ is said to be **integrable** if there exists an integral surface for Δ through every point in M.

Theorem 5.4 has the following corollary:

Theorem 6.7. *An n-dimensional involutive distribution is integrable.*

Proof. $n = 1$ is Theorem 2.2. $n = 2$ is Theorem 5.4. Now proceed by induction. We do enough of the case $n = 3$ to suggest the path; and much of this is copied from the proof of Theorem 5.4.

Choose $x_0 \in M$. Let X, Y, and Z be the transverse arc fields guaranteed in the definition of a 3-dimensional distribution. If we find an integral surface S for $\Delta(X,Y,Z)$ through x_0 then obviously S is an integral surface for Δ. Let F, G, and H be the local flows of X, Y, and Z and define

$$S := \{F_t G_s H_r(x_0) : |r|, |s|, |t| < \delta\}$$

with $\delta > 0$ chosen small enough as in the proof of Theorem 5.4 so that S is a three dimensional manifold. Again we may assume δ is also chosen small enough so that throughout S the functions $|a_k|$ are bounded by A, the constants Λ, Ω, and ρ for X, Y and Z hold uniformly, and the closure of $B(x, 3\delta(\rho+1))$ is complete. Notice

$$\underline{S} := \{G_s H_r(x_0) : |r|, |s| < \delta\}$$

is an integral surface through x_0 for $\Delta(Y,Z)$ by the proof of Theorem 5.4. Notice $S \sim X$ by construction, but it is not immediately clear $S \sim a'X + b'Y + c'Z$ for arbitrarily chosen $a', b', c' \in \mathbb{R}$. Again we really only need to show $S \sim a'F + b'G + c'H$ for an arbitrary point $z := F_t G_s H_r(x_0) \in S$, and again it is sufficient to prove

$$(F_t)^* \left(a'F + b'G + c'H\right) \sim S \qquad \text{at} \qquad y = G_s H_r(x_0)$$

by the construction of S. Continue as above adapting the same tricks from the proof of Theorem 5.4 to the extra dimension. □

Similar to the definition for a surface, an arbitrary set $S \subset M$ is defined to be **locally uniformly tangent** to X if

$$d(X_t(y), S) = o(t)$$

locally uniformly for $y \in S$, denoted $S \sim X$.

Lemma 6.8. *Let $\sigma_x : (\alpha, \beta) \to U \subset M$ be an integral curve to X which meets Condition E1 with uniform constant Λ on a neighborhood U. Assume $S \subset U$ is a closed set with $S \sim X$. Then*

$$d(\sigma_x(t), S) \leq e^{\Lambda|t|} d(x, S) \textit{ for all } t \in (\alpha, \beta).$$

Proof. (Adapted from the proof of Theorem 2.6 given in [5].)

We check only $t > 0$. Define

$$g(t) := e^{-\Lambda t} d(\sigma_x(t), S).$$

For $h \geq 0$, we have

$$\begin{aligned}
&g(t+h) - g(t) \\
&= e^{-\Lambda(t+h)} d(\sigma_x(t+h), S) - e^{-\Lambda t} d(\sigma_x(t), S) \\
&\leq e^{-\Lambda(t+h)} \left[d(\sigma_x(t+h), X_h(\sigma_x(t))) + d(X_h(\sigma_x(t)), X_h(y)) + d(X_h(y), S)\right] \\
&- e^{-\Lambda t} d(\sigma_x(t), S)
\end{aligned}$$

for any $y \in S$, which in turn is

$$\begin{aligned}
&\le e^{-\Lambda(t+h)}\left[d\left(X_h\left(\sigma_x\left(t\right)\right),X_h\left(y\right)\right)+o\left(h\right)\right]-e^{-\Lambda t}d\left(\sigma_x\left(t\right),S\right)\\
&\le e^{-\Lambda t}e^{-\Lambda h}d\left(\sigma_x\left(t\right),y\right)\left(1+\Lambda h\right)-e^{-\Lambda t}d\left(\sigma_x\left(t\right),S\right)+o\left(h\right)\\
&=\left[e^{-\Lambda h}\left(1+\Lambda h\right)d\left(\sigma_x\left(t\right),y\right)-d\left(\sigma_x\left(t\right),S\right)\right]e^{-\Lambda t}+o\left(h\right).
\end{aligned}$$

Therefore

$$g\left(t+h\right)-g\left(t\right)\le\left[e^{-\Lambda h}\left(1+\Lambda h\right)-1\right]e^{-\Lambda t}d\left(\sigma_x\left(t\right),S\right)+o\left(h\right)$$

since y was arbitrary in S. Thus

$$\begin{aligned}
&g\left(t+h\right)-g\left(t\right)\\
&\le o\left(h\right)e^{-\Lambda t}d\left(\sigma_x\left(t\right),S\right)+o\left(h\right)=o\left(h\right)\left(g\left(t\right)+1\right).
\end{aligned}$$

Hence, the upper forward derivative of $g\left(t\right)$ is nonpositive; i.e.,

$$D^+g\left(t\right):=\overline{\lim_{h\to0^+}}\left(\frac{g\left(t+h\right)-g\left(t\right)}{h}\right)\le0.$$

Consequently, $g\left(t\right)\le g\left(0\right)$ or

$$d\left(\sigma_x\left(t\right),S\right)\le e^{\Lambda t}d\left(\sigma_x\left(0\right),S\right)=e^{\Lambda t}d\left(x,S\right).$$

□

Choosing $x \in S$ in Lemma 6.8 gives the following metric space generalization of the Nagumo-Brézis Invariance Theorem (Example 2.3 shows how this generalizes the Banach space setting).

Theorem 6.9. *Let X satisfy E1 and E2 and assume a closed set $S \subset M$ has $S \sim X$. Then for any $x \in S$ we have $F_t\left(x\right) \in S$ for all $t \in (\alpha_x, \beta_x)$. I.e., S is an* ***invariant set*** *under the flow F.*

The case for forward flows is easily achieved *mutatis mutandis*. Cf. [10] for an exposition on general invariance theorems.

Next, local integral surfaces are pieced together to get global integral surfaces.

Proposition 6.10. *If S_1 and S_2 are integral surfaces through $x \in M$, then*
(i) $S_1 \cap S_2$ is an integral surface
(ii) $S_1 \cup S_2$ is an integral surface.
Further, there is a ***unique maximal integral surface*** *S through x, meaning $S \cap S_1 = S_1$ for any integral surface S_1 through x.*

Proof. The case $n = 1$ is true by the uniqueness of integral curves.

