3.9 Differential equations and local flows

Vector fields are intimately related to differential equations. But we argue that as *functions* with clearly defined domain and codomain, they are mathematically both more precise and more versatile. E.g. are the equations $\dot{y} = y$, $\dot{y} - y = 0$, or $\frac{\dot{y}}{y} = 1$ *"the same"*?. To translate between these languages, consider in classical terminology a system of ordinary differential equations on $\mathbb{R}^n$

$$\begin{aligned} \dot{x}_1 &= f_1(x) \\ \dot{x}_2 &= f_2(x) \\ &\vdots \\ \dot{x}_n &= f_n(x). \end{aligned} \tag{72}$$

One says that *"a solution of this system of differential equations is a function $x(t)$ that* satisfies *the system, i.e. makes it true when $x(t)$ is substituted for x ..."* Clearly there is nothing lost when going from writing out equations with symbols such as $\dot{x}_i$ to focusing on the vector field $X = \sum_{i=1}^n f^i D_i$ (we change the subscripts to traditional superscripts of differential geometry). Then the notion of *a solution* of a differential equation is made precise:

Definition 3.13 *Suppose $X \in \Gamma^\infty(M)$ is a smooth vector field on a smooth manifold M. An* integral curve *of X is a curve $\sigma\colon I \mapsto M$* (for some interval $I \subseteq \mathbb{R}$) *such that for all $t \in I$*

$$\sigma_{t*}\big(\tfrac{d}{dt}\big|_t\big) = (X \circ \sigma)(t), \quad \textit{or in shorthand notation: } \dot{\sigma} = X \circ \sigma. \tag{73}$$

To connect this with the notion of tangent vectors as first order partial differential operators note that this definition is equivalent to saying that σ is an integral curve if for every point $p \in M$ and every smooth function $\phi \in C^\infty(p)$ defined in some neighborhood of p

$$\tfrac{d}{dt}\,(\phi \circ \sigma) = (X\phi) \circ \sigma. \tag{74}$$

In particular, in a chart (u, U) taking $\phi = u^k$ yields $\frac{d}{dt}(u^k \circ \sigma) = (Xu^k) \circ \sigma$.

For later reference we formally distinguish a vector field along a curve from the restriction of a vector field on the manifold to the image of a curve. In the case that the curve is an imbedding, one easily can go back and forth. However, if the curve is not one-to-one, a vector field along a curve need not arise from a vector field on the manifold.

Definition 3.14 *Suppose $\sigma\colon I \mapsto M$ is a smooth curve. A smooth function $Y\colon I \mapsto TM$ such that $\pi \circ Y = \sigma$ is called a vector field along the curve σ.*

Note that if $X\colon M \mapsto TM$ is a smooth vector field on manifold M, and $\sigma\colon I \mapsto M$ is a smooth curve, then the composition $X \circ \sigma$ is a vector field along the curve σ. However, if $Y\colon I \mapsto TM$ is a vector field along the curve $\sigma\colon I \mapsto M$, then there need not exist any smooth vector field X on M such that Y arises from X in the afore described way.

Exercise 3.31 *Discuss vector fields along curves as opposed to vector fields on manifolds in the special cases where the curve is not one-to-one, or is an immersion, but not an imbedding. Are there any further problems with vector fields along curves in nonorientable manifolds such as the Möbius strip? Suggest conditions under which a smoooth vector field along a curve* extends *to a smooth vector field on a neighborhood of the (image of the) curve.*

More interesting than just isolated integral curves is the collection of all integral curves of a vector field. However, in general some technical difficulties arise due to the common lack of common domains. We start with a preliminary notion of *solutions* as *parameterized families* of integral curves, parameterized by the initial conditions.

Definition 3.15 (Preliminary definition) *The/a* flow *of a vector field $X\colon M \mapsto TM$ is a function Φ defined on a suitable (maximal ?) subset of $\mathbb{R} \times M$ with values in M that satisfies*

$$\begin{cases} \textit{for all } p \in M, & \Phi(0,p) = p \\ \textit{for all } f \in C^\infty(M), & \frac{d}{dt}(f \circ \Phi) = (Xf) \circ \Phi \end{cases} \tag{75}$$

When *holding the initial condition $\Phi(0,p) = p \in M$ fixed*, the map $\Phi(\cdot,p)\colon t \mapsto \Phi(t,p)$ is an integral curve of X. On the other hand, when holding the time $t \in \mathbb{R}$ fixed the map $\Phi(t,\cdot)\colon M \mapsto M$ is a map between manifolds. In different places either point of view may be more suitable. Often the critical step is to mix the roles of the *time* and *space* variables, e.g. consider the *flow-box theorem* 3.17 and the geodesic (normal) coordinates (or the exp-map) (theorem **??**). The basic technical questions concern the existence and uniqueness of a *maximal* flow (i.e. with maximal domain), and about its smoothness. A few examples will clearly delineate the limitations – most importantly, for a generic vector field, no *global flow* with all desirable properties need to exist. On the other hand, locally there is hardly any difference between flows on Euclidean spaces and on manifolds. Thus the remainder of this section provides a brief review of basic properties and background material of ordinary differential equations in $\mathbb{R}^n$.

Exercise 3.32 *For each $p \in \mathbb{R}$ solve the initial value problem $\dot{y} = y^2$, and $\dot{y} = 1 + y^2$, $y(0) = x$. In each case, precisely describe the maximal domain of the* flow.

The phenomenon exemplified by the solution curves in the preceding exercise is termed *finite escape time*.

Exercise 3.33 *Verify that for each pair of* times $t_1, t_2 \in \mathbb{R}$ *with $t_1 \leq t_2$ the function*

$$y(t) = \begin{cases} -(t_1 - t)^{3/2} & \textit{if } t \leq t_1 \\ 0 & \textit{if } t_1 \leq t \leq t_2 \\ (t - t_2)^{3/2} & \textit{if } t \geq t_2 \end{cases} \tag{76}$$

is a solution *of the differential equation $\dot{y} = \frac{3}{2} y^{1/3}$*
Precisely describe the maximal domain of the flow *– pay special attention to the solutions which are not of the form above, i.e. with one or both of $t_1 = -\infty$ and $t_2 = \infty$.*

Exercise 3.34 *Let $f\colon \mathbb{R} \mapsto \mathbb{R}$ be defined by $f(x) = 1$ of $x \geq 0$ and $f(x) = -1$ else. Discuss and suggest notions of* solutions *for the differential equations $\dot{x} = f(x)$ and $\dot{x} = -f(x)$. Focus on the maximal domains on which solutions* make sense*, and on initial conditions $x(t_0) = 0$.*

Project 3.35 *Prove that for every initial condition $y(0) = x \in \mathbb{R}$ the differential equation $\dot{y} = 1 + y^{1/3}$ has a unique solution.* [[There are lots of less-known conditions in the classical literature for the uniqueness of solutions to initial value problems which are much weaker than the traditionally taught *Lipschitz*-conditions.]] *Hint: Formally interchange the roles of t and y and discuss why the corresponding new differential equation has the* same *solutions. How does this procedure work in the case of a system like $\dot{x} = -y$, $\dot{y} = x + y^{1/3}$?*

The examples in these exercises illustrate that if the vector field (alas differential equation) is not sufficiently smooth then even locally neither existence and nor uniqueness of solutions are guaranteed. Moreover, without stringent *growth conditions* there is no hope for global existence. The standard existence and uniqueness result for solutions of ordinary differential equation relies on the Gronwall inequality and Picard iteration.

Lemma 3.12 (Gronwall inequality) *Suppose $f\colon [0,T] \mapsto [0,\infty)$ is continuous and $C, K \geq 0$ are such that for all $t \in [0,T]$, $f(t) \leq C + \int_0^t Kf(s)\,ds$. Then for all $t \in [0,T]$, $f(t) \leq Ce^{Kt}$.*

Proof. Suppose $f\colon [0,T] \mapsto [0,\infty)$ is continuous and $C, K \geq 0$ are such that for all $t \in [0,T]$, $f(t) \leq C + \int_0^t Kf(s)\,ds$. Define $u\colon [0,T] \mapsto [0,\infty)$ by $u(t) = C + \int_0^t Kf(s)\,ds$. Then u is continuously differentiable, and $0 \leq u' = Kf \leq u$ and hence for all $t \in [0,T]$, $\frac{d}{dt}(u(t)e^{-Kt}) \leq 0$, and, $0 \leq u(t) \leq u(0)e^{Kt}$ from which the desired inequality follows readily.■

Definition 3.16 *Suppose $U \subseteq \mathbb{R}^m$ is open. A function $f\colon U \mapsto \mathbb{R}^n$ is called* locally Lipschitz continuous *on U of for every $p \in U$ there exists an open neighborhood $W \subseteq U$ of p and a constant L, called the* Lipschitz constant*, such that for all $q, q' \in W$*

$$\|f(q) - f(q')\|_{\mathbb{R}^n} \leq L \cdot \|q - q'\|_{\mathbb{R}^m}. \tag{77}$$

