Lecture 7. Lie brackets and integrability

In this lecture we will introduce the Lie bracket of two vector fields, and interpret it in several ways.

7.1 The Lie bracket.

Definition 7.1.1 Let X and Y by smooth vector fields on a manifold M. The **Lie bracket** $[X,Y]$ of X and Y is the vector field which acts on a function $f \in C^\infty(M)$ to give $XYf - YXf$.

To make sense of this definition we need to check that this does indeed define a derivation. Linearity is clear, but we need to verify the Leibniz rule:

$$\begin{aligned}
[X,Y](fg) &= X(Y(fg)) - Y(X(fg)) \\
&= X(fY(g) + Y(f)g) - Y(fX(g) + X(f)g) \\
&= f(XYg) + (Xf)(Yg) + (XYf)g + (Yf)(Xg) \\
&\quad - (Yf)(Xg) - f(YXg) - (YXf)g - (Xf)(Yg) \\
&= f(XYg - YXg) + (XYf - YXf)g \\
&= f[X,Y]g + ([X,Y]f)g.
\end{aligned}$$

It is useful to write $[X,Y]$ in terms of its components in some chart: Write $X = X^i\partial_i$ and $Y = Y^j\partial_j$ (here I'm using the summation convention: Sum over repeated indices). Note that

$$[\partial_i, \partial_j]f = \partial_i\partial_j f - \partial_j\partial_i f = \frac{\partial^2}{\partial_i\partial_j}\left(f\circ\varphi^{-1}\right) - \frac{\partial^2}{\partial_j\partial_i}\left(f\circ\varphi^{-1}\right) = 0.$$

Therefore we have

$$[X,Y] = X^i\partial_i Y^j\partial_j - Y^j\partial_j X^i\partial_i = \sum_{i,j=1}^{n}\left(X^i\partial_i Y^i - Y^i\partial_i X^j\right)\partial_j. \qquad (7.1)$$

The Lie bracket measures the extent to which the derivatives in directions X and Y do not commute. The following proposition makes this more precise:

Proposition 7.1.1 *Let $X, Y \in \mathcal{X}(\mathcal{M})$, and let Ψ and be the local flow of X in some region containing the point $x \in M$. Then*

$$[X,Y]_x = \frac{d}{dt}\left(\left(D_x\Psi_t\right)^{-1} Y_{\Psi_t(x)}\right)\Big|_{t=0}.$$

The idea is this: The flow Ψ_t moves us from x in the direction of the vector field X. We look at the vector field Y in this direction, and use the map $D_x\Psi_t : T_xM \to T_{\Psi_t(x)}M$, which is non-singular, to bring Y back to T_xM. This gives us a family of vectors in the same vector space, so we can compare them by differentiating. In particular, this gives us the nice interpretation: The Lie bracket $[X,Y]$ vanishes if and only if Y is invariant under the flow of X. For this reason, the Lie bracket is also often called the *Lie derivative*, and denoted by $\mathcal{L}_X Y$.

Proof (Method 1): I will first give a proof which is quite illustrative, but which has the drawback that we need to consider two separate cases. Fix $x \in M$.
Case 1: Suppose $X_x \neq 0$. I will construct a special chart about x as follows: First take any chart $\varphi : U \to V$ about x, and by composition with a linear map assume that $X_x = \partial_n(x)$. Let

$$\Sigma = \varphi^{-1}\left(\{(w^1,\dots,w^n) \in V : w^n = \varphi^n(x)\}\right).$$

Then Σ is a smooth $(n-1)$-dimensional submanifold of M which passes through x, and is transverse to the vector field X on some neighbourhood of x (i.e. $T_y\Sigma \oplus \mathbb{R}X_y = T_yM$ for all $y \in \Sigma$ sufficiently close to x). Now consider the map $\tilde{\Psi} : \Sigma \times (-\delta,\delta) \to M$ given by restricting the flow Ψ of X to $\Sigma \times (-\delta,\delta)$: We have

$$D_{(x,0)}\tilde{\Psi} = \begin{bmatrix} \mathrm{I}_{n-1} & 0 \\ 0 & 1 \end{bmatrix}$$