For (i) in higher dimensions n, Theorem 6.9 guarantees S_1 and S_2 contain local integral curves for $\sum_{k=1}^{n} a_k X^k$ for all choices of $a_k \in \mathbb{R}$ with initial condition x. Since the X^k are transverse, there is a small neighborhood of x on which all the choices of the parameters a_k give local non-intersecting curves in M which fill up n dimensions giving an integral surface in $S_1 \cap S_2$ (precisely the argument in the second paragraph of the proof of Theorem 5.4).

For (ii) since $S_1 \cap S_2$ is an integral surface inside $S_1 \cup S_2$ the only question is whether the union is still an n-dimensional manifold. Pick $x \in S_1 \cup S_2$ and for $i = 1,2$ let $U_i \subset S_i$ be the n-dimensional neighborhood of x guaranteed by the fact that S_i is an integral surface. As with (i) each of these neighborhoods are manifolds filled by the flows of $\sum_{k=1}^{n} a_k X^k$. By Nagumo's invariance result, Theorem 6.9, they coincide near x.

The maximal integral surface is the union of all integral surfaces through x. □

Definition 6.11. A **foliation** is a partition of M into a set of subsets $\Phi := \{\mathcal{L}_i\}_{i\in I}$ for some indexing set I, where the subsets $\mathcal{L}_i \subset M$ (called **leaves**) are disjoint, connected topological manifolds each having the same dimension.

A foliation Φ is **tangent** to a distribution Δ if the leaves are integral surfaces; in this case we say Δ **foliates** M.

Collecting all these results we have the following version of the Global Frobenius Theorem.

Theorem 6.12. *Let Δ be an n-dimensional distribution on a locally complete metric space M.*

(i) *If Δ is involutive, then Δ is integrable.*

(ii) *If Δ is integrable, then Δ foliates M.*

(iii) *If Δ foliates M into $\Phi := \{\mathcal{L}_x\}_{x\in M}$ then for any $X, Y \in \Delta$, we have $[X,Y]_x(t) \in \mathcal{L}_x$ for $t \in [-1,1]$.*

(iv) *Δ is involutive if and only if Δ has a local frame at each $x \in M$ with commutative flows.*

Proof. (i) is Proposition 6.7.

(ii) is Proposition 6.10.

(iii) follows from Theorem 6.9 and the definition of the bracket.

(iv) $(\Leftarrow)$ This is automatic since the bracket is trivial if the flows commute.

$(\Rightarrow)$ Pick any local frame $\{X^i\}$ at $x \in M$ and construct a commutative frame as follows. Let $\sigma_x^1 : (\alpha, \omega) \to M$ be the integral curve of X^1. (Remember $F_t^1(x) = \sigma_x^1(t)$.) Define a 2-dimensional surface by pushing σ_x^1 along the transverse flow F^2

$$S := \left\{ F_{t_2}^2\left(\sigma_x^1(t_1)\right) \middle| |t_1| < \delta, |t_2| < \delta \right\}$$

for suitably small $\delta > 0$. Restricted to S define $\widetilde{X^1}$ at $s = F_{t_2}^2\left(\sigma_x^1(t_1)\right) \in S$ by

$$\widetilde{X^1}(s) = \widetilde{X^1}_t\left(F_{t_2}^2\left(\sigma_x^1(t_1)\right)\right) := F_{t_2}^2 F_t^1\left(\sigma_x^1(t_1)\right) = \left(F_{t_2*}^2 \circ F_t^1\right)(s)$$

and keep $\widetilde{X^2}\Big|_S := X^2$. This forces

(a) $\widetilde{F^1}$ and $\widetilde{F^2}$ commute

(b) $\widetilde{X^1}$ and $\widetilde{X^2}$ span a surface (S) locally

(c) $\Delta\left(\widetilde{F^1},\widetilde{F^2}\right) \subset \mathcal{L}_x$.

Continue with $n = 3$, etc., pushing forward $\Delta\left(\widetilde{F^1},\widetilde{F^2}\right)$ with F^3 to extend $\widetilde{X^1}$ and $\widetilde{X^2}$ on a local 3-D set and $\widetilde{X^3}\Big|_{3-D\text{ set}} := X^3$. In the end we have an n-dimensional surface (property (b)) in $\mathcal{L}_x$ (property (c), which is the key point of the proof and requires the assumption of the statement of the theorem) and so fills $\mathcal{L}_x$ locally and commutes (property (a)). □

Part (iii) of Theorem 6.12 is as close to a converse of (i) as we have been able to achieve. The bracket is tangent to the distribution in the sense given in the theorem, but not necessarily locally uniformly tangent to a single arc field in the distribution—which is the definition of $\sim$ required for involutivity.

The local frame with commutative flows gives local coordinates on the leaves of the foliation. If the foliation is trivial, having only one leaf, then the flows give local coordinates near each point in M—called **flow coordinates**—in which case M is a topological manifold.

7 Commutativity of Flows

Theorem 7.1. *Assume X and Y satisfy E1 and E2 on a locally complete metric space M. Let F and G be the local flows of X and Y. Then $[F,G] \sim 0$ if and only if F and G commute, i.e.,*

$$F_t G_s(x) = G_s F_t(x), \qquad \text{i.e.,} \qquad F_t^*(G) = G.$$

Proof. The assumption $[F,G] \sim aX + bY$ with $a = b = 0$ allows us to copy the approach in the proof of Theorem 5.4. Let $\delta > 0$ be chosen small enough so

1. the functions $|a|$ and $|b|$ are bounded
2. the constants Λ, Ω, and ρ for X and Y hold uniformly
3. $[F,G] \sim 0$ uniformly

all on $S := B(x, 2\delta(\rho+1))$ and that S is also complete. We check $t > 0$. Since $F_t^*(G)$ and G are both local flows, we only need to show they are tangent to each other and then they must be equal by uniqueness of integral curves.

Imagine being in the context of differentiable manifolds. There, for vector fields f and g with local flows F and G, we would have

$$\lim_{h\to 0}\frac{F_h^*(g) - g}{h} = \mathcal{L}_f g = [f,g] = 0$$

so $F_h^*(g) = g + o(h)$ and thus we expect

$$F_h^*(g) = g + o(h).$$

We might use this idea as before with the linearity of pull-back (Proposition 4.2) to get

$$F_t^*(g) = \lim_{n\to\infty} F_{t/n}^{*(n)}(g) = \lim_{n\to\infty} g + no(1/n) = g$$

as desired.