Theorem 3.13 (Picard Lindelöf) *Let $U \subseteq \mathbb{R}^m$ be a connected open set and suppose $f\colon U \mapsto \mathbb{R}^m$ is locally Lipschitz continuous. Then there exists an open set $W \subseteq \mathbb{R} \times U$ such that for each fixed $y \in U$ the set $\{t \in \mathbb{R} : (t,y) \in W\}$ is an interval $(a(y), b(y))$, and there exists a unique continuous map $\Psi\colon W \mapsto U$ with maximal domain W such that*

$$\begin{aligned} \text{for all } (t,y) \in W, \qquad \tfrac{\partial}{\partial t}\Psi(t,y) &= f(\Psi(t,y)) \\ \text{for all } y \in U, \quad \Psi(0,y) &= y. \end{aligned} \tag{78}$$

Definition 3.17 *The map $\Psi\colon W \mapsto U$ of the preceding theorem is called the* flow *of the ordinary differential equation $\dot{y} = f(y)$.*

The basic idea behind the proof is to use Picard-iteration, that is to transform the differential equation into an integral equation that is solved by successive iteration:

$$\begin{aligned} y_0(t) &= y_0 \\ y_{k+1}(t) &= y_0 + \textstyle\int_0^t f(s, y_k(s))\,ds. \end{aligned} \tag{79}$$

Proof still to be typeset....

The standard argument uses that for sufficiently small $\varepsilon > 0$ the map $y_n \mapsto y_{n+1}$ on the metric space $C^0([t_0, t_0 + \varepsilon])$ is a *contraction*, and hence must have a fixed point.

With some additional work one may establish some additional regularity. The basic idea is that due to the integration in the Picard-iteration scheme the function y_{k+1} is at least one degree smoother than the lesser of the degrees of smoothness of y_n and f. leading to a *boots-trapping argument* ... But with all technical details this is a fairly laborious proof ...

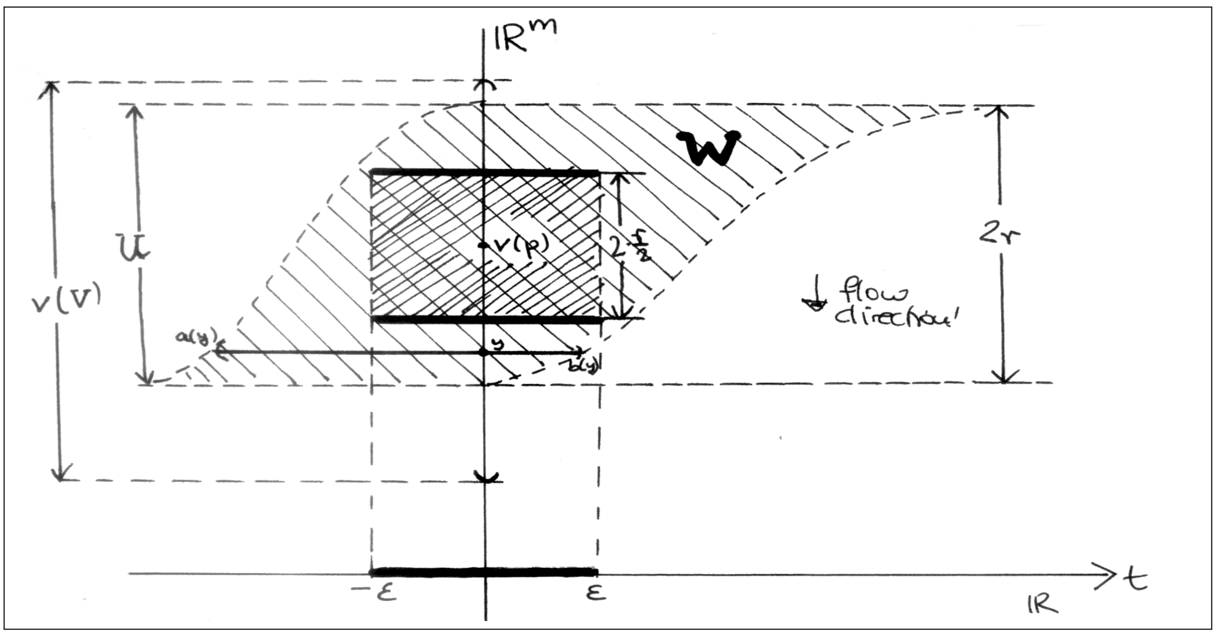
Proof still to be typeset....

Theorem 3.14 (Smoothness of the flow) *Suppose f, U, W, Ψ are as in the preceding theorem. Then if $f \in C^r(U, \mathbb{R}^n)$ for $1 \leq r \leq \infty$ then $\Psi \in C^r(W, U)$ (i.e. Ψ is C^r in* both *arguments).*

Exercise 3.36 *Carry out the Picard iteration scheme for the initial value problem $\dot{y} = y$, $y(0) = c \in \mathbb{R}$. Explicitly calculate the first three iterates, and by induction find a formula for the n-th iterate. Verify that this sequence of functions converges globally (i.e. for all $t \in \mathbb{R}$ to an analytic function which is* the *solution of the initial value problem.*

Exercise 3.37 *Carry out the Picard iteration scheme for the initial value problem $\dot{y}_1 = -y_2$, $\dot{y}_2 = y_1$, $y(0) = (c_1, c_2) \in \mathbb{R}^2$, e.g. explicitly calculate the first five iterates, and by induction find a formula for the n-th iterate. Verify that this sequence of functions converges globally (i.e. for all $t \in \mathbb{R}$) to an analytic function which is* the *solution of the initial value problem.*

Suitably adapted to our settings of smooth vector fields on manifolds one obtains the following theorem which guarantees that the time interval does not depend on the initial point. This is an important feature that allows one to fix t and consider the (local) diffeomorphism Φ_t.



Theorem 3.15 (Existence of smooth, local flows) *Suppose $X \in \Gamma^\infty(M)$ is a smooth vector field and $p \in M$. Then there exists an open neighborhood $O \subseteq M$ of p, $\varepsilon > 0$ and a one-parameter family of difeommorphisms $\Phi_t \colon O \mapsto \Phi_t(A) \subseteq M$ onto their respective images, parameterized by $t \in (-\varepsilon, \varepsilon)$ such that*

(i) $\Phi \colon (-\varepsilon, \varepsilon) \times O \mapsto M$ *is* C^∞.

(ii) If $|s|$, $|t|$, $|s + t| < \varepsilon$, and $q, \Phi_t(q) \in O$ then $\Phi_s(\Phi_t(q)) = \Phi_{s+t}(q)$ and $\Phi_0(q) = q$.

(iii) For all $q \in O$, X_q is tangent *to the curve $\gamma \colon (-\varepsilon, \varepsilon) \mapsto O$ where $\gamma(t) = \Phi_t(q)$ at $t = 0$.*

Proof still to be typeset....

The proof is a rather straightforward construction, using at some place compactness in order to achieve a constant time interval, i.e. that is the same for all points in some neighborhood.

Proof still to be typeset....

Corollary 3.16 *If $K \subseteq M$ is compact and $X \in \Gamma^\infty M$ vanishes on $M \setminus K$ then X generates a unique one parameter family of diffeomorphisms (i.e. the flow is* global, *defined for all $t \in \mathbb{R}$).*

Most useful for theoretical purposes is the following result about the possibility to *stratify* any vector field at any points where it does not vanish.

Theorem 3.17 (Flow box theorem) *Suppose $X \in \Gamma^\infty(M)$ and $p \in M$ is such that $X_p \neq 0$. Then there exists a chart (u, U) about p, w.l.o.g. $u(p) = 0$, such that*

$$\text{for all} \quad q \in U, \quad \mathcal{X}\big|_U = \tfrac{\partial}{\partial u^1} \quad \text{i.e.} \quad X_q = \tfrac{\partial}{\partial u^1}\big|_q. \tag{80}$$

Note that this effectively says that modulo a local coordinate change every (nonlinear!) system of differential equations is in the neighborhood of any regular point locally equivalent to the system

$$\begin{aligned} \dot{x}_1 &= 1 \\ \dot{x}_2 &= 0 \\ &\vdots \\ \dot{x}_n &= 0. \end{aligned} \tag{81}$$

Of couse, in general there is no hope of finding explicit, so called *"closed-form"* solutions for nonlinear differential equations – which is rally synonymous with finding a *formula* for the coordinates u, or more practically a change of coordinates to (u, U).