with respect to the natural bases $\{\partial_1,\dots,\partial_{n-1},\partial_t\}$ for $T_{(x,0)}(\Sigma \times \mathbb{R})$ and $\{\partial_1,\dots,\partial_{n-1},\partial_n = X_x\}$ for T_xM. In particular, $D_{(x,0)}\tilde{\Psi}$ is non-singular, and so by the inverse function theorem there is a neighbourhood of $(x,0)$ in $\Sigma\times\mathbb{R}$ on which $\tilde{\Psi}$ is a diffeomorphism. We take $\tilde{\varphi}$ to be the chart obtained by first taking the inverse of $\tilde{\Psi}$ to obtain a point in $\Sigma \times \mathbb{R}$, and then applying φ to the point in Σ to obtain a point in $\mathbb{R}^{n-1} \times \mathbb{R}$. The special feature of this chart is that $X_y = \partial_n$ on the entire chart.

Now we can compute in this chart: The flow of X can be written down immediately: $\tilde{\varphi} \circ \Psi_t \circ \tilde{\varphi}^{-1}(x^1,\dots,x^n) = (x^1,\dots,x^{n-1},x^n+t)$. If we write $Y = Y^k\partial_k$, then we have

$$\left(\left(D_x\Psi_t\right)^{-1} Y_{\Psi_t(x)}\right)^k = Y^k(x^1,\dots,x^n+t),$$

and so

$$\frac{d}{dt}\left((D_x\Psi_t)^{-1}Y_{\Psi_t(x)}\right)^k\Big|_{t=0} = \frac{\partial Y^k}{\partial x^n} = X^j\frac{\partial Y^k}{\partial x^j} = [X,Y]^k$$

by (7.1), since $\partial_i X^j = 0$ for all j.

Case 2: Now suppose $X_x = 0$. Then we have $\Psi_t(x) = x$ for all t, and the local group property implies that the linear maps $\{D_x\Psi_t\}$ form a one-parameter group in the Lie group $GL(T_xM) \simeq GL(n)$. Hence $D_x\Psi_t = e^{tA}$ for some A, and we get

$$(D_x\Psi_t)^{-1}Y_{\Psi_t(x)} = e^{tA}Y_x,$$

and

$$\frac{d}{dt}\left((D_x\Psi_t)^{-1}Y_{\Psi_t(x)}\right)\Big|_{t=0} = -AY_x.$$

It remains to compute A: Working in local coordinates,

$$A(\partial_j) = \partial_t\partial_j\Psi_t = \partial_j\partial_t\Psi_t = \partial_j X.$$

Therefore we get

$$\frac{d}{dt}\left((D_x\Psi_t)^{-1}Y_{\Psi_t(x)}\right)\Big|_{t=0} = -Y^j\partial_j X^k\partial_k = [X,Y]$$

by (7.1), since $X^k = 0$ for all k. □

Proof (Method 2): Here is another proof which is somewhat more direct but perhaps less illuminating: Choose any chart φ for M about x. In this chart we can write uniquely $X = X^j\partial_j$ and $Y = Y^k\partial_k$, and here the coefficients X^j and Y^k are smooth functions on a neighbourhood of x. Then

$$\begin{aligned}
\frac{d}{dt}\left(D_{\Psi_t(x)}\Psi_{-t}Y_{\Psi_t(x)}\right)^j\Big|_{t=0} &= \left(\partial_t\partial_k\Psi^j_{-t}\right)\Big|_{t=0}Y^k_x + \left(\partial_k\Psi^j_{-t}\right)\partial_t Y^k(\Psi_t x)\Big|_{t=0}\\
&= \left(\partial_k\partial_t\Psi^j_{-t}\right)\Big|_{t=0}Y^k_x + \delta^j_k X^i_x\partial_i Y^k_x\\
&= -Y^k_x\partial_k X^j_x + X^i_x\partial_i Y^j_x\\
&= [X,Y]^j_x.
\end{aligned}$$

□

We could now deduce the following result on the naturality of the Lie bracket directly from Proposition 7.6.2. However I will give a direct proof:

Proposition 7.1.2 *Let $F : M \to N$ be a smooth map, $X, Y \in \mathcal{X}(M)$, and $W, Z \in \mathcal{X}(N)$ with W F-related to X and Z F-related to Y. Then for every $x \in M$,*

$$[W,Z]_{F(x)} = D_xF([X,Y]_x).$$

Think in particular about the case where F is a diffeomorphism: Then the proposition says that the push-forward of the Lie bracket of X and Y is the Lie bracket of the push-forwards of X and Y.