Now in our context of metric spaces with $t>0$, line (4.2) again gives

$$F^{*}_{t/n}(G)_{t/n}(x)=\left(\tfrac{t}{n}[F,G]+G\right)_{t/n}(x).$$

For $t<0$ one would use (4.3). Also we again have

$$[F,G]\sim 0 \qquad \text{implies}$$
$$d\left(\left(\tfrac{t}{n}[F,G]\right)_{t/n}(x),x\right)=o\left(\tfrac{1}{n^2}\right).$$

Using these tricks (and Theorem 2.6 in the fourth line following) gives

$$\begin{aligned}
&d\left(\left(F^{*}_{t}(G)\right)_{t/n}(x),G_{t/n}(x)\right)=d\left(\left(F^{*(n-1)}_{t/n}F^{*}_{t/n}(G)\right)_{t/n}(x),G_{t/n}(x)\right)\\
&=d\left(F^{*(n-1)}_{t/n}\left(\tfrac{t}{n}[F,G]+G\right)_{t/n}(x),G_{t/n}(x)\right)\\
&\leq d\left(F^{*(n-1)}_{t/n}\left(G_{t/n}\tfrac{t}{n}[F,G]_{t/n}(x)\right),F^{*(n-1)}_{t/n}G_{t/n}(x)\right)+d\left(F^{*(n-1)}_{t/n}G_{t/n}(x),G_{t/n}(x)\right)\\
&\leq d\left(G_{t/n}\tfrac{t}{n}[F,G]_{t/n}(y),G_{t/n}(y)\right)e^{\Lambda_X\frac{t(n-1)}{n}}+d\left(F^{*(n-1)}_{t/n}G_{t/n}(x),G_{t/n}(x)\right)
\end{aligned}$$

where $y:=F_{(n-1)t/n}(x)$

$$\leq d\left(\tfrac{t}{n}[F,G]_{t/n}(y),y\right)e^{\Lambda_Y t/n}e^{\Lambda_X\frac{t(n-1)}{n}}+d\left(F^{*(n-1)}_{t/n}G_{t/n}(x),G_{t/n}(x)\right)$$

and so

$$\begin{aligned}
&d\left(\left(F^{*}_{t}(G)\right)_{t/n}(x),G_{t/n}(x)\right)\\
&\quad\leq d\left(F^{*(n-1)}_{t/n}G_{t/n}(x),G_{t/n}(x)\right)+e^{\Lambda_Y t/n+\Lambda_X\frac{t(n-1)}{n}}o_1\left(\tfrac{1}{n^2}\right)
\end{aligned}$$

where $o_1\left(\frac{1}{n^2}\right):=d\left(\frac{t}{n}[F,G]_{t/n}(y),y\right)$.

Iterating this result gives

$$\begin{aligned}
&d\left(\left(F^{*n}_{t/n}(G)\right)_{t/n}(x),G_{t/n}(x)\right)\\
&\leq d\left(F^{*(n-1)}_{t/n}G_{t/n}(x),G_{t/n}(x)\right)+e^{\Lambda_Y t/n+\Lambda_X\frac{t(n-1)}{n}}o_1\left(\tfrac{1}{n^2}\right)\\
&\leq d\left(F^{*(n-2)}_{t/n}G_{t/n}(x),G_{t/n}(x)\right)+e^{\Lambda_Y t/n+\Lambda_X\frac{t(n-2)}{n}}o_2\left(\tfrac{1}{n^2}\right)+e^{\Lambda_Y t/n+\Lambda_X\frac{t(n-1)}{n}}o_1\left(\tfrac{1}{n^2}\right)\\
&\leq\ldots\leq d\left(F^{0}_{t/n}G_{t/n}(x),G_{t/n}(x)\right)+e^{\Lambda_Y t/n}\sum_{i=1}^{n}o_i\left(\tfrac{1}{n^2}\right)e^{\Lambda_X\frac{t(n-i)}{n}}\\
&=e^{\Lambda_Y t/n}\sum_{i=1}^{n}o_i\left(\tfrac{1}{n^2}\right)e^{\Lambda_X\frac{t(n-i)}{n}}
\end{aligned}$$

where $o_i\left(\frac{1}{n^2}\right):=d\left(\frac{t}{n}[F,G]_{t/n}(y_i),y_i\right)$ and $y_i:=F_{(n-i)t/n}(x)$. Since $d\left(\frac{t}{n}[F,G]_{t/n}(y),y\right)=$

$o\left(\frac{1}{n^2}\right)$ uniformly for $y \in B\left(x, 2\delta\left(\rho+1\right)\right)$ we have

$$
\begin{aligned}
& d\left(\left(F_{t/n}^{*n}\left(G\right)\right)_{t/n}\left(x\right), G_{t/n}\left(x\right)\right) \\
& \leq e^{\Lambda_Y t/n} \sum_{i=1}^{n} o_i\left(\tfrac{1}{n^2}\right) e^{\Lambda_X \frac{t(n-i)}{n}} = o\left(\tfrac{1}{n^2}\right) e^{\Lambda_Y t/n} \sum_{i=1}^{n} e^{\Lambda_X \frac{t(n-i)}{n}} \\
& = o\left(\tfrac{1}{n^2}\right) e^{\Lambda_Y t/n} e^{\Lambda_X t} \sum_{i=1}^{n} \left(e^{-\frac{t}{n}}\right)^i = o\left(\tfrac{1}{n^2}\right) e^{\Lambda_Y t/n + \Lambda_X t} \frac{1-\left(e^{-\frac{t}{n}}\right)^{n+1}}{1-\left(e^{-\frac{t}{n}}\right)}.
\end{aligned}
$$

So

$$
d\left(\left(F_t^*\left(G\right)\right)_{t/n}\left(x\right), G_{t/n}\left(x\right)\right) = o\left(\tfrac{1}{n}\right)
$$

and $F_t^*\left(G\right) \sim G$ by the same argument at the last paragraph of the proof of Theorem 5.4.

The converse is trivial. □

Using Example 2.3, this theorem applies to the non-locally compact setting with nonsmooth vector fields. [13], another paper which inspires this one, obtains similar results in the setting of manifolds with a very different approach.

8 Examples

Example 8.1. Let M be a Banach space. First let X and Y be translations in the directions of u and $v \in M$

$$
X_t\left(x\right) := x + tu \qquad Y_t\left(x\right) := x + tv
$$

then $F = X$ and $G = Y$ for $|t| \leq 1$. Obviously $[F, G] = 0$ and the flows commute.