Proof. Start with any coordinates (w, W) about p and suppose $\dot{w}(X_p) = (a^1, \ldots, a^m) = a \in \mathbb{R}^m$, and $w(p) = 0$. Without loss of generality we may assume that $a^1 \neq 0$ (else permute the coordinates in w). Define a new coordinate chart $(v, V = W)$ by setting for $q \in W$

$$v^1(q) = \tfrac{1}{a^1} w^1(q), \text{ and for } j \geq 2, \quad v^j(q) = w^j(q) - \tfrac{a^j}{a^1} w^1(q). \tag{82}$$

Verify that this constant linear coordinate change aligns the new coordinates with the vector field at the point p, i.e., $X_p = \frac{\partial}{\partial v^1}\big|_p$.

$$\left.\frac{\partial}{\partial v^1}\right|_p = \sum_{j=1}^m \left.\frac{\partial w^j}{\partial v^1}\right|_p = \left.\frac{\partial}{\partial w^j}\right|_p = \sum_{j=1}^m a^j \left.\frac{\partial}{\partial w^j}\right|_p = X_p. \tag{83}$$

In a second step, use the local flow $\Phi\colon (-\varepsilon, \varepsilon) \times \tilde{U} \mapsto M$ of X defined on some neighborhood $\tilde{U} \subseteq V$ of p, to define a map $u^{-1}\colon v(\tilde{U}) \mapsto M$ by

$$u^{-1}(a^1, a^2, \ldots a^m) = \Phi(a^1, v^{-1}(0, a^2, a^3, \ldots a^m)). \tag{84}$$

This map u^{-1} is well-defined on some open set $v(\tilde{U})$ about $(0, v(p) \subseteq \mathbb{R}^{m+1})$, and it is clearly C^∞, but we need to verify that it is invertible (possibly after further restricting its domain), i.e. verify that there exists a map u such that $(u^{-1})^{-1} = u$ indeed defines coordinates on some neighborhood of p. First calculate for $a \in v(\tilde{U}) \subseteq \mathbb{R}^m$ and $f \in C^\infty(v(\tilde{U}))$

$$\begin{aligned} \left((u^{-1})_* D_1|_a\right)(f) &= \lim_{h\to 0} \tfrac{1}{h}\left(f \circ u^{-1}(a^1 + h, a^2, \ldots a^m) - f \circ u^{-1}(a)\right) \\ &= \lim_{h\to 0} \tfrac{1}{h}\left(f \circ \Phi_{a^1+h} \circ v^{-1}(0, a^2, \ldots a^m) - f \circ u^{-1}(a)\right) \\ &= \lim_{h\to 0} \tfrac{1}{h}\left((f \circ \Phi_h)(u^{-1}(a)) - f(u^{-1}(a))\right) \\ &= X_{u^{-1}(a)} f. \end{aligned} \tag{85}$$

which shows that for all $a \in v(\tilde{U})$, $\left((u^{-1})_* D_1\right|_a = X_{u^{-1}(a)}$. Similarly, for $k > 1$ calculate

$$\begin{aligned}
\left((u^{-1})_{*0}\, D_k|_0\right)(f) &= \lim_{h\to 0} \tfrac{1}{h}\left(f \circ u^{-1}(0,\ldots,0,h,0,\ldots 0) - f \circ u^{-1}(0)\right) \\
&= \lim_{h\to 0} \tfrac{1}{h}\left(f \circ \circ v^{-1}(0,\ldots,0,h,0,\ldots 0) - f \circ v^{-1}(0)\right) \\
&= \lim_{h\to 0} \tfrac{1}{h}\left(f \circ u^{-1}(0,\ldots,0,h,0,\ldots 0) - f \circ v^{-1}(0)\right) \\
&= D_k|_0\,(f \circ v^{-1}) = \left.\tfrac{\partial f}{\partial v^k}\right|_p .
\end{aligned} \tag{86}$$

Therefore for all $k \geq 1$, $\left((u^{-1})_{*0} D_k\right|_0 = \left.\frac{\partial}{\partial v^k}\right|_p$ showing that $(u^{-1})_{*0} \colon T_0\mathbb{R}^m \mapsto T_pM$ has full rank, indeed, $v_{*p} \circ (u^{-1})_{*0}$ is the identity on T_pM. By the inverse function theorem, there exist open neighborhoods $U \subseteq \tilde{U}$ of p and $Z \subseteq v(U)$ of 0 such that $u^{-1}|_Z \colon Z \mapsto U$ is a bijection and

$$u \stackrel{\text{def}}{=} (u^{-1}\big|_Z)^{-1} \colon (U,p) \mapsto (\mathbb{R}^m, 0) \tag{87}$$

defines a local coordinate chart about p in which $X \equiv \frac{\partial}{\partial u^1}$. ■

Example 3.2 *Consider $M = \mathbb{R}^2$ with standard coordinates $((x_1,x_2),\mathbb{R}^2)$, the vector field $X = \frac{\partial}{\partial x_1} + 2x_1 \frac{\partial}{\partial x_1}$ and the point $p = (0,0)$. At p, $X_p = \left.\frac{\partial}{\partial x_1}\right|_p$, i.e. no preliminary constant linear coordinate change is needed. To calculate the flow of X solve the initial value problems*

$$\left\{\begin{array}{lclclcl} \dot{x}_1 &=& 1 & \qquad & x_1(0) &=& a \\ \dot{x}_2 &=& 2x_1 & & x_1(0) &=& b \end{array}\right. \tag{88}$$

yielding the formula for the flow

$$\Phi_t(a,b) = (a+t, b+(a+t)^2 - a). \tag{89}$$

To define the new coordinates first construct the inverse

$$x = u^{-1}(y_1,y_2) \stackrel{\text{def}}{=} \Phi_{y_1}(0,y_2) = (y_1, y_2 + y_1^2). \tag{90}$$

Verify that in the new coordinates $u = (y_1,y_2) = (x_1, x_2 - x_1^2)$ the vector field is straightened out, indeed, $X = \frac{\partial}{\partial y_1}$ and $\frac{\partial}{\partial y_2} = \frac{\partial}{\partial x_2}$.

The next sections will demonstrate that it is generally impossible to simultaneously straighten out two vector fields, or straighten out what one might call *plane fields.*

3.10 Lie derivatives

The (local) flows obtained in section 3.9 provide some minimal analogue of an additive structure on a manifold. An important application is to use these for *calculus-like*, coordinate free characterization and definitions of Lie derivatives of functions, of vector fields, and in more generality, later of any tensors. Instead of starting with a vector field, it is often convenient to start with a family of diffeomorphisms. Formally define:

Definition 3.18 *A one-parameter family of diffeomorphisms on a manifold M is a smooth map $\Phi\colon \mathbb{R}\times M \mapsto M$ satisfying*

- *For every fixed $t\in\mathbb{R}$ the map $\Phi_t\colon M\mapsto M$, defined by $\Phi_t(p)=\Phi(t,p)$ for $p\in M$, is a diffeomorphism of M.*
- *$\Phi_s\circ\Phi_t=\Phi_{s+t}$ for all $s,t\in\mathbb{R}$.*

Similarly, a local one-parameter family of diffeomorphisms *is a smooth map $\Phi\colon(\varepsilon,\varepsilon)\times U\mapsto M$, defined for some $\varepsilon>0$ and some open set $U\subseteq M$, which satisfies*

- *For every fixed $|t|<\varepsilon$ the map $\Phi_t\colon U\mapsto \Phi_t(U)\subseteq M$, is a diffeomorphism of U onto its image.*
- *$\Phi_s\circ\Phi_t=\Phi_{s+t}$ for all s,t such that $|s|,|t|,|s+t|<\varepsilon$.*

It is easily recognized that such (local) one-parameter families of diffeomorphism arise is (local) flows of smooth vector fields X. In turn, given such a family, one can *recover* the vector field that locally generates this family as its flow, and one calls the *infinitesimal generator* of the family Φ of (local) diffeomorphism. This is based on the observation that one may characterize the *derivative* Xf of a function $f\in C^\infty(M)$ in the direction of a vector field X entirely in terms of the flow of X:

Definition 3.19 *The Lie derivative L_Xf of a smooth function $f\in C^\infty(M)$ with respect to a vector field $X\in\Gamma^\infty(M)$ with flow Φ is defined for $p\in M$ as*

$$(L_Xf)(p)=\lim_{h\to 0}\tfrac{1}{h}\Big((f\circ\Phi_h)(p)-f(p)\Big). \tag{91}$$

One readily sees that this *Lie derivative* L_Xf is the same as Xf – but the notation L_Xf is still very suggestive and nicely carries over to *Lie derivatives* of other objects:

Definition 3.20 *The Lie derivatives L_XY and $L_X\omega$ of a smooth (tangent) vector field $Y\in\Gamma^\infty(M)$ and of a smooth co-tangent vector field $\omega\in\Omega^1(M)$ with respect to $X\in\Gamma^\infty(M)$ with flow Φ are defined (at any point $p\in M$):*

$$\begin{array}{rcl}(L_XY)_p &=& \lim\limits_{h\to 0}\tfrac{1}{h}\Big(Y_p-(\Phi_{h*}Y)_p\Big)\\ (L_X\omega)_p &=& \lim\limits_{h\to 0}\tfrac{1}{h}\Big((\Phi_h^*\omega)_p-\omega_p\Big).\end{array} \tag{92}$$

Proposition 3.18 *Suppose $f\in C^\infty(M)$, $X,Y,Z\in\Gamma^\infty(M)$, and $\omega,\eta\in\Omega^1(M)$. Then*

$$\begin{array}{lrcl}\text{(i)} & L_X(Y+Z) &=& L_XY+L_XZ\\ \text{(ii)} & L_X(\omega+\eta) &=& L_X\omega+L_X\eta\\ \text{(iii)} & L_X(fY) &=& (L_Xf)\cdot Y+f\cdot L_XY\\ \text{(iv)} & L_X(f\omega) &=& (L_Xf)\omega+f\cdot(L_X\omega)\\ \text{(v)} & L_X(\omega(Y)) &=& (L_X\omega)(Y)+\omega(L_XY)\end{array}\quad . \tag{93}$$

The first two parts are immediate consequences of the $\mathbb{R}$-linearity of the maps Φ_{h*} and Φ_h^*. We will prove the product rules (iii) and (v) and leave (iv) as an exercise.