Proof. For any $f : N \to \mathbb{R}$, and $x \in M$, $y = F(x) \in N$,

$$\begin{aligned}
[W,Z]_y f &= W_y Zf - Z_y Wf \\
&= D_xF(X_x)Zf - D_xF(Y_x)Wf \\
&= X_x((Zf)\circ F) - Y_x((Wf)\circ F) \\
&= X_x(Y(f\circ F)) - Y_x(X(f\circ F)) \\
&= [X,Y]_x(f\circ F) \\
&= D_xF([X,Y]_x)f.
\end{aligned}$$

Here we used the fact that for every $p \in M$,

$$(Zf)\circ F_p = (Zf)_{F(p)} = D_pF(Y_p)f = Y_p(f\circ F),$$

and similarly $(Wf)\circ F = X(f\circ F)$. □

7.2 The Lie algebra of vector fields.

The Lie bracket gives the space of smooth vector fields $\mathcal{X}(M)$ on a manifold M the structure of a *Lie algebra*:

Definition 7.2.1 A **Lie algebra** consists of a (real) vector space E together with a multiplication $[.,.] : E \times E \to E$ which satisfies the three properties

$$\begin{aligned}
[u,v] &= -[v,u]; \\
[au+bv,w] &= a[u,w]+b[v,w]; \\
[[u,v],w]+[[v,w],u]+[[w,u],v] &= 0,
\end{aligned}$$

for all $u, v, w \in E$ and $a, b \in \mathbb{R}$.

The first two properties are immediate. The third (known as the Jacobi identity) can be verified directly:

$$\begin{aligned}
[[X,Y],Z]f &= [X,Y]Zf - Z[X,Y]f \\
&= XYZf - YXZf - ZXYf + ZYXf \\
[[X,Y],Z]f &+ [[Y,Z],X]f + [[Z,X],Y]f \\
&= XYZf - YXZf - ZXYf + ZYXf \\
&\quad + YZXf - ZYXf - XYZf + XZYf \\
&\quad + ZXYf - XZYf - YZXf + YXZf \\
&= 0.
\end{aligned}$$

The following exercise gives another interpretation of the Jacobi identity:

Exercise 7.2.1 Let $X, Y, Z \in \mathcal{X}(M)$. Let Ψ be the local flow of Z near a point $x \in M$. For each t proposition 7.1.2 gives

$$D\Psi_t([X,Y]) = [D\Psi_t(X), D\Psi_t(Y)].$$

Differentiate this with respect to t to get another proof of the Jacobi identity.

Lie algebras play an important role in the theory of Lie groups: Consider the space $\mathfrak{g}$ of left-invariant vector fields on a Lie group G. We have already seen that this is a finite-dimensional vector space isomorphic to the tangent space at the identity T_eG by the natural construction

$$v \in T_eG \mapsto V \in \mathfrak{g} : V_g = D_e l_g(v).$$

We will show that $\mathfrak{g}$ is a Lie algebra. It is sufficient to show that the vector subspace $\mathfrak{g}$ of $\mathcal{X}(M)$ is closed under the Lie bracket operation:

Proposition 7.2.2

$$[\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{g}.$$

Proof. Let $X, Y \in \mathfrak{g}$. Then for any $g \in G$, Proposition 7.1.2 gives

$$[X,Y] = [(l_g)_*X, (l_g)_*Y] = (l_g)_*[X,Y]$$

since X is l_g-related to itself, as is Y. Therefore $[X,Y] \in \mathfrak{g}$. □

The Lie algebra of a group captures the local or infinitesimal structure of a group. It turns out that the group G itself can be (almost) completely recovered just from the Lie algebra $\mathfrak{g}$ – at least, $\mathfrak{g}$ determines the universal cover of G. Different groups with the same universal cover have the same Lie algebra.

