Lie Groups and Lie Algebras
MATH 744

William Goldman's Lecture
Department of Mathematics
University of Maryland

January 29, 2002

Chapter 1

Overviews and Examples of Lie Groups and Lie Algebras

1.1 Hamiltonians $\mathbb{H}$

1.1.1 Notations

$\mathbb{H}$ is a group of quaternions.
$\mathbb{H}^\times$ is a group of nonzero quaternions.
$\mathbb{H}_0$ is a real vector space of purely imaginary quaternions, i.e., $\mathbb{H}_0 = \mathbb{R}\{i, j, k\} \cong \mathbb{R}^3$ and $\mathbb{H} = \mathbb{R} \oplus \mathbb{H}_0$.
$\mathbb{H}_1$ is a group of unitary quaternions, i.e., $\mathbb{H}_1 = \{u \in \mathbb{H} \mid ||u||^2 = 1\}$.
$\mathbf{SL}(n)$ is a group of unimodular automorphisms of $\mathbb{R}^n$. Here, note that a unimodular automorphism means a volume preserving automorphism.
$\mathbf{SL}(n, \mathbb{R}) = \{A \in \mathbf{M}_n(\mathbb{R}) \mid \mathbf{det} A = 1\}$.
$\mathbf{SO}(n)$ is a group of unimodular and orthogonal automorphisms of $\mathbb{R}^n$.
$\mathbf{SO}(n, \mathbb{R}) = \{A \in \mathbf{M}_n(\mathbb{R}) \mid \mathbf{det} A = 1 \text{ and } AA^T = I\}$.

1.1.2 Involution α in $\mathbb{H}$

Let $\alpha : \mathbb{H} \to \mathbb{H}$ be a conjugation, which means that $\alpha(a + bi + cj + dk) = a - bi - cj - dk$. That is, α is an involution of $\mathbb{H}$, i.e., $\alpha^2 = id$. It is easy to see that α enjoys the following properties:
1. $u \in \mathbb{H}$ is real $\Longleftrightarrow \alpha(u) = u$.
2. $u \in \mathbb{H}$ is purely imaginary $\Longleftrightarrow \alpha(u) = -u$.
3. $u\alpha(u) = ||u||^2$.
4. $\alpha(uv) = \alpha(v)\alpha(u)$. Note that the order is reversed.
5. $Re(u) = \frac{1}{2}(u + \alpha(u))$ and $Im(u) = \frac{1}{2}(u - \alpha(u))$.

1.1.3 An Action of $\mathbb{H}_1$ on $\mathbb{H}_0$ by conjugation

Recall that $\mathbb{H}_0 = \mathbb{R}\{i, j, k\} = \{u \in \mathbb{H} \mid Re(u) = 0\} \cong \mathbb{R}^3$ and $\mathbb{H}_1 = \{u \in \mathbb{H} \mid ||u||^2 = 1\} \cong S^3$.

Let $\varphi(u)v = uvu^{-1}$ for $u \in \mathbb{H}_1$ and $v \in \mathbb{H}_0$, where $u^{-1} = \alpha(u)$. Clearly, it is an action of $\mathbb{H}_1$ on $\mathbb{H}_0$. Note that $\mathbb{R}^\times \cap \mathbb{H}_1$ acts trivially.

Exercise 1.1.3.1. *1. Show that $\mathbb{H}_0$ is invariant under the action $\varphi(\mathbb{H}_1)$.*
2. The action $\varphi(u)$ preserves Euclidean geometry of $\mathbb{H}_0 \cong \mathbb{R}^3$, i.e., each element of $\mathbb{H}_1$ gives a rotation of $\mathbb{H}_0$ as $\mathbb{R}^3$.

Proof. In the first case, it suffices to show that $Re(\varphi(u)v) = Re(uvu^{-1}) = 0$ for all $u \in \mathbb{H}_1$ and $v \in \mathbb{H}_0$. By messy but trivial computations, it will be done.