Proof (of part (iii)). Suppose $f \in C^\infty(M)$, $X, Y \in \Gamma^\infty(M)$, $p \in M$ and let Φ be the local flow of X defined in a neighborhood of p. Using basic properties of the tangent maps Φ_{h*}, especially the linearity, and, as expected for product rules, add and subtract suitable terms:

$$\begin{aligned} (L_X(fY))_p &= \lim_{h\to 0} \tfrac{1}{h}\Big((fY)_p - (\Phi_{h*}(fY))_p\Big) \\ &= \lim_{h\to 0} \tfrac{1}{h}\Big(f(p)\cdot Y_p - \Phi_{h*\Phi_{-h}(p)}\big((f\circ\Phi_{-h})(p)\cdot Y_{\Phi_{-h}(p)}\big)\Big) \\ &= \lim_{h\to 0} \tfrac{1}{h}\Big(f(p)Y_p - (f\circ\Phi_{-h})(p)\cdot \Phi_{h*\Phi_{-h}(p)}Y_{\Phi_{-h}(p)}\Big) \\ &= \lim_{h\to 0} \tfrac{1}{h}\Big(f(p) - (f\circ\Phi_{-h})(p)\Big)\cdot Y_p \\ &\qquad\qquad + \lim_{h\to 0}(f\circ\Phi_{-h})(p)\cdot \lim_{h\to 0}\tfrac{1}{h}\left(Y_p - (\Phi_{h*}Y)_p\right) \\ &= (L_Xf)(p)\cdot Y_p + f(p)\cdot(L_XY)_p. \end{aligned} \tag{94}$$

As long as all the limits exist there is no problem with breaking them up. Moreover, from continuity of f at p it follows that $\lim\limits_{h\to 0}(f\circ\Phi_{-h})(p) = f(p)$. ■

Instrumental for the proofs of part (iv) and (v) of proposition 3.18 is the observation:

Lemma 3.19 *Suppose $\omega \in \Omega^1(M)$, $p \in M$ and Φ is the local flow of $X \in \Gamma^\infty(M)$ near p. Then*

$$\lim_{h\to 0}(\Phi_h^*\omega)_p = \omega_p \tag{95}$$

Proof. Suppose that (u, U) is a chart about p. Then for $|h|$ sufficiently small $\Phi_h(p) \in U$ and (69) (with $v = u$) yields

$$(\Phi_h^*\omega)_p = \sum_{i=1}^n \omega_i(\Phi_h(p))\Phi_{hp}^*\left(du^i_{\Phi_h(p)}\right) = \sum_{j=1}^m \Big(\sum_{i=1}^n \omega_i(\Phi_h(p))\,\frac{\partial(u^i\circ\Phi_h)}{\partial u^j}\Big|_p\Big)\cdot du^j_p. \tag{96}$$

Since ω is smooth, clearly $\lim\limits_{h\to 0}\omega_i(\Phi_h(p)) = \omega_i(p)$. Regarding the second factor, since Φ is smooth in both variables, and Φ_0 is the identity map on some neighborhood of p, it follows that $\lim\limits_{h\to 0}\frac{\partial(u^i\circ\Phi_h)}{\partial u^j}|_{\Phi_h(p)} = \delta^i_j$. [[Think carefully about this limit!]] ■

Exercise 3.38 *Prove the product rule (iv) of proposition 3.18 (use lemma 3.19).*

Proof (of part (v) of proposition 3.18). Let $X, Y \in \Gamma^\infty(M)$, $\omega \in \Omega^1(M)$ $p \in M$ and let Φ be the local flow of X defined in a neighborhood of p.

$$\begin{aligned}
L_X(\omega(Y))(p) &= \lim_{h\to 0} \tfrac{1}{h}\Big((\omega(Y))\circ\Phi_h(p) - (\omega(Y))(p)\Big)\\
&= \lim_{h\to 0} \tfrac{1}{h}\Big(\omega_{\Phi_h(p)}(Y_{\Phi_h(p)}) - \omega_p(Y_p)\Big)\\
&= \lim_{h\to 0} \tfrac{1}{h}\Big(\omega_{\Phi_h(p)}\Big(\ \Phi_{h*p}\circ\Phi_{-h*\Phi_h(p)}Y_{\Phi_h(p)}\ \Big) - \omega_p(Y_p)\Big)\\
&= \lim_{h\to 0} \tfrac{1}{h}\Big((\Phi_h^*\omega)_p(\Phi_{-h*}Y)_p - \omega_p(Y_p)\Big)\\
&= \lim_{h\to 0} \Big(\tfrac{1}{h}\Big(\ (\Phi_h^*\omega)_p - \omega_p\ \Big)(Y_p) + (\Phi_h^*\omega)_p\Big(\ \tfrac{1}{-h}\Big(\ Y_p - (\Phi_{-h*}Y)_p\ \Big)\Big)\Big)\\
&= \Big(\lim_{h\to 0} \tfrac{1}{h}\Big(\ (\Phi_h^*\omega)_p - \omega_p\ \Big)\Big)(Y_p)\\
&\qquad + \Big(\lim_{h\to 0}(\Phi_h^*\omega)_p\Big)\Big(\lim_{h\to 0}\Big(\tfrac{1}{-h}\Big)\Big(\ Y_p - (\Phi_{-h*}Y)_p\ \Big)\Big)\\
&= (L_X\omega)_p(Y_p) + \omega_p(L_XY)_p \qquad . \qquad \blacksquare
\end{aligned} \tag{97}$$

The next major goal is to prove that the Lie derivative of a vector field is the same as the Lie bracket defined earlier: $L_XY = [X,Y]$. Given the obvious anti-symmetry $[Y,X] = -[X,Y]$ of the Lie bracket this sheds some important light on the Lie derivative which at first view seems to assign very different roles to the vector fields X and Y, evaluating Y along the integral curves of X. We will eventually provide an alternative coordinate-free proof, but first will use this as an opportunity to figure out how to calculate Lie derivatives in coordinates.

Lemma 3.20 *Suppose (u,U) is a coordinate chart and $X \in \Gamma^\infty(M)$. Then*

$$L_X(du^i) = \sum_{j=1}^m \tfrac{\partial(Xu^i)}{\partial u^j}\, du^j. \tag{98}$$

Proof. Suppose (u,U) is a coordinate chart about $p \in M$ and Φ is the local flow of $X \in \Gamma^\infty(M)$ near p. For sufficiently small $|h|$ we may assume that $\Phi_h(p) \in U$. The key step is the interchange of the order of differentiation in time and space directions, and thus requires the use of the definition of X as the generator of the flow Φ.

$$\begin{aligned}
\big(L_X(du^i)\big)_p &= \lim_{h\to 0}\tfrac{1}{h}\Big(\ \big(\Phi_h^*(du^i)_p - (du^i)\big)_p\ \Big)\\
&= \lim_{h\to 0}\tfrac{1}{h}\sum_{j=1}^m\Big(\ \tfrac{\partial(u^i\circ\Phi_h)}{\partial u^j}\Big|_p\ (du^j)_p - \delta^i_j\,(du^j)_p\ \Big)\\
&= \sum_{j=1}^m \lim_{h\to 0}\tfrac{1}{h}\Big(\ \tfrac{\partial(u^i\circ\Phi_h)}{\partial u^j}\Big|_p - \tfrac{\partial(u^i\circ\Phi_0)}{\partial u^j}\Big|_p\ \Big)\ (du^j)_p\\
&= \sum_{j=1}^m \lim_{h\to 0}\tfrac{1}{h}\Big(\ \tfrac{\partial}{\partial u^j}(u^i\circ\Phi)\Big|_{(h,p)} - \tfrac{\partial}{\partial u^j}(u^i\circ\Phi)\Big|_{(0,p)}\ \Big)\ (du^j)_p\\
&= \sum_{j=1}^m \tfrac{\partial}{\partial t}\tfrac{\partial}{\partial u^j}(u^i\circ\Phi)\Big|_{(0,p)}\ (du^j)_p\\
&= \sum_{j=1}^m \tfrac{\partial}{\partial u^j}\,\tfrac{\partial}{\partial t}(u^i\circ\Phi)\Big|_{(0,p)}\ (du^j)_p\\
&= \sum_{j=1}^m \tfrac{\partial(Xu^i)}{\partial u^j}(du^j)_p \qquad . \qquad \blacksquare
\end{aligned} \tag{99}$$