Exercise 7.2.2 Show that:

(1). A smooth group homomorphism $\rho : G \to H$ induces a homomorphism from the Lie algebra $\mathfrak{g}$ of G to the Lie algebra $\mathfrak{h}$ of H;

(2). If H is a subgroup of G, then $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$;

(3). A diffeomorphism of manifolds induces a homomorphism between their vector field Lie algebras;

(4). A submersion $F : M \to N$ induces a Lie subalgebra of $\mathcal{X}(M)$ (consisting of those vector fields which are mapped to zero by DF).

7.3 Integrability of families of vector fields

The significance of the Lie bracket, for our purposes, comes from its role in indicating when a family of vector fields can be simultaneously integrated to give a map: Suppose we have k vector fields $E_1, \ldots, E_k$ on M, and we want to try to find a map F from $\mathbb{R}^k$ to M for which $\partial_i F = E_i$ for $i = 1, \ldots, k$. The naive idea is that we try to construct such a map as follows: Pick some point x in M, and set $F(0) = x$. Then we can arrange $\partial_k F = E_k$, by setting $F(te_k) = \Psi_{E_k}(x, t)$ for each t. The next step should be to follow the integral curves of the vector field E_{k-1} from $F(te_k)$ for time s to get the point $F(te_k + se_{k-1})$:

$$F(te_k + se_{k-1}) = \Psi_{E_{k-1}}\left(\Psi_{E_k}(x, t), s\right),$$

and so on following $E_{k-2}, ..., E_2$, and finally E_1. Unfortunately that doesn't always work:

Example 7.3.1 Take $M = \mathbb{R}^2$, $E_1 = \partial_1$, and $(E_2)_{(x,y)} = (1 + x)\partial_2$. Following the recipe outlined above, we set $F(0, 0) = (0, 0)$, and follow the flow of E_2 to get $F(0, t) = (0, t)$, and then follow E_1 to get $F(s, t) = (s, t)$. However, this gives $DF(\partial_1) = E_1$ but not $DF(\partial_2) = E_2$. Note that if we change our procedure by first integrating along the vector field E_1, then the vector field E_2, then we don't get the same result: Instead we get $F(s, t) = (s, t + st)$.

In fact, there is an easy way to tell this could not have worked: If there was such a map, then we would have

$$[E_1, E_2] = [DF(\partial_1), DF(\partial_2)] = DF([\partial_1, \partial_2]) = 0.$$

But in the example we have $[E_1, E_2] = \partial_2 \neq 0$. It turns out that this is the only obstruction to constructing such a map:

Proposition 7.3.2 *Suppose $E_1, \ldots, E_k \in \mathcal{X}(M)$ are vector fields which commute: $[E_i, E_j] = 0$ for $i, j = 1, \ldots, k$. Then for each $x \in M$ there exists a neighbourhood U of 0 in $\mathbb{R}^k$ and a unique smooth map $F : U \to M$ satisfying $F(0) = x$ and $D_y F(\partial_i) = (E_i)_{F(y)}$ for every $y \in U$ and $i \in \{1, \ldots, k\}$.*

Proof. We construct the map F exactly as outlined above. In other words, we set

$$F(y^1, \ldots, y^k) = \Psi_{E_1, y^1} \circ \Psi_{E_2, y^2} \circ \ldots \circ \Psi_{E_k, y^k}(x).$$

Where $\Psi_{E_i, t}$ is the flow of the vector field E_i for time t. This gives immediately that $DF(\partial_1) = E_1$ everywhere, but $DF(\partial_2)$ is not so clear.