Note that $\varphi : \mathbb{H} = \mathbb{R}^4 \to \mathbb{H} = \mathbb{R}^4$ and the action $\varphi(u) : \mathbb{H} = \mathbb{R}^4 \to \mathbb{H} = \mathbb{R}^4$ is given by $\varphi(u)v = uvu^{-1}$. Since the action of $\varphi(u)$ is obviously a linear isomorphism and

$$uvu^{-1} = 0 \text{ for } u \in \mathbb{H}^\times \iff v = 0,$$

we have $\varphi(u) \in \mathbf{GL}(4, \mathbb{R})$ for $u \in \mathbb{H}^\times$. In the second case, the key thing to show is that $\varphi(u)$ preserves norms for $u \in \mathbb{H}_1$. If the action preserves norms, by the formula,

$$\langle v, w\rangle = \frac{||v+w||^2 - ||v||^2 - ||w||^2}{2}$$

we conclude that it is a unimodular and orthogonal action on $\mathbb{R}^4$. Since

$$||\varphi(u)v|| = ||uvu^{-1}|| = ||u||||v||||u^{-1}|| = ||v|| \text{ for } u \in \mathbb{H}_1, v \in \mathbb{R}^4,$$

we have $\varphi(u) \in \mathbf{SO}(4)$ for $u \in \mathbb{H}_1$.

Suppose that $r \in \mathbb{R}$ and $u \in \mathbb{H}_1$. We have

$$\varphi(u) \cdot r = uru^{-1} = r||u||^2 = r.$$

If we consider the canonical inclusions $\mathbf{SO}(3) \subset \mathbf{SO}(4) \subset \mathbf{O}(4, \mathbb{R})$, by the first problem and the above discussion, we can think the action of $\varphi(u)$ on $\mathbb{H}_0$ as the usual action of an element in $\mathbf{SO}(3)$. ☐

1.1.4 One parameter Subgroup

A **homomorphism** $\varphi : \mathbb{R} \to \mathbb{H}^\times$ is a one parameter subgroup, i.e., $\varphi(s+t) = \varphi(s)\varphi(t)$. More generally, we can replace $\mathbb{H}^\times$ by a Lie group G.

Remark 1.1.4.1. *By the uniqueness and existence of a solution of an O.D.E, one parameter subgroup φ is determined by $\varphi'(0)$ Also, φ determines a flow $\lambda_t : G \to G$ where $\lambda_t = l_{\varphi(t)}$, i.e., $l_{\varphi(t)}(g) = \varphi(t)g$. Note that an infinitesimal left multiplication corresponds to a right invariant vector field.*

Example 1.1.4.1. *Assume that $\varphi'(0) = v \in \mathbb{H}_0$ and let $\varphi'(s) = \varphi(s)$. So,*

$$\varphi(s) = \exp(\varphi'(0)s) = \sum \frac{v^n}{n!}s^n.$$

Since $||v^n|| = ||v||^n, ||\varphi(s)|| \leq \exp(||v||s)$. So, it converges uniformly and absolutely on a compact subset of $\mathbb{R}$. Let $v \in \mathbb{H}_0$, the group of purely imaginary quaternions. So,

$$v = ||v||u \textit{ where } u \in \mathbb{H}_0 \cap \mathbb{H}_1 \cong S^2.$$

Note that $v^n = ||v||^n u^n$ *and* $u^2 = -1$. *So we have*

$$v^n = \begin{cases} (-1)^m ||v||^{2m} & \text{if } n = 2m \\ (-1)^m ||v||^{2m+1} u & \text{if } n = 2m+1. \end{cases}$$

From the above observation, we have an interesting formula for $v \in \mathbb{H}_0$*:*

$$\begin{aligned}\exp(vs) &= \sum \frac{v^n}{n!} s^n = \sum_{n=2m} (-1)^m \frac{||v||^{2m}}{n!} s^{2m} + \sum_{n=2m+1} (-1)^m \frac{||v||^{2m+1}}{n!} s^{2m+1} u \\ &= \cos(||v||s)1 + \sin(||v||s)u.\end{aligned}$$

Generally, $\exp(i\theta)$ *acts on* $\mathbb{H}$ *in three ways, right multiplication, left multiplication, and conjugation. Let us examine an action of* $\exp(i\theta)$ *on* $\mathbb{H}$ *by conjugation. Note that we can think* $\mathbb{C}$ *spanned by* $\{1, i\}$ *as a subalgebra of* $\mathbb{H}$ *and by the relation* $ij = k$, $\{j, k\}$ *spans* $\mathbb{C}j$. *If* $u \in \mathbb{C} \subset \mathbb{H}$, $\varphi(e^{i\theta})u = u$. *Observe that*

$$\begin{aligned}\varphi(e^{i\theta})j = e^{i\theta} j e^{-i\theta} &= (\cos\theta + i\sin\theta)j(\cos\theta - i\sin\theta) \\ &= (\cos^2\theta - \sin^2\theta)j + 2(\sin\theta\cos\theta)k \quad \text{using } ijk = -1 \\ &= (\cos 2\theta)j + (\sin 2\theta)k.\end{aligned}$$