Next consider the dilations X and Y about u and $v \in M$

$$
X_t\left(x\right) := \left(1+t\right)\left(x-u\right) + u \qquad Y_t\left(x\right) := \left(1+t\right)\left(x-v\right) + v.
$$

The flows are computable with little effort using Euler curves, e.g.,

$$
F_t\left(x\right) = \lim_{n\to\infty} X_{t/n}^{(n)}\left(x\right) = e^t x - \left(e^t - 1\right) u.
$$

Then for $t \geq 0$

$$
\begin{aligned}
& [F, G]_{t^2}\left(x\right) \\
& = G_{-t} F_{-t} G_t F_t\left(x\right) \\
& = e^{-t}\left[e^{-t}\left(e^t\left[e^t x - \left(e^t - 1\right) u\right] - \left(e^t - 1\right) v\right) - \left(e^{-t} - 1\right) u\right] - \left(e^{-t} - 1\right) v \\
& = x - u + e^{-t} u - e^{-t} v + e^{-2t} v - e^{-2t} u + e^{-t} u - e^{-t} v + v \\
& = x + \left(v - u\right)\left(e^{-t} - 1\right)^2
\end{aligned}
$$

so $[F,G] \sim Z$ where Z is the translation $Z_t(x) := x + t(v-u)$ since, for instance with $t > 0$

$$d\left([F,G]_t(x), Z_t(x)\right)$$
$$= |v-u| \left| \left(e^{-\sqrt{t}} - 1\right)^2 - t \right| = |t|\,|v-u| \left| \left(\frac{e^{-\sqrt{t}} - 1}{\sqrt{t}}\right)^2 - 1 \right| = o(t).$$

Hence the distribution $\Delta(X,Y)$ is not involutive. However, this shows the set of all dilations generates all translations using brackets. Using the same tricks we've just employed, it is easy to check the bracket of a dilation and a translation is tangent to a translation, e.g., if $F_t(x) := x + tu$ and $G_t(x) := e^t x$ (dilation about 0) then $[F,G] \sim F$ since for $t > 0$

$$[F,G]_{t^2}(x) = G_{-t}F_{-t}G_tF_t(x) = e^{-t}\left[e^t[x+tu] - tu\right] = x + tu\left(1 - e^{-t}\right)$$

and so

$$d\left([F,G]_t(x), F_t(x)\right) = |tu| \left| \tfrac{1-e^{-\sqrt{t}}}{\sqrt{t}} - 1 \right| = o(t).$$

Example 8.2. Now consider the metric space $(H(\mathbb{R}^n), d_H)$ where $H(\mathbb{R}^n)$ is the set of non-void compact subsets of $\mathbb{R}^n$ for $n \in \mathbb{N}$ and d_H is the **Hausdorff metric** given by

$$d_H(a,b) := \max\left\{ \sup_{x\in a}\left[\inf_{y\in b}\{d(x,y)\}\right], \sup_{y\in a}\left[\inf_{x\in b}\{d(x,y)\}\right] \right\}.$$

To aid intuition it might help to see one more equivalent definition:

$$d_H(a,b) = \inf\{r \geq 0 \,|\, b \subset B(a,r) \text{ and } a \subset B(b,r)\}$$

where $B(a,r) := \bigcup_{x\in a} B(x,r)$.

$H(\mathbb{R}^n)$ has several useful topological properties in common with $\mathbb{R}^n$. It is separable, complete and even locally compact. However, this space is incapable of accepting any natural linear structure. $H(\mathbb{R}^n)$ is a particularly strange space topologically because, despite being locally compact, $H(\mathbb{R}^n)$ is infinite dimensional by most any measure we can attempt to apply. We can even find infinitely many transverse flows.

First define arc fields X and $Y : H(\mathbb{R}^n) \times [-1,1] \to H(\mathbb{R}^n)$ to be the translations in the directions of u and $v \in \mathbb{R}^n$

$$X_t(a) := a + tu \qquad Y_t(a) := a + tv$$

where $a + tu := \{x + tu |\ x \in a\}$, e.g. Then X and Y are their own flows for $|t| \leq 1$. It is straightforward to check X and Y are transverse, they verge and $[F,G] = 0$. This shows the flows commute, and the 2-dimensional distribution generated by X and Y is involutive, integrable and foliates $H(\mathbb{R}^n)$ with 2-dimensional topological manifolds—in a space which is decidedly not locally linear.

Next try applying the theory described in this paper to X, Y and the dilation $Z_t(a) := e^t a = \{e^t x |\ x \in a\}$. The form of the calculations in Example 8.1 will be repeated in this new context.

Finally consider the arc fields

$$X_t^1(a) := \left\{ x + t\left(0, e^{-x_1^2}\right) \middle| \; x = (x_1, x_2) \in a \right\}$$
$$X_t^2(a) := a + t(\cos\theta, \sin\theta)$$
$$X_t^3(a) := e^t(a - \overline{a}) + \overline{a}.$$

where $\theta := \sup_{x,y \in a} \{|x-y|\}$ is the diameter of a, and $\overline{a}$ is the center of mass of the convex hull of a. The distribution $\Delta\left(X, X^1\right)$ where $X_t(a) := a + tu$ where $u := (1,0)$ is called the "Taffy distribution" because the effect of the bracket $\left[X, X^1\right]_t(a)$ is similar to a taffy-pulling machine on a blob $a \in H\left(\mathbb{R}^2\right)$, as will become clearer after studying Example 8.3.

In this next example we explore non-involutive metric space distributions on classical Hilbert space $L^2(\mathbb{R})$.

Example 8.3 (two parameter decomposition of L^2)**.** Now let M be real Hilbert space $L^2(\mathbb{R})$. Since M is Banach the results of the previous example hold. In this example, though, we only consider translation in $L^2(\mathbb{R})$ by a function $h \in C^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$ denoted

$$X_t(f) := f + th.$$

Here on $M := L^2(\mathbb{R})$ there is another obvious candidate for an elementary flow: translation with respect to the variable x, i.e.,

$$Y_t(f)(x) := f(x+t).$$

Unlike dilation and translation, the dynamic engendered by Y seemingly has nothing to do with the vector space structure of $L^2(\mathbb{R})$. In fact, despite appearances, Y is a nonsmooth flow: notice for example with the characteristic function χ as initial condition,

$$\left.\frac{d}{dt} Y_t\left(\chi_{[0,1]}\right)\right|_{t=0} \notin L^2(\mathbb{R}).$$

Interpreted as a flow on a metric space, however, this is no obstacle. We refer to X as **vector space translation** and Y as **function translation**. Notice X and Y are their own flows (for $|t| \leq 1$). It is straightforward to check X & Y verge when, for example, $h \in C^1(\mathbb{R})$ with derivative $h' \in L^2(\mathbb{R})$:

$$\begin{aligned}
& d\left(Y_s X_t(f), X_t Y_s(f)\right) \\
&= \sqrt{\int \left(f(x+s) + th(x+s) - [f(x+s) + th(x)]\right)^2 dx} \\
&= |st| \sqrt{\int \left(\frac{h(x+s) - h(x)}{s}\right)^2 dx} \\
&= O(|st|)
\end{aligned}$$

uniformly for $f \in M$. Since they obviously satisfy E1 and E2, Theorem 3.9 promises a unique flow for their sum. This was introduced by Colombo and Corli in [7, section 5.2] with other interesting function space examples with motivation from partial differential equations.