Lemma 7.3.3 *If $[X, Y] \equiv 0$, then the flows of X and Y commute:*

$$\Psi_{X,t} \circ \Psi_{Y,s} = \Psi_{Y,s} \circ \Psi_{X,t}.$$

Proof. This is clearly true for $t = 0$, since $\Psi_{X,0}(y) = y$ for all y. Therefore it suffices to show that

$$\partial_t \Psi_{X,t} \circ \Psi_{Y,s} = \partial_t \Psi_{Y,s} \circ \Psi_{X,t}.$$

The left-hand side is equal to the vector field X by assumption, while the right is equal to $D\Psi_{Y,s}(X)$. We have

$$\begin{aligned} D\Psi_{Y,s}(X) &= D\Psi_{Y,0}(X) + \int_0^s \frac{d}{dr} D\Psi_{Y,r}(X)\, dr \\ &= X + \int_0^s D\Psi_{Y,r} \frac{d}{dr'} D\Psi_{Y,r'}(X)\Big|_{r'=0} dr, \end{aligned}$$

where I used the local group property of the flow of Y to get $D\Psi_{Y,r+r'} = D\Psi_{Y,r} \circ D\Psi_{Y,r}$. By Proposition 6.1.1, we have

$$D\Psi_{Y,r'}(X)\Big|_{r'=0} = -[Y, X] = 0,$$

and so $D\Psi_{Y,s}(X) = X$. □

Thus we have for every $j \in \{1, \ldots, k\}$ the expression

$$F(y^1, \ldots, y^k) = \Psi_{E^j, y^j} \circ \Psi_{E^1, y^1} \circ \ldots \circ \Psi_{E_{j-1}, y^{j-1}} \circ \Psi_{E_{j+1}, y^{j+1}} \circ \ldots \circ \Psi_{E_k, y^k}(x)$$

and differentiation in the y^j direction immediately gives $DF(\partial_j) = E_j$. □

A closely related but slightly more general integrability theorem is the theorem of Frobenius, which we will prove next.

Definition 7.3.4 A distribution $\mathcal{D}$ of k-planes on M is a map which associates to each $x \in M$ a k-dimensional subspace $\mathcal{D}_x$ of T_xM, and which is smooth in the sense that for any $V \in \mathcal{X}(M)$ and any chart with coordinate tangent basis $\{\partial_1, \ldots, \partial_n\}$, the vector field $V_{\mathcal{D}}$ given by projecting V onto $\mathcal{D}$ (orthogonally with respect to the given basis) is smooth.

Remark. One can make this definition more natural (i.e. with less explicit reference to charts) as follows: Given a manifold M, there is a natural manifold $G_k(TM)$ constructed from M, called the Grassmannian bundle (of k-planes) over M

$$G_k(TM) = \{(x, E) : \ x \in M, \ E \text{ a } k\text{-dimensional subspace of } T_xM\}.$$

Note that the k-dimensional subspaces of T_xM are in one-to-one correspondence with the space of equivalence classes of rank $(n-k)$ linear maps from T_xM to $\mathbb{R}^{n-k}$, under the equivalence relation

$$M_1 \sim M_2 \Longleftrightarrow M_1 = LM_2, \ L \in GL(n-k) :$$

We denote the equivalence class of M by $[M]$. Given a k-dimensional subspace E, we choose any $n-k$ independent linear functions $f_1, \ldots, f_{n-k}$ on T_xM which vanish on E (for example, choose a basis $\{e_1, \ldots, e_n\}$ for T_xM such that $\{e_1, \ldots, e_k\}$ is a basis for E, and take f_i to be the linear function which takes e_{k+i} to 1 and all the other basis elements to zero. Then $f = (f_1, \ldots, f_{n-k})$ defines a rank $n-k$ linear map from T_xM to $\mathbb{R}^{n-k}$ vanishing on E. This correspondence is well-defined as a map into the space of equivalence classes, since if we choose a different basis $\{e'_1, \ldots, e'_n\}$ of this form, then it is related to the first by $e'_i = A(e_i)$ for some non-singular matrix, and we have $f'_j = \sum_{m=1}^{n-k} A_{k+j}^{k+m} f_m$, i.e. $f' = Af \sim F$.

Conversely, given a rank $n-k$ linear function M, we associate to it the k dimensional linear subspace $\ker M$. Clearly if $M_1 \sim M_2$ then $\ker M_1 = \ker M_2$.