So, since $\varphi(e^{i\theta})k = e^{i\theta} k e^{-i\theta} = (\cos 2\theta)k + (\sin 2\theta)j$, *we have* $\varphi(e^{i\pi}) = id$ *on* $\mathbb{H}$.

Exercise 1.1.4.1. *Let* $u \in S^2 = \mathbb{H}_0 \cap \mathbb{H}_1$. *Find* $v \in S^3 \cong \mathbb{H}_1$ *such that* $\varphi(v)i = u$. *In other words,* $\mathbb{H}_1$ *acts transitively on* S^2.

Proof. By the similar computation, we deduce that the action of $\varphi(e^{j\theta})$ on i is a just rotation through the axis j, $\varphi(e^{k\theta})$ on i is a just rotation through the axis k and $\varphi(e^{i\theta})$ on i is an identity. So, geometrically, we can imagine that for any $u \in S^2$, there exist rotations $\varphi(e^{j\theta_1})$ and $\varphi(e^{k\theta_2})$ which moves i to $u \in S^2$. Hence,

$$v = e^{j\theta_1} \cdot e^{k\theta_2} \in S^3 \text{ where } e^{j\theta_1} = \cos\theta_1 + j\sin\theta_1 \text{ and } e^{k\theta_2} = \cos\theta_2 + k\sin\theta_2.$$

□

1.1.5 Relationship between $\mathbb{H}$ and $\mathbb{C}^2$

First, notice that $\mathbf{SU}(2) \cong \mathbb{H}_1$ and $\mathbf{SU}(2) = \{A \in \mathbf{M}_n(\mathbb{C}) \mid \mathbf{det}A = 1 \text{ and } AA^T = I\}$. Since $\mathbf{SL}(n, \mathbb{R}) = \{A \in \mathbf{M}_n(\mathbb{R}) \mid \mathbf{det}A = 1\}$, $\mathbf{SL}(n, \mathbb{C}) = \{A \in \mathbf{M}_n(\mathbb{C}) \mid \mathbf{det}A = 1\}$, and **det** is a functional of each fields, it is easy to see that $\mathbf{SL}(n, \mathbb{R})$ is a real differentiable manifold of dimension $n^2 - 1$ and $\mathbf{SL}(n, \mathbb{C})$ is a real differentiable manifold of dimension $2n^2 - 2$.

Since $\mathbb{C} \subset \mathbb{H}$ is a subalgebra, we can think $\mathbb{H}$ is a vector space over $\mathbb{C}$ spanned be $\{1, j\}$. So, $\mathbb{H} \cong \mathbb{C}^2$ as complex vector spaces: Suppose that $\mathbb{C}^2$ is spanned by $\{e_1, e_2\}$. A complex linear isomorphism f between $\mathbb{H}$ and $\mathbb{C}^2$ is given by

$$f(a + bi + cj + dk) = ae_1 + bie_1 + ce_2 + die_2 \text{ where } a, d, c, d \in \mathbb{R}.$$

Now, we want to know how $\varphi(j)$ acts on $\mathbb{C}$ in this description. Since $ji(-j) = -i$ and $j1(-j) = 1$, we conclude that the conjugation by j is just a restriction of the complex conjugation, i.e., $z \to \overline{z}$ for $z \in \mathbb{C}$.

Remark 1.1.5.1. *We have a dichotomy.*

$$\mathbf{SL}(2,\mathbb{R}) \subset \mathbf{GL}(2,\mathbb{R}) \subset \mathbf{M}_2(\mathbb{R}) \hookrightarrow \mathbf{M}_2(\mathbb{C}) \Rightarrow k = -1 \text{ Hyperbolic geometry}$$
$$\mathbb{H}_1 \subset S^3 \cong \mathbb{H}_1 \times \mathbb{R}_+ = \mathbb{H}^\times \subset \mathbb{H} \hookrightarrow \mathbf{M}_2(\mathbb{C}) \Rightarrow k = 1 \text{ Elliptic geometry} .$$

The algebra structure of $\mathbf{M}_2(\mathbb{R})$ *is* $\mathbb{R}[i,j]/(i^2 = j^2 = 1, ij = -ji = k)$ *and the algebra structure of* $\mathbb{H}$ *is* $\mathbb{R}[i,j]/(i^2 = j^2 = -1, ij = -ji = k)$.