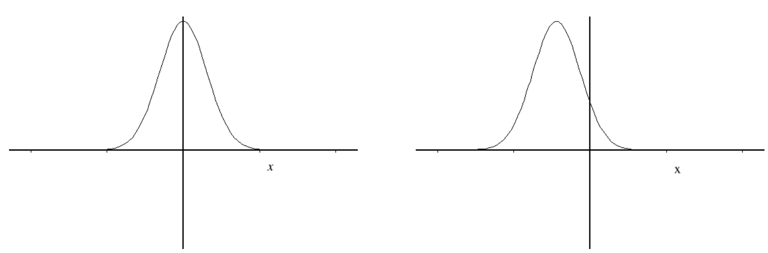
Let us now compute the bracket. We check $t > 0$ explicitly, skipping the case $t \leq 0$ though this is just as easy.

$$\begin{aligned}
&[X,Y]_{t^2}(f)(x) \\
&= Y_{-t}X_{-t}Y_tX_t(f)(x) = Y_{-t}X_{-t}\left[f(x+t) + th(x+t)\right] \\
&= f(x) + th(x) - th(x-t) = f(x) + t^2\left[\frac{h(x) - h(x-t)}{t}\right].
\end{aligned}$$

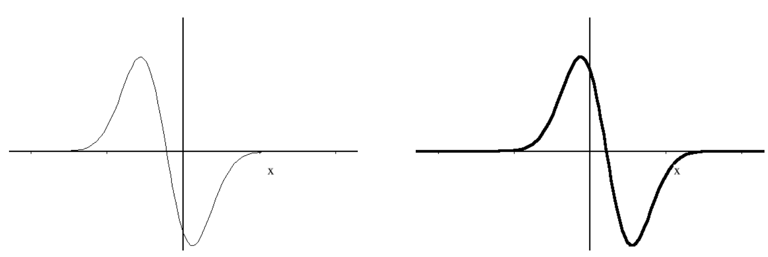
Defining a new arc field $Z_t(f) := f + th'$ (which is its own flow) we therefore have

$$d\left([X,Y]_t(f), Z_t(f)\right) = |t|\sqrt{\int_{\mathbb{R}}\left(\frac{h(x) - h(x-t)}{t} - h'(x)\right)^2 dx} = o(t)$$

when $h \in C^1(\mathbb{R})$ with $h' \in L^2(\mathbb{R})$. Thus $[X,Y] \sim Z$. (See Figures 4 and 5 with $h(x) = e^{-x^2}$.)



$X_1(0) = e^{-x^2}$ $\qquad$ $Y_1X_1(0) = e^{-(x+1)^2}$



$X_{-1}Y_1X_1(0) = e^{-(x+1)^2} - e^{-x^2}$ $\qquad$ $Y_{-1}X_{-1}Y_1X_1(0) \neq 0$

Figure 4: X and Y do not commute.

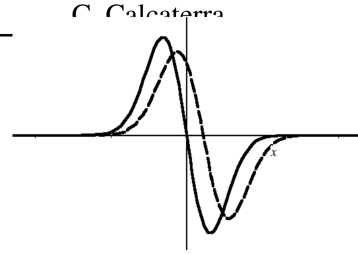


Figure 5: $\left\| Y_{-\sqrt{t}} X_{-\sqrt{t}} Y_{-\sqrt{t}} X_{-\sqrt{t}} (0) - t\left(\frac{d}{dx} e^{-x^2}\right) \right\|_2 = o(t)$

This has remarkable consequences. Using the idea of Chow's Theorem from control theory (also called the Chow-Rashevsky Theorem or Hermes' Theorem), if the $(n+1)$-st derivative $h^{[n+1]}$ is not contained in $span\left\{h^{[i]} : 0 \le i \le n\right\}$ then iterating the process of bracketing X and Y generates a large space reachable via repeated compositions of X and Y. Denoting

$$\overset{n}{Z}_t(f) := f + t h^{[n]} \tag{8.1}$$

successive brackets of X and Y are

$$\begin{aligned} [X,Y] &\sim Z =: \overset{1}{Z} \\ \left[X\overset{2}{,}Y\right] &:= [[X,Y],Y] \sim \overset{2}{Z} \\ [X\overset{n}{,}Y] &:= \underbrace{[[\ldots[[X,Y],Y],\ldots,Y],Y]}_{n \text{ times}} \sim \overset{n}{Z}. \end{aligned} \tag{8.2}$$

For notational purposes we set $\left[X\overset{0}{,}Y\right] := X$. In the particular case $h(x) := e^{-x^2}$ all of $L^2(\mathbb{R})$ is reachable by X and Y.

To see this we apply the theory of orthogonal functions with the Hermite[1] polynomials

$$H_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} = (-1)^n e^{x^2} h^{[n]}(x)$$

which have dense span in $L^2(\mathbb{R})$ when multiplied by $e^{-x^2/2}$. Those familiar with orthogonal expansions can predict the rest; we review some of the details.

$$\left\{ \frac{1}{\sqrt{n!2^n\sqrt{\pi}}} H_n(x) e^{-x^2/2} : n \in \mathbb{N} \right\}$$

is a basis of $L^2(\mathbb{R})$ and is orthonormal since

$$\int_{\mathbb{R}} H_m(x) H_n(x) e^{-x^2} dx = n! 2^n \sqrt{\pi} \delta_{mn}. \tag{8.3}$$

The Hermite polynomials also satisfy some useful relations

$$H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x) \qquad \text{and} \qquad H_n'(x) = 2n H_{n-1}(x). \tag{8.4}$$

[1]We may of course use other orthogonal families with a different choice of h, particularly when the domain of interest is other than $\mathbb{R}$; e.g., scaled Chebyshev polynomials for $[0,2\pi)$, etc. We expect many choices of h give controllable systems whether the brackets generate orthogonal sets or not.