Rather than performing a similar direct calculation of the Lie derivative $L_X\left(\frac{\partial}{\partial u^j}\right)$, we use the product rule of part (v) of proposition 3.18, and observe

$$0 = L_X\delta^i_j = L_X\Big((du)^i \Big(\tfrac{\partial}{\partial u^j} \Big) \Big) = \big(L_X(du)^i\big) \Big(\tfrac{\partial}{\partial u^j} \Big) + (du^i) \Big(L_X \tfrac{\partial}{\partial u^j} \Big) \tag{100}$$

Using the lemma 3.20 this establishes:

Corollary 3.21 *Suppose (u, U) is a coordinate chart and $X \in \Gamma^\infty(M)$. Then*

$$L_X \tfrac{\partial}{\partial u^j} = \sum_{j=1}^{m} - \tfrac{\partial(Xu^i)}{\partial u^j} \tfrac{\partial}{\partial u^i} \tag{101}$$

We are now ready to state and give a direct proof of the theorem:

Theorem 3.22 *If $X, Y \in \Gamma^\infty(M)$ then $L_XY = [X, Y]$.*

Corollary 3.23 *If $X, Y \in \Gamma^\infty(M)$ and $f \in C^\infty(M)$ then $L_{fX}Y = fL_XY - (L_Yf)X$.*

Proof. Work locally in a coordinate chart (u, U) and use the corollary 3.23.

$$\begin{aligned} L_XY &= L_X \Big(\sum_{j=1}^{m} (Yu^j) \tfrac{\partial}{\partial u^j} \Big) \\ &= \sum_{j=1}^{m} \Big(L_X(Yu^j) \tfrac{\partial}{\partial u^j} + (Yu^j) L_X \Big(\tfrac{\partial}{\partial u^j} \Big) \Big) \\ &= \sum_{j=1}^{m} \Big((XYu^j) \tfrac{\partial}{\partial u^j} + (Yu^j) \textstyle\sum_{i=1}^m \Big(- \tfrac{\partial(Xu^i)}{\partial u^j} \tfrac{\partial}{\partial u^i} \Big) \Big) \\ &= \sum_{j=1}^{m} \Big((XYu^j) \tfrac{\partial}{\partial u^j} - \textstyle\sum_{i=1}^m \Big(\underbrace{(Yu^j) \tfrac{\partial}{\partial u^j}}_{Y} Xu^i \Big) \tfrac{\partial}{\partial u^i} \Big) \\ &= \sum_{j=1}^{m} (XY - YX)u^j \tfrac{\partial}{\partial u^j} \quad = \quad [X, Y]. \qquad \blacksquare \end{aligned} \tag{102}$$

Proof (coordinate-free version, following Spivak vol.1, p.213).
Suppose $p \in M$, $f \in C^\infty(p)$, and Φ is a local flow of X defined in a neighborhood of p. Begin with some preliminary constructions. For $|\tau|$ sufficiently small and q sufficiently near p define

$$g(\tau, q) = \int_0^1 \tfrac{\partial(f\circ\Phi)}{\partial t}(s\tau, q)\, ds. \tag{103}$$

Integration of $\tau g(\tau, q)$ yields

$$\begin{aligned} \tau g(\tau, q) &= \textstyle\int_0^1 \tfrac{\partial(f\circ\Phi)}{\partial t}(s\tau, q)\, \tau ds \\ &= (f \circ \Phi)(s\tau, q) \Big|_{s=0}^{s=1} \\ &= (f \circ \Phi_\tau)(q) - f(q). \end{aligned} \tag{104}$$

Consequently $(f \circ \Phi_t)(q) = f(q) + tg_t(q)$, and, moreover,

$$g_0(q) = \int_0^1 \tfrac{\partial(f\circ\Phi)}{\partial t}(0\tau, q)\, ds = \tfrac{\partial(f\circ\Phi)}{\partial t}|_{(0,q)} = (X_qf) = (Xf)_q. \tag{105}$$

This allows us to write

$$(\Phi_{h*}Y)\, f = Y_{\Phi_{-h}(p)}(f \circ \Phi_h) = Y_{\Phi_{-h}(p)}f + h\, Y_{\Phi_{-h}(p)}g_h. \tag{106}$$

Finally we are ready to use these ingredients to prove the theorem

$$\begin{aligned}
(L_X Y)_p f &= \lim_{h\to 0} \tfrac{1}{h}\Big(Y_p - (\Phi_{h*}Y)_p\Big)\, f \\
&= \lim_{h\to 0} \tfrac{1}{h}\Big((Y_p f) - Y_{\Phi_{-h}(p)} f - h\, Y_{\Phi_{-h}(p)} g_h\Big) \\
&= \lim_{h\to 0} \Big(\tfrac{1}{-h}\Big)\Big(((Yf)\circ\Phi_{-h})(p) - (Yf)(p)\Big) - \lim_{h\to 0}((Yg_h)\circ\Phi_{-h})(p) \\
&= (L_X(Yf) - Y(Xf))(p) \quad = \quad [X,Y]_p f \quad . \quad \blacksquare
\end{aligned} \tag{107}$$

Corollary 3.24 *If $X \in \Gamma^\infty(M)$ then $L_X X \equiv 0$.*

Proof. This is an immediate consequence of $L_X X = [X,X] = -[X,X]$, using the anticommutativity of the Lie bracket. However, as a useful exercise in working with the definitions we provide a direct proof in terms of flows. Thus let $p \in M$, $f \in C^\infty(p)$ and let Φ be the local flow of X defined near p. Then

$$(L_X X)_p f = \lim_{h\to 0} \tfrac{1}{h}\Big((X_p f - (\Phi_{h*}X)_p\, f\Big)$$

We will show directly that the difference inside the limit is identically zero.

$$\begin{aligned}
(\Phi_{h*}X)_p\, f &= \big(\Phi_{h*\Phi_{-h}(p)} X_{\Phi_{-h}(p)}\big)\, f \\
&= X_{\Phi_{-h}(p)}(f\circ\Phi_h) \\
&= \lim_{s\to 0} \tfrac{1}{s}\Big(\,((f\circ\Phi_h)\circ\Phi_s)\,(\Phi_{-h}(p)) - (f\circ\Phi_h)\,(\Phi_{-h}(p))\Big) \\
&= \lim_{s\to 0} \tfrac{1}{s}\Big((f\circ\Phi_s)(p) - f(p)\Big) \\
&= X_p f \quad . \qquad \blacksquare
\end{aligned} \tag{108}$$

Exercise 3.39 *Suppose that $X, Y \in \Gamma^\omega(M)$ are analytic vector fields, $p \in M$, and Φ is the local flow of X defined near p. For $\varepsilon > 0$ sufficiently small consider the curve $\sigma\colon (-\varepsilon,\varepsilon) \mapsto T_pM$ defined by $\sigma(t) = (\Phi_{t*}Y)_p$. Rewrite the definition 3.20 of the Lie derivative $L_X Y$ in the form $(L_X Y)_p f = \frac{d}{dt}|_{t=0}\,(\Phi_{t*}Y)\, f(p)$ to establish (via repeated differentiation) the Taylor series expansion $\sigma(t) = \sum_{k=0}^\infty \frac{t^k}{k!}(L_X^k Y)_p$.* This is often suggestively written as

$$\Phi_{t*}\circ Y \circ \Phi_{-t} = \sum_{k=0}^{\infty} \frac{t^k}{k!}(\mathrm{ad}_X^k Y)_p \quad \textit{where } (\mathrm{ad}_X^{k+1}, Y) = [X, (\mathrm{ad}_X^k Y)]. \tag{109}$$

[[An easily accessible reference is Theorem A.1 in *Nilpotent and high-order approximations of vector field systems* by H. Hermes SIAM Rev. **33** (1991), no. 2, 238–264. `http://www.jstor.org.ezproxy1.lib.asu.edu/stable/pdfplus/2031143.pdf`]]

Compare also proposition 3.28 for the flow of the vector field Φ_{t*}.

3.11 Lie derivatives and integrability

The main goal of this section is to prove a first *integrability theorem* which asserts that the vanishing of the Lie brackets $[X_i, X_j]$ is also sufficient for the (local) existence of a submanifold such that X_i are tangent vector fields for this submanifold. This complements proposition 3.9 which asserted that the vanishing of the Lie bracket is necessary for the existence of local coordinates such that given vector fields are coordinate vector fields. Applications to controllability and observability will be discussed in a alter section.