Now we can choose charts for $G_k(TM)$: Let $\mathcal{A}$ be an atlas for M, $\varphi : U \to V$ a chart of $\mathcal{A}$. Define an open subset $\tilde{U}$ of $G_k(TM)$ by

$$\tilde{U} = \left\{ (x, [F]) \in U \times L(T_xM, \mathbb{R}^{n-k}) : \ F\Big|_{\mathrm{span}(\partial_1, \ldots, \partial_{n-k})} \text{ nonsingular} \right\}$$

and define $\varphi_\sigma : U_\sigma \to \mathbb{R}^n \times \mathbb{R}^{k(n-k)}$ as follows:

$$\varphi_\sigma(x, [F]) = (\varphi, N)$$

where $N \in L_{k,n-k} \simeq \mathbb{R}^{k(n-k)}$ is the $(n-k) \times k$ matrix constructed as follows: Set $F_i^j = F^j(\partial_{\sigma(i)})$. Then by assumption the first $(n-k)$ columns of the matrix F form a non-singular $(n-k) \times (n-k)$ matrix, which we denote by G. Then we have $[F] = [G^{-1}F]$, and $G^{-1}F$ has the form

$$[\mathrm{I}_{n-k} \quad N]$$

for some $(n-k) \times k$ matrix N. N is well-defined as a function of the equivalence class $[F]$, since if $\tilde{F} = AF$, then $\tilde{G} = GA^{-1}$, so $\tilde{G}^{-1}\tilde{F} = G^{-1}F$. Explicitly,

$$N_i^j = \frac{\sum_{\sigma,\tau \in S_{n-k}} \mathrm{sgn}\sigma \mathrm{sgn}\tau F_{\sigma(1)}^{\tau(1)} \ldots F_{\sigma(n-k-1)}^{\tau(n-k-1)} F_{n-k+i}^{\tau(n-k)} \delta_{\sigma(n-k)}^j}{\sum_{\sigma,\tau \in S_{n-k}} \mathrm{sgn}\sigma \mathrm{sgn}\tau F_{\sigma(1)}^{\tau(1)} \ldots F_{\sigma(n-k-1)}^{\tau(n-k-1)} F_{\sigma(n-k)}^{\tau(n-k)}}$$

for $i = 1, \ldots, k$ and $j = 1, \ldots, n-k$, where the sums are over all pairs of permutations of $n-k$ objects.

Note that these charts cover $G_k(TM)$, because even if $[F] \in G_k(T_xM)$ does not have its first $n-k$ columns non-singular with respect to the coordinate tangent basis of φ, we can choose a new chart for which this is true, of the form $A \circ \varphi$ for some $A \in GL(n)$.

Then we can define a k-dimensional distribution to be a smooth map $\mathcal{D}$ from M to $G_k(TM)$ such that $\mathcal{D}_x \in G_k(T_xM)$ for all $x \in M$.

The question which the Frobenius theorem addresses is the following: Given a distribution $\mathcal{D}$, can we find a k-dimensional submanifold Σ through each point, such that $T_x\Sigma = \mathcal{D}_x$ for every $x \in \Sigma$? This is what is meant by *integrating* a distribution.

For convenience we will denote by $\mathcal{X}(\mathcal{D})$ the vector space of smooth vector fields X on M for which $X_x \in \mathcal{D}_x$ for every $x \in M$ (i.e. vector fields tangent to the distribution).

Proposition 7.3.5 (Frobenius' Theorem) *A distribution $\mathcal{D}$ is integrable if and only if $\mathcal{X}(\mathcal{D})$ is closed under the Lie bracket operation.*

Proof. One direction is clear: Suppose there is such a submanifold N, embedded in M by a map F. Then for any smooth vector fields $X, Y \in \mathcal{X}(\mathcal{D})$ there are unique vector fields $\tilde{X}$ and $\tilde{Y}$ in $\mathcal{X}(N)$ such that X is F-related to $\tilde{X}$ and Y is F-related to $\tilde{Y}$. Then Proposition 7.1.2 gives $[X, Y] = DF([\tilde{X}, \tilde{Y}])$ is in $\mathcal{X}(\mathcal{D})$.