Given any $g \in L^2(\mathbb{R})$ it is possible to write

$$g(x) = \sum_{n=0}^{\infty} a_n \frac{1}{\sqrt{n!2^n\sqrt{\pi}}} H_n(x) e^{-x^2/2} \tag{8.5}$$

(equality in the L^2 sense) where

$$a_n := \frac{1}{\sqrt{n!2^n\sqrt{\pi}}} \int_{\mathbb{R}} g(x) H_n(x) e^{-x^2/2}(x)\, dx \in \mathbb{R}.$$

The necessity of this formula for a_n can easily be checked by multiplying both sides of (8.5) by $H_n(x) e^{-x^2/2}$, integrating and applying (8.3). However, we want

$$g = \sum_{n=0}^{\infty} c_n h^{[n]}$$

so apply the above process to $g(x) e^{x^2/2}$ instead[2]. Then

$$\begin{aligned} g(x) e^{x^2/2} &= \sum_{n=0}^{\infty} b_n \frac{1}{\sqrt{n!2^n\sqrt{\pi}}} H_n(x) e^{-x^2/2} \qquad \text{so} \\ g &= \sum_{n=0}^{\infty} c_n h^{[n]} \end{aligned}$$

where

$$\begin{aligned} b_n &:= \frac{1}{\sqrt{n!2^n\sqrt{\pi}}} \int_{\mathbb{R}} g(x) e^{x^2/2} H_n(x) e^{-x^2/2}(x)\, dx \qquad \text{so that} \\ c_n &:= \frac{(-1)^n}{n!2^n\sqrt{2\pi}} \int_{\mathbb{R}} g(x) h^{[n]}(x) e^{x^2}(x)\, dx. \end{aligned}$$

Therefore when N is large, g is approximated by

$$\sum_{n=0}^{N} c_n h^{[n]} = F_1(0)$$

where F is the flow of the arc field

$$\widetilde{X} := \sum_{n=0}^{N} c_n [X^n; Y]$$

which we follow for unit time starting with initial condition $0 \in L^2(\mathbb{R})$. F can, of course, be approximated by Euler curves

$$F_1(0) = \lim_{n\to\infty} \widetilde{X}^{(n)}_{1/n}(0)$$

and since $\widetilde{X}$ is merely a (complicated) composition of X and Y, this gives us a simple algorithm for approximating any function g with only two simple flows.

[2]The function $g(x) e^{x^2/2}$ is no longer necessarily L^2, of course, but here we lapse into the habit of ignoring convergence issues as they are important for the theoretical proof that all of $L^2(\mathbb{R})$ is reachable with X and Y, but not central to this demonstration. This theoretical lapse is easily remedied by multiplying by the characteristic function $\chi_{[-m,m]}$ to guarantee all of the following integrals converge, then letting $m \to \infty$ at the end.

Let us compute a basic example to illustrate this surprising fact. Choosing at random $g(x) := \chi_{[0,1]}(x)$, the characteristic function of the unit interval, we have

$$c_n := \tfrac{(-1)^n}{n!2^n\sqrt{2\pi}} \int_0^1 H_n(x)\,dx = \tfrac{(-1)^n}{2(n+1)n!2^n\sqrt{2\pi}} \left[H_{n+1}(1) - H_{n+1}(0)\right] \qquad \text{so, e.g.,}$$
$$c_0 = \tfrac{1}{\sqrt{2\pi}},\ c_1 = \tfrac{-1}{2\sqrt{2\pi}},\ c_2 = \tfrac{1}{12\sqrt{2\pi}},\ c_3 = \tfrac{1}{12\sqrt{2\pi}},\ c_4 = \tfrac{1}{480\sqrt{2\pi}},\ \text{etc.}$$

by (8.4). Then stopping for the purposes of illustration at $N = 3$ our function g is approximated by

$$\sum_{n=0}^{3} c_n h^{[n]}.$$

Notice the flow of $\overset{i}{Z}$ from (8.1) is locally the same as $\overset{i}{Z}$ since it is just vector space translation, so we will use the same symbol. All vector space translations commute under (arc field) addition, and the arc field

$$\widetilde{Z}_t(f) := \left(c_0\overset{0}{Z} + c_1\overset{1}{Z} + c_2\overset{2}{Z} + c_3\overset{3}{Z}\right)_t (f)$$

is locally equal to its flow. Obviously

$$\widetilde{Z}_1(0) = \sum_{n=0}^{3} c_n h^{[n]}$$

and $\widetilde{Z} \sim \widetilde{X}$ where

$$\begin{aligned}\widetilde{X}_t(f) &:= \left(c_0 X + c_1 [X,Y] + c_2 [[X,Y],Y] + c_3 [[[X,Y],Y],Y]\right)_t (f)\\ &= \left(\sum_{n=0}^{3} c_n [X \overset{n}{,} Y]\right)_t (f).\end{aligned}$$

Remember the arc field bracket and the arc field sum are defined as nothing more than compositions of arc fields, e.g.,

$$\begin{aligned}&(c_0 X + c_1 [X,Y] + c_2 [[X,Y],Y])_t\\ &= [[X,Y],Y]_{c_2 t} [X,Y]_{c_1 t} X_{c_0 t}\end{aligned}$$

and, e.g., when $t > 0$

$$\begin{aligned}&c_2 [[X,Y],Y]_t\\ &= Y_{-\sqrt{c_2 t}} \left(X_{-\sqrt{c_2 t}} Y_{-\sqrt{c_2 t}} X_{\sqrt{c_2 t}} Y_{\sqrt{c_2 t}}\right) Y_{\sqrt{c_2 t}} (Y_{-\sqrt{c_2 t}} X_{-\sqrt{c_2 t}} Y_{\sqrt{c_2 t}} X_{\sqrt{c_2 t}}).\end{aligned}$$

Therefore this approximation of g is achieved by computing the Euler curves for $\widetilde{X}$ which is a complicated process (with a simple formula) of composing the elementary operations of function translation (Y) and vector space translation by the Gaussian (X).

Continuing the example, for choices of h other than the Gaussian it may be the case that $h^{[n+1]} \in span\left\{h^{[i]} : 0 \le i \le n\right\}$. Then the space reachable by X and Y is precisely limited. E.g., when h is a trigonometric function from the orthogonal Fourier decomposition of L^2

the parameter space is two-dimensional, or when h is an n-th order polynomial in the context of $M = L^2[a,b]$ then the parameter space is $(n+1)$-dimensional.