Theorem 3.25 *Suppose $p \in M^m$ and $X_i \in \Gamma^\infty(M)$, $i = 1, \ldots k$ are smooth vector fields such that $\{X_{1p}, \ldots X_{kp}\} \subseteq T_pM$ are linearly independent. If the vector fields commute pairwise, i.e. $[X_i, X_j] \equiv 0$, in an open neighborhood of p, then there exists a chart (u, U) about p such that $X_i|_U = \frac{\partial}{\partial u^i}$, $i = 1, \ldots k$.*

Towards the end of this section we shall provide a constructive proof (is it?). The basic idea is to use the flows Φ^j of the vector fields X_j and define a set of local coordinates u^j basically as the inverse of the map $(x^1, \ldots x^k) \mapsto \Phi^k_{x_k} \circ \ldots \circ \Phi^2_{x_2} \circ \Phi^1_{x_2}(p)$. (It is rather straightforward to accommodate the case that $k < m$.) The most interesting facet of this construction is how the *times* spent flowing along each integral curve become *spatial* coordinates. The bulk of this section is devoted to prove some basic facts about Lie derivatives and their flows, which are both illuminating, and helpful in the eventual proof of the theorem.
We begin with some geometrical explorations: We defined the Lie bracket $[X, Y]$ of two vector fields $X, Y \in \Gamma^\infty(M)$ in terms of the action on smooth functions $f \in C^\infty(M)$, i.e. $[X, Y]f = X(Yf) - Y(Xf)$. The following relates the bracket to the commutative properties (or the lack thereof) of the associated flows. More specifically, suppose $p \in M$ and Φ and Ψ are the local flows of X and Y defined near p. It is natural to ask how the Lie bracket $[X, Y]_p$ relates e.g. to how $\Psi_s \circ \Phi_t(p)$ and $\Phi_t \circ \Psi_s(p)$ compare. It turns out to be more convenient to instead compare $\Psi_{-s} \circ \Phi_{-t} \circ \Psi_s \circ \Phi_t(p)$ with p. In the following, let $s = t$ and consider the terminal points of this concatenation of flows for small values of t. Since everything is smooth, the endpoints as a function of t define a smooth curve, which *passes through* p at $t = 0$. The basic observation is that the *tangent direction* of this curve at p is basically the Lie bracket of $[X, Y]_p$.

Proposition 3.26 *Suppose $p \in M$ and $X, Y \in \Gamma^\infty(M)$ generate the local flows Φ and Ψ near p. For $\varepsilon > 0$ sufficiently small define the curve $\sigma\colon (-\varepsilon, \varepsilon) \mapsto M$ by $\sigma(t) = \Psi_{-t} \circ \Phi_{-t} \circ \Psi_t \circ \Phi_t(p)$. Then $\sigma'(0) = 0$ and $\sigma''(0) = 2[X, Y]_p$.*

Note that in general second derivatives do not define tangent vectors – it is only because the first derivative vanishes that the second derivative is a tangent vector! Before proving the proposition, establish the following intermediate step from the Lie bracket as a derivations on functions and to a more symmetric description of the Lie derivative in terms of flows.

Lemma 3.27 *Suppose $p \in M$ and $X, Y \in \Gamma^\infty(M)$ generate the local flows Φ and Ψ near p and $f \in C^\infty(p)$. Then*

$$[X, Y]_p f = \lim_{s,t \to 0} \tfrac{1}{st} \Big((f \circ \Psi_s \circ \Phi_t)(p) - (f \circ \Phi_t \circ \Psi_s)(p) \Big). \tag{110}$$

Proof (of the lemma). Let p, X, Y, Φ, Ψ and f be as in the lemma. Consider the first term in $[X,Y]_p f = X_p(Yf) - Y_p(Xf)$, using the definition for the Lie derivative of a function twice:

$$\begin{aligned} X_p(Yf) &= \lim_{t\to 0} \tfrac{1}{t}\Big((Yf)(\Phi_t(p)) - (Yf)(p)\Big) \\ &= \lim_{t\to 0}\ \tfrac{1}{t}\Big(\lim_{s\to 0}\ \tfrac{1}{s}\Big((f\circ\Psi_s\circ\Phi_t)(p) - (f\circ\Phi_t)(p)\Big) - \lim_{s\to 0}\ \tfrac{1}{s}\Big((f\circ\Psi_s)(p) - f(p)\Big)\Big) \\ &= \lim_{s,t\to 0}\ \tfrac{1}{st}\Big((f\circ\Psi_s\circ\Phi_t)(p) - (f\circ\Phi_t)(p) - (f\circ\Psi_s)(p) + f(p)\Big) \end{aligned}$$

Combine this with the analogous expression for $Y_p(Xf)$, namely

$$Y_p(Xf) = \lim_{s,t\to 0} \tfrac{1}{st}\Big((f\circ\Phi_t\circ\Psi_s)(p) - (f\circ\Psi_s)(p) - (f\circ\Phi_t)(p) + f(p)\Big)$$

to obtain the desired result. ■

Exercise 3.40 *Suppose $f\colon \mathbb{R}\mapsto\mathbb{R}$ is twice continuously differentiable at $a\in\mathbb{R}$. Show that if $f'(a) = 0$ then the usual difference quotient $f''(a) = \lim\limits_{h\to 0}\frac{1}{h^2}(f(a+h) - 2f(a) + f(a-h))$ simplifies to $f''(a) = \lim\limits_{h\to 0}\frac{2}{h^2}(f(a+h) - f(a))$*

Proof (of the proposition). Let p, X, Y, Φ, and Ψ be as in the proposition, and suppose $f \in C^\infty(p)$. Calculate

$$\begin{aligned} \sigma_0' f = (f\circ\sigma)'(0) &= \tfrac{d}{dt}\Big|_0 (f\circ\Psi_{-t}\circ\Phi_{-t}\circ\Psi_t\circ\Phi_t)(p) \\ &= \lim_{t\to 0}\tfrac{1}{t}\Big((f\circ\Psi_{-t}\circ\Phi_{-t}\circ\Psi_t\circ\Phi_t)(p) - f(p)\Big) \end{aligned}$$

Apply the chain-rule repeatedly, or directly, add and subtract suitable terms and take the limits of each difference separately

$$\begin{aligned} \sigma_0' f &= \lim_{t\to 0}\tfrac{1}{t}\Big((f\circ\Psi_{-t})(\Phi_{-t}\circ\Psi_t\circ\Phi_t(p)) - f(\Phi_{-t}\circ\Psi_t\circ\Phi_t(p))\Big) \\ &\quad + \lim_{t\to 0}\tfrac{1}{t}\Big((f\circ\Phi_{-t})(\Psi_t\circ\Phi_t(p)) - f(\Psi_t\circ\Phi_t(p))\Big) \\ &\quad + \lim_{t\to 0}\tfrac{1}{t}\Big((f\circ\Psi_t)(\Phi_t(p)) - f(\Phi_t(p))\Big) \\ &\quad + \lim_{t\to 0}\tfrac{1}{t}\Big((f\circ\Phi_t)(p) - f(p)\Big) \\ &= (-Yf - Xf + Yf + Xf)(p) \quad = \quad 0 \end{aligned}$$

This last step uses the continuity of Xf and Yf, e.g. that $\lim\limits_{t\to 0}(Xf)(\Psi_t\circ\Phi_t(p))) = (Xf)(p)$, compare exercise 3.41.
Regarding the second part of the proposition, recognize that it suffices to consider the difference quotient $\frac{2}{t^2}(f(\sigma(t)) - f(p))$ because the first derivative vanishes at $t = 0$, compare exercise 3.40. To directly employ lemma 3.27 insert the identity written as $\Phi_{-t}\circ(\Psi_{-t}\circ\Psi_t)\circ\Phi_t$, and use that $[X,Y]$ is continuous:

$$\begin{aligned} \tfrac{d^2}{dt^2}\Big|_0 (f\circ\sigma)(t) &= \lim_{t\to 0}\tfrac{2}{t^2}\Big((f\circ\Psi_{-t})(\Phi_{-t}\circ\Psi_t\circ\Phi_t)(p) - f(p)\Big) \\ &= \lim_{t\to 0}\tfrac{2}{t^2}\Big((f\circ\Psi_{-t}\circ\Phi_{-t})(\Psi_t\circ\Phi_t(p)) - (f\circ\Phi_{-t}\circ\Psi_{-t})(\Psi_t\circ\Phi_t(p))\Big) \\ &= 2\,[X,Y]_p\, f \quad . \qquad ■ \end{aligned}$$

Exercise 3.41 *Suppose $f, g\colon \mathbb{R} \mapsto \mathbb{R}$ with $g(0) = p \in \mathbb{R}$, $g \in C^0(0)$, and $f \in C^1(p)$. Rigorously prove that $f'(p) = \lim_{h\to 0} (f(g(h) + h) - f(g(h)))$. It may be convenient to employ the function $F\colon (s,t) \mapsto F(g(s) + t)$, or alternatively estimate the integral $\frac{1}{h}\int_{g(h)}^{g(h)+h} (f'(s) - f(p))\, ds$.)*
Explain where your argument breaks down in the case of $p = 0$ and f defined by $f(0) = 0$ and $f(x) = x^2 \cdot \sin\frac{1}{x}$ else (which is differentiable, but not continuously differentiable at zero).
Use this to rigorously justify the limits in the calculation of $\sigma_0' f$ in the proof of the proposition. E.g. first consider the simple case of functions