The other direction takes a little more work: Given a distribution of k-planes $\mathcal{D}$ which is involutive (closed under Lie brackets), and any point $x \in M$, we want to construct a submanifold through x tangent to the distribution. Choose a chart φ for M near x, and assume that $\mathcal{D}_x$ is the subspace of T_xM generated by the first k coordinate tangent vectors. Then by the smoothness of the distribution we can describe $\mathcal{D}_y$ for y sufficiently close to x as follows:

$$\mathcal{D}_y = \left\{ \sum_{i=1}^{k} c^i \left(\partial_i + \sum_{j=k+1}^{n} a_i^j(y)\partial_j \right) \,:\, (c^1, \ldots, c^k) \in \mathbb{R}^k \right\}.$$

Here the functions a_i^j are smooth functions on a neighbourhood of x, for $i = 1, \ldots, k$ and $j = k+1, \ldots, n$. In particular we have k non-vanishing vector fields $E_1, \ldots, E_k$ tangent to $\mathcal{D}$, given by

$$(E_i)_y = \partial_i + \sum_{j=k+1}^{n} a_i^j(y)\partial_j.$$

Now we compute:

$$\begin{aligned}
[E_i, E_j] &= \left[\partial_i + \sum_{p=k+1}^{n} a_i^p(y)\partial_p,\ \partial_j + \sum_{q=k+1}^{n} a_j^q(y)\partial_q \right] \\
&= \sum_{p=k+1}^{n} \left(\partial_i a_j^p - \partial_j a_i^p + \sum_{q=k+1}^{n} \left(a_i^q \partial_q a_j^p - a_j^q \partial_q a_i^p \right) \right) \partial_p.
\end{aligned}$$

But by assumption $[E_i, E_i] \in \mathcal{X}(\mathcal{D})$, so it is a linear combination of the vector fields $E_1, \ldots, E_k$. But $[E_i, E_j]$ has no component in the direction $\partial_1, \ldots, \partial_k$. Therefore $[E_i, E_j] = 0$.

Proposition 7.3.2 gives the existence of a map F from a region of $\mathbb{R}^k$ into M with $DF(\partial_i) = E_i$ for $i = 1, \ldots, k$. In particular DF is of full rank, and F is an immersion, hence locally an embedding, and the image of F is a submanifold with tangent space equal to $\mathcal{D}$ everywhere. □

Example 7.3.2 The Heisenberg group In this example we investigate a situation where a distribution $\mathcal{D}$ is not involutive. Let G be the group of 3×3 matrices of the form

$$\begin{bmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{bmatrix},$$

which we identify with three-dimensional space $\mathbb{R}^3$. The left-invariant vector fields corresponding to the standard basis $\{e_1, e_2, e_3\}$ at the identity $0 \in \mathbb{R}^3$ are given by

$$E_1 = \begin{bmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \simeq e_1;$$
$$E_2 = \begin{bmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & x_1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \simeq e_2 + x_1 e_3;$$
$$E_3 = \begin{bmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \simeq e_3.$$

We take a distribution $\mathcal{D}$ on $\mathbb{R}^3$ to be the subspace at each point spanned by E_1 and E_2. Then $\mathcal{D}$ is not involutive, since we have

$$[E_1, E_2] = [e_1, e_2 + x_1 e_3] = e_3 = E_3.$$

The behaviour in this case turns out to be as far as one could imagine from that in the involutive case : If there were a submanifold tangent to the distribution, then only points in that submaniofld could be reached along curves tangent to the distribution. In contrst, we have

Proposition 7.3.6 *For any $y \in \mathbb{R}^3$ there exists a smooth $\gamma : [0,1] \to \mathbb{R}^3$ with $\gamma'(t) \in \mathcal{D}_{\gamma(t)}$ for all t, $\gamma(0) = 0$, and $\gamma(1) = y$.*

In other words, *every* point can be reached by following curves tangent to $\mathcal{D}$.