Restating these results in different terminology: Controlling amplitude and phase the 2-parameter system is holonomically constrained. Controlling phase and superposition perturbation (Y and X) generates a larger space of signals; how much Y and X deviate from holonomy depends on the choice of perturbation function h. Consequently, a result for signal analysis is: controlling two parameters is enough to generate any signal.

We collect some of the results of the previous example. Denote the **reachable set** of X and Y by

$$R(X,Y) := \left\{ Y_{s_n} X_{t_n} Y_{s_{n-1}} X_{t_{n-1}} \ldots Y_{s_1} X_{t_1}(0) \in L^2(\mathbb{R}) : s_i, t_i \in \mathbb{R},\ n \in \mathbb{N} \right\}$$

where $0 \in L^2(\mathbb{R})$ is the constant function. $R(X,Y)$ is the set of all finite compositions of X and Y.

Theorem 8.4. *Let $h \in L^2(\mathbb{R})$ be the Gaussian $h(x) := e^{-x^2}$ and define*

$$X_t(f) := f + th \qquad \text{and} \qquad Y_t(f)(x) := f(x+t).$$

Then $R(X,Y)$ is dense in $L^2(\mathbb{R})$.

This result is a constructive approach related to Wiener's Tauberian Theorem and the Müntz-Szász Theorem (see [16] and [9]) as will be detailed in a forthcoming paper.

Algorithm 8.5. Let $g \in L^2(\mathbb{R})$ be such that $\int_{\mathbb{R}} \left[g(x) e^{x^2/2} \right]^2 dx < \infty$. Then

$$g = \lim_{n\to\infty} \widetilde{X}^{(n)}_{1/n}(0)$$

where

$$\widetilde{X} := \sum_{n=0}^{\infty} c_n [X \overset{n}{,} Y] \qquad \text{with} \qquad c_n := \frac{(-1)^n}{n! 2^n \sqrt{2\pi}} \int_{\mathbb{R}} g(x) h^{[n]}(x) e^{x^2}(x)\, dx$$

$$\text{and} \qquad [X \overset{n}{,} Y] := \underbrace{[[\ldots[[X,Y],Y],\ldots,Y],Y]}_{n \text{ times}}$$

and

$$[X,Y](f,t) := \begin{cases} Y_{-\sqrt{t}} X_{-\sqrt{t}} Y_{\sqrt{t}} X_{\sqrt{t}}(f) & \text{for } t \geq 0 \\ X_{-\sqrt{|t|}} Y_{-\sqrt{|t|}} X_{\sqrt{|t|}} Y_{\sqrt{|t|}}(f) & \text{for } t < 0 \end{cases}$$

for any $f \in L^2(\mathbb{R})$.

Let us recast this result in a setting more familiar in infinite-dimensional control theory. Consider the bang-bang control system given by the partial differential equation (PDE)

$$u_t = \phi(t) u_x + \psi(t) h \tag{8.6}$$

where $u = u(x,t)$, $h = h(x)$ is as above a smooth function with square integrable derivatives of all orders, while ϕ and ψ are the "bang-bang" controls with range limited to $\{-1,0,1\}$. Notice when $\phi(t) = 0$ and $\psi(t) = \pm 1$ on an interval, the solution to the PDE coincides with X. When $\phi(t) = \pm 1$ and $\psi(t) = 0$ on an interval, the solution to the PDE coincides with Y. So to put the above results in still other terminology, this bang-bang control system is controllable when $h(x) = e^{-x^2}$.

Example 8.6. Let us continue Example 8.3 with $M = L^2(\mathbb{R})$ and

$$X_t(f) := f + th \qquad \text{and} \qquad Y_t(f)(x) := f(x+t)$$

which are vector space translation and function translation. Define the arc fields

$$V_t(f) := e^t f \qquad \text{and} \qquad W_t(f)(x) := f\left(e^t x\right)$$

which may be thought of as **vector space dilation** (about the point $0 \in M$) and **function dilation** (about the point $0 \in \mathbb{R}$). Again, V and W are coincident with their own flows. Using the same approach as in Example 8.3 it is easy to check the brackets satisfy

$$\begin{array}{ll}
[X,Y]_t(f) = f + th' + o(t) & [X,V]_t(f) = f + th + o(t) \\
[X,W]_t(f)(x) = f(x) + txh'(x) + o(t) & [Y,V] = 0 \\
[Y,W]_t(f)(x) = f(x-t) + o(t) & [V,W] = 0
\end{array}$$

assuming for the $[X,Y]$ and $[X,W]$ calculations that $h \in C^1(\mathbb{R})$ and $h' \in L^2(\mathbb{R})$. Consequently

$\Delta(X,Y)$ may be highly non-involutive depending on h,
$\Delta(X,V)$ is involutive, but X and V do **not** commute,
$\Delta(X,W)$ may be highly non-involutive depending on h,
$\Delta(Y,V)$ is involutive; Y and V commute,
$\Delta(Y,W)$ is involutive, but Y and W do **not** commute,
$\Delta(V,W)$ is involutive; V and W commute.

When h is chosen correctly, X and W control many function spaces, similarly to X and Y. The four involutive distributions foliate $L^2(\mathbb{R})$.

Finally we give two examples on probability spaces.

Example 8.7. (Probability distributions and geometry of the infinite-dimensional unit sphere) Begin with $M := L^p(\mathbb{R})$ and modify the previous example with the arc fields

$$V_t(f)(x) := \frac{e^{tg} f}{\|e^{tg} f\|_p}$$

and

$$W_t(f)(x) := e^{tc/p} f\left(e^{tc} x\right)$$

where $c \in \mathbb{R}$ and we pick $g \in L^p(\mathbb{R})$ satisfying $\frac{d^n}{dx^n} g \in L^p(\mathbb{R})$ for all $n \in \mathbb{N}$. V and W are well-defined arc fields restricted to the subsets

$$S := \left\{ f \in L^p(\mathbb{R}) \mid \ \|f\|_p = 1 \right\} \text{ and } S^+ := \left\{ f \in L^p(\mathbb{R}) \mid \ f \geq 0 \wedge \|f\|_p = 1 \right\}.$$

S is the unit sphere; and the "hemisphere", S^+, corresponds to a space of probability distributions on $\mathbb{R}$. These arc fields are their own flows. $Y_t(f)(x) := f(x+t)$ also restricts to an arc field on S and on S^+. Further, $X_t(f) := (1-t) f + th$ restricts to an arc field on S if we choose $h \in S$, but on S^+ we must restrict X to $t \geq 0$. Then picking $h \in L^p(\mathbb{R})$ satisfying $\frac{d^n}{dx^n} h \in L^p(\mathbb{R})$ for all $n \in \mathbb{N}$ we get the following table of brackets