Exercise 3.42 *Under the hypotheses of lemma 3.27 show that*

$$[X, Y]_p f = \lim_{t\to 0} \tfrac{1}{t}\Big((f \circ \Psi_{-\sqrt{t}} \circ \Phi_{-\sqrt{t}} \circ \Psi_{\sqrt{t}} \circ \Phi_{\sqrt{t}})(p) - f(p) \Big) \tag{111}$$

Proof (of theorem 3.25).
Suppose $p \in M^m$, $X_i \in \Gamma^\infty(M)$, $i = 1, \ldots k$ are such that $\{X_{1p}, \ldots X_{kp}\}$ are linearly independent, and that $[X_i, X_j] \equiv 0$ on an open neighborhood $O \subseteq M$ of p.
Since the tangent vectors $\{X_{1p}, \ldots X_{kp}\}$ are linearly independent at p they necessarily are independent in some neighborhood $O' \subseteq O$ of p. Now suppose (v, V) is a chart about p such that $V \subseteq O'$ and w.l.o.g $\frac{\partial}{\partial v^i}|_p = X_{ip}$ for $i = 1, \ldots k$.

Exercise 3.43 *Verify that one always can obtain such an* adapted *coordinate chart after performing, if necessary, a constant* linear *change of coordinates.*

Let Φ^i, $i = 1, \ldots k$ be the local flows of the vector fields X_i defined near p. Without loss of generality, after restricting the domains as necessary, we may assume that each flow Φ^i is defined for all $t \in (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$ and for all q in some open neighborhood $V' \subseteq V$ of p. Let $W = \{x \in \mathbb{R}^m \colon |x^i| < \varepsilon \text{ for } i = 1, \ldots m\}$ and define $V'' = V' \cap v^{-1}(W)$, which is still an open neighborhood of p. Define the map $\Psi\colon v(V') \cap W \mapsto M$ by

$$\Psi(x) = \Phi^k_{x^k} \circ \ldots \circ \Phi^2_{x^2} \circ \Phi^1_{x^1} \circ v^{-1}(0, \ldots, 0, x^{k+1}, \ldots x^m) \tag{112}$$

Clearly at $a = 0$ the tangent vectors $\Psi_{*0}(D_i|_0) = \frac{\partial}{\partial v^i}|_p$ for $i = 1, \ldots m$ are aligned with the coordinates. Hence Ψ_{*0} has full rank, and thus Ψ_{*x} has full rank for x near $0 \in \mathbb{R}^m$. Consequently the restriction of Ψ to some open neighborhood $W' \subseteq W$ of $0 \in \mathbb{R}^m$ is a diffeomorphism (onto its image). Let $U = \Psi(W') \subseteq M$ and define $u\colon U \mapsto \mathbb{R}^m$ by $u = (\Psi|_{W'})^{-1}$. Clearly for all $k \leq n$, $\frac{\partial}{\partial u^k} \equiv X_k|_U$. Using that the flows of the vector fields X_i commute on U, write

$$\Psi(x) = \Phi^i_{x^i} \circ (\Phi^k_{x^k} \circ \ldots \Phi^{i+1}_{x^{i+1}}) \circ (\Phi^{i-1}_{x^{i-1}} \circ \ldots \circ \Phi^2_{x^2} \circ \Phi^1_{x^1}) \circ v^{-1}(0, \ldots, 0, x^{k+1}, \ldots x^m) \tag{113}$$

and one readily sees that also $\frac{\partial}{\partial u^i} \equiv X_i|_U$ for all $i = 1, \ldots k$. ■

3.12 Distributions and integrability theorems

This section generalizes the first integrability theorem proven in the preceding section. It relaxes the condition that the vector fields pairwise commute to the weaker condition that the Lie brackets are linear combinations of the set of given vector fields. This main theorem is known as *Frobenius integrability theorem.*
The natural language to address this situation is in terms of the linear spans of the vector fields, termed *distributions*. The section introduces some elegant language and terminology, but the main ideas – exploiting Lie brackets and commutativity – come from the preceding section.

Definition 3.21 *A* distribution *on a manifold M is a function Δ that assigns to every $p \in M$ a subspace $\Delta_p(M) \subseteq T_pM$. If $\dim T_pM = k$ for all $p \in M$ then Δ is called a k-*dimensional distribution. *A distribution Δ is called* smooth *if for every $p \in M$ there exist an open neighborhood U of p and k smooth vector fields $X_k \in \Gamma^\infty(U)$ such that for every $q \in U$, Δ_q is the linear span of$X_{1q}, \ldots X_{kq}$. A vector field $X \in \Gamma^\infty(m)$ is said to* belong to *a distribution Δ if for all $p \in M$, $X_p \in \Delta(p)$. A k-dimensional submanifold $N \subseteq M$ is called an* integral manifold *of Δ if for all $q \in M$, $\imath_*(T_qN) = \Delta_q$ (where $\imath\colon N \mapsto M$ is the inclusion map.)*

A few comments are in order:

- Integral manifolds of distributions are similar to integral curves of vector fields, but without the *time*-parameterization.
- A collection of vector fields always determines a distribution. But in general this distribution need not have a well-defined dimension as is illustrated by a (single) vector field that vanishes at some (but not all) points.
- Not every k-dimensional distribution is determined by a collection of k vector fields: Consider the M-strip M as the quotient of $\mathbb{R}^2$ under the equivalence relation $(x, y) \sim (x', y')$ if and only if $x' - x = 2\pi k$ and $y' = (-)^k y$ for some integer k and let $\pi\colon \mathbb{R}^2 \mapsto M$ be the associated quotient map. Note that the vector field D_2 on $\mathbb{R}^2$ does not map to a vector field on M. However, the distribution Δ generated by D_2 maps to a distribution $\pi_*(\Delta)$ on M which is not generated by any single vector field on M.
- While for every smooth vector field there exist unique integral curves through every point, in general a smooth k-dimensional distribution will not have any integral manifold if its dimension is larger than one. A most simple example is the smooth distribution on $\mathbb{R}^3$ spanned by the two vector fields $X = D_1 + yD_3$ and $Y = D_2 - xD_3$. As will become clear in the following this is an immediate consequence of the Lie bracket $[X, Y]$ being linearly independent of $\{X, Y\}$ (over $C^\infty(M)$).

Exercise 3.44 *Use a computer algebra system (or other drawing software) to generate images of the distribution Δ on $\mathbb{R}^3$ that is spanned by the vector fields $X = D_1 + x^2 D_3$ and $Y = D_2 - x^1 D_3$ (or $Y = D_2$ is an even simpler example). Explain graphically why Δ does not have any integral manifold.*

We next provide a few useful notions before proceeding to Frobenius' theorem.

Definition 3.22 *Suppose $\Phi \in C^\infty(M^m, N^n)$. Two vector fields $X \in \Gamma^\infty(M)$ and $Y \in \Gamma^\infty(N)$ are called* Φ-related *if $\Phi_{*p}X_p = Y_{\Phi(p)}$ for all $p \in M$.*

Proposition 3.28 *Suppose $\Phi\colon M \mapsto N$ is a diffeomorphism and $Y \in \Gamma^\infty(M)$ generates the local flow Ψ_t. The $(\Phi_* Y) \in \Gamma^\infty(N)$ generates the local flow $\Phi \circ \Psi_t \circ \Phi^{-1}$.*

A typical application is when $M = N$ and Φ is the local flow of a vector fields X (evaluated at a fixed time t). Compare exercise 3.39 for a Taylor expansion of the infinitesimal generator $(\Phi_* Y)$.