Proof. We consider curves tangent to the distribution, given by prescribing the tangent vector at each point:

$$\gamma'(t) = \alpha(t) E_1 + \beta(t) E_2.$$

This gives the system of equations

$$\begin{aligned} x' &= \alpha; \\ y' &= \beta; \\ z' &= x\beta; \end{aligned}$$

which gives

$$\begin{aligned} x(t) &= \int_0^t \alpha(s)\, ds; \\ y(t) &= \int_0^t \beta(s)\, ds; \\ z(t) &= \int_0^t \beta(s) \int_0^s \alpha(s')\, ds'\, ds. \end{aligned}$$

Then if $y = (X, Y, Z)$ we make the choice $\alpha(t) = X + af(t)$ and $\beta(t) = Y + bf(t)$, where a and b are constants to be chosen, and f is a smooth function which satisfies $\int_0^1 f(t)dt = 0$ and $\int_0^1 tf(t)ds = 0$, but $\Gamma = \int_0^1 \int_0^t f(t)f(s)ds\,dt \neq 0$ (for example, $f(t) = t^2 - t + 1/6$ will work). Then we have $x(1) = X$, $y(1) = Y$, and $z(1) = XY/2 + ab\Gamma$. Finally, we can choose a and b to ensure that $z(1) = Z$ (e.g. $a = 1$, $b = (Z - XY/2)/\Gamma$). $\square$

Example 7.3.3 Subgroups of Lie groups As an application of the Frobenius theorem, we will consider subgroups of Lie groups. We have seen that the left-invariant vector fields on a Lie group G form a finite-dimenensional Lie algebra $\mathfrak{g}$. We will see that subspaces of $\mathfrak{g}$ which are closed under the Lie bracket correspond exactly to connected smooth subgroups of G. In one direction this is clear: Suppose H is a Lie group which is contained in G, with the inclusion map being an immersion. Then the Lie algebra $\mathfrak{h}$ of H is naturally included in $\mathfrak{g}$, since for any left-invariant Lie algebra V on H, we can extend to a left-invariant Lie algebra $\tilde{V}$ on G by setting

$$\tilde{V}_g = D_0 l_g(V_e).$$

$\tilde{V}$ agrees with V on H. The image of this inclusion is a vector subspace of $\mathfrak{g}$. Furthermore, if $X, Y \in \mathfrak{h}$ then $[X, Y]\tilde{} = [\tilde{X}, \tilde{Y}]$ is in this subspace of $\mathfrak{g}$. So we have a subspace of $\mathfrak{g}$ which is closed under Lie brackets.

Proposition 7.3.7 *Let $\mathfrak{h}$ be any vector subspace of $\mathfrak{g}$ which is closed under Lie brackets. Then there exists a unique connected Lie group H and an inclusion $i : H \to G$ which is an injective immersion, such that $D_h i(T_h H)$ is the subspace of $T_{i(h)}G$ given by the left-invariant vector fields in $\mathfrak{h}$, for every $h \in H$.*

Proof. Since $\mathfrak{h}$ is closed under Lie brackets, the distribution $\mathcal{D}$ defined by the vectors in $\mathfrak{h}$ is involutive: If we take a basis $\{E_1, \ldots, E_k\}$ for $\mathfrak{h}$, then any $X, Y \in \mathcal{X}(M)$ can be written in the form $X = X^i E_i$, $Y = Y^j E_j$ for some smooth functions $X^1, \ldots, X^k$ and $Y^1, \ldots, Y^k$. Then we have

$$[X, Y] = [X^i E_i, Y^j E_j] = (X^i E_i(Y^j) - Y^i E_i(X^j))E_j + X^i Y^j [E_i, E_j]$$

which is in $\mathcal{D}$. By Frobenius' Theorem, there is a submanifold Σ passing through $e \in G$ with tangent space $\mathcal{D}$.

To show that Σ is a subgroup, we show that $xy \in \Sigma$ for all x and y in Σ: Write $y = \exp(sY)$ for $Y \in \mathcal{D}_e$. Then

$$\frac{d}{ds}\left(x \exp(sY)\right) = \frac{d}{dr}\left(x \exp(sY) \exp(rY)\right)\Big|_{r=0} = D_0 l_{x\exp(sY)} Y \in \mathcal{D}$$

for each s. Thus the curve $s \mapsto x\exp(sY)$ starts in Σ and is tangent to $\mathcal{D} = \mathcal{T}\Sigma$, and so stays in Σ. Therefore $xy \in \Sigma$, and Σ is a subgroup. $\square$