$[\cdot,\cdot] \sim$	W	V
X	$X_t^1(f) := f(x) + t\left[cxg_1'(x) + \left(1 + \frac{c}{p}\right) g_1(x)\right]$	$X_t^2(f) := \frac{f + tg_1 \cdot g_9}{\Vert f + tg_1 \cdot g \Vert_p}$
Y	$-cY$	$X_t^3(f) := \frac{e^{-tg_9'} f}{\left\Vert e^{-tg_9'} f \right\Vert_p}$

All of these arc fields give 1-dimensional foliations of $M \setminus \{0\}$ and since they are each invariant on the unit ball $B(0,1) \subset M$ they also foliate $B \setminus \{0\}$ by a family of isometries as well as S. Interestingly the length of the integral curves are infinite (whenever the initial condition is not $f \equiv 0$) and not geodesics (locally minimal-length curves). An insight into the infinite dimension of L^2 comes from comparing these foliations to the 1-dimensional foliation of $\mathbb{R}^2 \setminus \{0\}$ given by rotations which also foliates the ball $B_{\mathbb{R}^2}(0,1) \setminus \{0\}$. These rotations are a family of isometries, where the integral curves have finite length $2\pi r$.

Since the bracket satisfies $[Y,W] \sim -Y$ is involutive, we know $\Delta(Y,W)$ gives a 2-dimensional foliation of $M \setminus \{0\}$ and $B \setminus \{0\}$ and S. The area (with any reasonable definition) of each leaf is again infinite. Higher dimensional foliations are given by adding transverse arc fields.

Finally, the results of Example 8.7 extend immediately to more abstract spaces.

Example 8.8. Given any complete metric space we let $\mathcal{B}(M)$ denote the σ-algebra of Borel sets in M (i.e., all countable unions, intersections and complements of open subsets of M). A **Radon measure** on a metric space is inner regular (any Borel set A satisfies $\mu(A) = \sup\{\mu(S) |\ \text{compact } S \subset A\}$) and locally finite (for any $x \in M$ there exists $r > 0$ such that $\mu(B(x,r)) < \infty$) on $\mathcal{B}(M)$. The **support** of a Radon measure μ is the set

$$supp(\mu) := \{x \in M |\ \mu(B(x,r)) > 0,\ \forall r > 0\}.$$

Denote $\|\mu\| := \mu(M) \in [0,\infty]$. A metric space (M,d) with a radon measure μ is called a **metric measure space** (M,d,μ).

On any complete separable metric space (M,d) define the spaces of measures

$$\begin{aligned} \mathcal{M}(M) &:= \{\text{Radon measures on } M\} \\ \mathfrak{F}(M) &:= \left\{\mu \in \mathcal{M}(M) \big|\ \|\mu\| = 1 \ \&\ supp(\mu) \text{ is bounded in } M\right\} \end{aligned}$$

For $\phi : M \to \mathbb{R}$ denote $\mu(\phi) := \int_M \phi d\mu$. Denote the set of Lipschitz functions $f : M \to \mathbb{R}$ with Lipschitz constant less than or equal to 1 by $Lip_1(M)$. The **Wasserstein metric** d_W on $\mathfrak{F}(M)$ is defined by

$$d_W(\mu,\nu) := \sup\{\mu(\phi) - \nu(\phi) |\ \phi \in Lip_1(M)\}.$$

On $\mathfrak{F}(M)$ define the arc fields

$$V_t(\mu)(x) := \frac{e^{tg}\mu}{\|e^{tg}\mu\|}$$

for $g : M \to \mathbb{R}$ and $Y_t(\mu)(\phi) := \mu(F_t^* \phi)$ where F is any flow on M and F^* denotes the pullback, as defined above.

The bracket $[V,Y]$ is then an arc field on $\mathfrak{F}(M)$ which gives geometric information about the distribution $\Delta(V,Y)$ on $\mathfrak{F}(M)$.

References

[1] R. Abraham, J. Marsden and T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, 2nd Ed., Springer-Verlag, 1988.

[2] J.-P. Aubin, *Mutational and Morphological Analysis*, Birkhäuser, Boston, 1999.

[3] C. Calcaterra, *Arc Fields*, University of Hawaii Doctoral Dissertation, 1999.

[4] C. Calcaterra, *Continuous Dynamics on Metric Spaces*, draft available at www.math.metrostate.edu/craig

[5] C. Calcaterra and D. Bleecker, Generating Flows on a Metric Space, *J. Math. Anal. Appl.*, **248**, 645-677 (2000)

[6] C. Calcaterra, A. Boldt, M. Green, D. Bleecker, Metric Coordinate Systems, *Commun. Math. Anal.* **6** (2009), no. 2, pp. 79–108.

[7] R. M. Colombo and A. Corli, A Semilinear Structure on Semigroups in a Metric Space, *Semigroup Forum*, **68** (2004), pp. 419-444.

[8] R. M. Colombo and A. Corli, Differential Equations in Metric Spaces with Applications, *Discrete Contin. Dyn. Syst.* **23** (2009), no. 3, pp. 733-753.

[9] B. Faxén, On Approximation by Translates and Related Problems in Function Theory, *Ark. Mat.* **19** (1981), no. 2, pp. 271–289.

[10] D. Motreanu and N. H. Pavel, *Tangency, Flow Invariance for Differential Equations, and Optimization Problems*, Marcel Decker, 1999.

[11] A. I. Panasyuk, Quasidifferential Equations in a Complete Metric Space under Conditions of the Caratheodory Type. I, *Differ. Uravn.* **31** (1995), pp. 901-910.

[12] A. I. Panasyuk, Quasidifferential Equations in a Complete Metric Space under Caratheodory-type Conditions. II, *Differ. Uravn* **31,** no. 8 (1995), pp. 1308-1317.

[13] F. Rampazzo and H. J. Sussman, Commutators of Flow Maps of Nonsmooth Vector Fields, (2006), *J. Differential Equations* (to appear).

[14] S. Simić, Lipschitz Distributions and Anosov Flows. *Proc. Amer. Math. Soc.* **124** (1996), no. 6, pp. 1869-1877.

[15] E. D. Sontag, *Mathematical Control Theory*, 2nd Ed., Springer-Verlag, 1998.

[16] R. A. Zalik, The Müntz-Szász Theorem and the Closure of Translates, J. Math. Anal. Appl. **82** (1981), no. 2, pp. 361–369.