Proof. Suppose $p \in M$, $q = \Phi(p) \in N$, $f \in C^\infty(q)$, and calculate

$$\begin{aligned} (\Phi_* Y)_q(f) &= (\Phi_{*\Phi^{-1}(q)} Y_{\Phi^{-1}(q)})(f) \\ &= Y_{\Phi^{-1}(q)}(f \circ \Phi) \\ &= \lim_{h\to 0} \tfrac{1}{h}\left(((f\circ\Phi)\circ\Psi_h)(\Phi^{-1}(q)) - (f\circ\Phi)(\Phi^{-1}(q))\right) \\ &= \lim_{h\to 0} \tfrac{1}{h}\left((f\circ(\Phi\circ\Psi_h\circ\Phi^{-1}))(q) - f(q)\right). \ \blacksquare \end{aligned} \tag{114}$$

Proposition 3.29 *Suppose $\Phi \in C^\infty(M^m, N^n)$, $X_1 \in \Gamma^\infty(M)$ is Φ-related to $Y_1 \in \Gamma^\infty(N)$, and $X_2 \in \Gamma^\infty(M)$ is Φ-related to $Y_2 \in \Gamma^\infty(N)$. Then $[X_1, X_2] \in \Gamma^\infty(M)$ is Φ-related to $[Y_1, Y_2] \in \Gamma^\infty(N)$.*

Proof. Suppose $\Phi \in C^\infty(M^m, N^n)$, $p \in M$, $f \in C^\infty(\Phi(p))$ and X_i, Y_i are Φ-related as in the proposition. Thus, by hypothesis, $(\Phi_{*p} X_{ip}) = Y_{i\Phi(p)}$, i.e. $X_{ip}(f\circ\Phi) = Y_{i\Phi(p)} f = (Yf)(\Phi(p))$ for $i = 1, 2$. Calculate:

$$\begin{aligned} [X_1, X_2]_p(f\circ\Phi) &= \Big(X_1(X_2(f\circ\Phi) - X_2(X_1(f\circ\Phi)\Big)(p) \\ &= \Big(X_1(Y_2 f)\circ\Phi) - X_2(Y_1 f)\circ\Phi)\Big)(p) \\ &= \Big(Y_1(Y_2 f)) - Y_2(Y_1 f))\Big)(\Phi(p)) \\ &= [Y_1, Y_2]_{\Phi(p)} f \ . \quad \blacksquare \end{aligned}$$

Definition 3.23 *A smooth distribution Δ on a manifold M is called* involutive *if whenever $X, Y \in \Gamma^\infty(M)$ belong to Δ, then $[X, Y]$ belongs to Δ.*

Proposition 3.30 *Suppose Δ is a k-dimensional smooth distribution on M. Then Δ is involutive iff for every $p \in M$ there exist a neighborhood $U \subseteq M$ of p and vector fields $X_1, \ldots X_k \in \Gamma^\infty(M)$ which span Δ on U, and there exist functions $c_{ij}^\ell \in C^\infty(U)$ such that*

$$[X_i, X_j] = \sum_{\ell=1}^{k} c_{ij}^\ell X_\ell \quad \textit{for all } i, j = 1, \ldots k \tag{115}$$

Proof. The *only if* part is clear from the definition. Conversely suppose $X, Y \in \Gamma^\infty(U)$ belong to Δ. This means that there are smooth functions $f^i, g^j \in C^\infty(U)$ (compare the following exercise) such that $X = \sum_{i=1}^k f^i X_i$ and $Y = \sum_{j=1}^k g^j X_j$. Consequently,

$$\begin{aligned} [X, Y] &= \textstyle\sum_{i,j=1}^k \left(f^i g^j [X_i, X_j] + f^i (X_i g^j) X_j - g^j (X_j f^i) X_i\right) \\ &= \textstyle\sum_{j=1}^k \Big(\sum_{i,\ell=1}^k f^i g^\ell c_{i\ell}^j + \sum_{i=1}^k \left(f^i (X_i g^j) - g^i (X_i a^j)\right)\Big) X_j \end{aligned} \tag{116}$$

which clearly belongs to Δ. $\blacksquare$

Exercise 3.45 *Elaborate why the functions f^i and g^j in the preceding proof are smooth.*

Theorem 3.31 (Frobenius) *Suppose Δ is an involutive k-dimensional smooth distribution on M and $p \in M$. Then there exists a chart (u, U) about p with $u(p) = 0$ and $u(U) = (-\varepsilon, \varepsilon)^m \subseteq \mathbb{R}^m$ such that for every $c = (c^{k+1}, \ldots c^m) \in (-\varepsilon, \varepsilon)^{m-k}$ the set $N_c = \{q \in U \colon u^j(q) = c^j,\ j = k+1, \ldots m\}$ is an integral manifold of Δ. Moreover, every connected integral manifold of the restriction of Δ to U is contained in some N_c as above.*

Proof. Suppose Δ is an involutive k-dimensional smooth distribution on M and $p \in M$. Start with any chart (v, V) about p such that $\Delta(p)$ is spanned by $\{\frac{\partial}{\partial v^1}|_p, \ldots \frac{\partial}{\partial v^k}|_p\}$. (As in the previous section, this may always be achieved via a (constant) linear change of coordinates).
Let $W = \{q \in V \colon v^{k+1}(q) = \ldots = v^m(q) = 0\}$ and consider the *projection* $\pi \colon V \mapsto W$ defined by $\pi(q) = v^{-1}(v^1(q), \ldots v^k(q), 0, \ldots 0)$. Then $\pi_{*p}\big|_{\Delta(p)} \colon \Delta(p) \mapsto T_pW \subseteq T_pM$ is the identity map. Hence, by continuity, $\pi_{*q}\big|_{\Delta(q)} \colon \Delta(q) \mapsto T_qW \subseteq T_qM$ is one-to-one for q sufficiently close to p. This allows us to define vector fields X_j by setting X_{jq} to be the unique tangent vector

$$X_j(q) \in \pi_{*q}^{-1}\left(\frac{\partial}{\partial v^j}\Big|_{\pi(q)}\right) \cap \Delta(q) \tag{117}$$

Note that this defines *smooth* vector fields X_j near p. [[An alternative approach starts with vector fields Y_j that locally span Δ and then find $a_i^j \in C^\infty(V)$ such that $\sum_{j=1}^k a_i^j \pi_{*q} Y_{jq} = \frac{\partial}{\partial v^i}\Big|_{\pi(q)}$.]]
The crux is that the vector fields X_j and $\frac{\partial}{\partial v^i}$ are π-related (for each $j = 1, \ldots k$. Since

$$\pi_{*q}[X_i, X_j]_q = [\tfrac{\partial}{\partial v^i}, \tfrac{\partial}{\partial v^j}]_{\pi(q)} = 0 \tag{118}$$

and $\pi_{*q}|_{\Delta(q)}$ is one-to-one this implies that $[X_i, X_j]_q = 0$. Since $\{X_{1q}, \ldots X_{kq}\}$ also form a basis for $\Delta(q)$ for q sufficiently close to p, the integrability theorem 3.25 of the previous section applies and yields a coordinate chart (u, U) about p with $U, \subseteq V$ such that $X_j\big|_U \equiv \frac{\partial}{\partial u^j}$ for $j = 1, \ldots k$. The manifolds $N_c = \{q \in U \colon u^{k+1} = c^{k+1}, \ldots u^m(q) = c^m\}$ are integral manifolds for $\Delta\big|_U$.
Conversely, suppose $N \subseteq U$ is any connected integral manifold of $\Delta\big|_U$, and $\imath \colon N \hookrightarrow U$ is the inclusion map. If $q \in N$ and $Y_q \in T_qN$ then $Y_q(u^j \circ \imath) = (\imath_{*q} Y_q) u^j = 0$ for $j = k+1, \ldots m$ because $\imath_{*q}(Y_q) \in \Delta(q)$ by construction. This shows that u^j, $j = k+1, \ldots m$ are constant on N and hence $N \subseteq N_c$ for some c as above. ■

Exercise 3.46 *Consider $M = \mathbb{R}^3 \setminus \{(0,0,0)\}$ and consider the analytic distribution Δ spanned by the vector fields $X_1 = -x^2 D_1 + x^1 D_2$, $X_2 = -x^3 D_2 + x^2 D_3$, and $X_3 = -x^1 D_3 + x^3 D_1$. Verify that $\Delta_p \subseteq T_pM$ is a 2-dimensional subspace for every $p \in M$. Explicitly calculate the associated flows Φ^1, Φ^2, and Φ^3. Verify that these exist globally on $\mathbb{R} \times M$.*
Calculate a multiplication table for the Lie brackets $[X_i, X_j]$ to verify that Δ is involutive, and conclude that M is "foliated" *by 2-dimensional integral manifolds of Δ (identify these!).*

Exercise 3.47 *(raw and untested, continuation of exercise 3.46).*
Use the notation of the preceding exercise and consider the set G of all compositions $G = \{\Phi^3_{x^3} \circ \Phi^2_{x^2} \circ \Phi^1_{x^1} \colon x^1, x^2, x^3 \in \mathbb{R}^3\} \subseteq \mathrm{Diff}(M)$ as a subset of all diffeomorphisms of M. Verify that G is a Lie group, i.e. it combines the structures of a smooth manifold and a group in such a way that multiplication and inverses are smooth operations (with respect to the C^∞-topology on G [[xxx reference xxx]]).
Find a Lie group isomorhism (i.e. a diffeomorphism that preserves the group operations) from G to the set $SO(3)$ of special orthogonal 3×3 matrices, i.e. 3×3 matrices A satisfying $A^T A = I_{3\times 3}$. Exhibit a local coordinate chart (u, U) of $SO(3)$ about $p = I_{3\times 3}$ and explicitly find formulae for the images of the vector fields X_i under the bijection from above.

_____ in progress _____ in progress _____

the word: foliation infinitesimal symmetries revisit exercise 3.19

_____ in progress _____ in progress _____

3.13 Controllability

expand this section with many examples