Exercise 1.1.5.1. *Prove that* $\mathbf{O}(n) = \{A \in \mathbf{M}_n(\mathbb{R}) \mid AA^T = I\}$ *is compact, but* $\mathbf{O}(n,\mathbb{C})$ *is not compact if* $n > 1$.

Proof. It suffices to show that $\mathbf{O}(n)$ is closed and bounded. By the relation $AA^T = I$, we have

$$\sum_{j=1}^{n} a_{ij}^2 = 1 \text{ where } (a_{ij})_{n\times n} = A \in \mathbf{O}(n) \text{ and for } i = 1,\ldots,n.$$

Since a finite intersection of compact sets is compact, $\mathbf{O}(n)$ is closed and bounded in $\mathbb{R}^{n^2}$. However, if we think $A \in \mathbf{O}(n,\mathbb{C})$, we only deduce that $\mathbf{O}(n,\mathbb{C})$ is closed in $\mathbb{C}^{n^2}$ from the equation

$$\sum_{j=1}^{n} a_{ij}^2 = 1 \text{ where } (a_{ij})_{n\times n} = A \in \mathbf{O}(n) \text{ and for } i = 1,\ldots,n.$$

Actually, it is not bounded: For a counter example, we give

$$\infty \leftarrow \begin{pmatrix} \sqrt{n+1} & -\sqrt{n}i \\ \sqrt{n}i & \sqrt{n+1} \end{pmatrix} \in \mathbf{O}(2,\mathbb{C}) \text{ as } n \to \infty.$$

□

Let $\mathbb{C} = \mathbb{R}[i]/(i^2 = -1)$. So we have

$$\mathbb{C} \hookrightarrow \mathbb{H} = \mathbb{R}[i,j]/(i^2 = j^2 = -1, ij = -ji = k).$$

That is, $\mathbb{H}$ has a basis $\{1, j\}$ over $\mathbb{C}$. So, $\mathbb{H} \cong \mathbb{C}^2$ as vector spaces.

Question 1.1.5.1. *If we take* $\mathbb{C}[i,j]/(i^2 = j^2 = -1, ij = -ji = k)$*, which algebra do we get?*

1.1.6 Action of $\mathbb{H}$ on itself by right multiplications r_z

Note that since $\mathbb{H}$ is an associative algebra, right multiplications commute with left multiplications.

Let $\varphi : \mathbb{H} \to \mathbf{End}_{\mathbb{C}}(\mathbb{C}^2)$ by $\varphi(h)z = z\alpha(h)$ for $z \in \mathbb{C}^2$.

$$\begin{aligned}\varphi(h_1h_2)z &= z\alpha(h_1h_2) = z(\alpha(h_2)\alpha(h_1)) = (z\alpha(h_2))\alpha(h_1) \\ &= \varphi(h_1)(z\alpha(h_2)) = \varphi(h_1)(\varphi(h_2)z) = (\varphi(h_1)\varphi(h_2))z.\end{aligned}$$

That is, $\varphi(h_1h_2) = \varphi(h_1)\varphi(h_2)$. Let us examine φ more. On $\mathbb{C}^2$, we have

$$\begin{aligned}
\varphi(1) &= id = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\varphi(i) &= r_{\alpha(i)} = r_{-i} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \text{ since } 1 \mapsto -i \text{ and } j \mapsto j(-i) = ij \\
\varphi(j) &= r_{\alpha(j)} = r_{-j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \text{ since } 1 \mapsto -j \text{ and } j \mapsto 1 \\
\varphi(k) &= r_{\alpha(k)} = r_{-k} = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, \text{ since } 1 \mapsto -k = -ij \text{ and } j \mapsto j(-k) = -i.
\end{aligned}$$

So, we have a representation φ of $\mathbb{H}$ as an algebra in $\mathbf{M}_2(\mathbb{C})$, i.e., for $a, b, c, d \in \mathbb{R}$,

$$\varphi(a + bi + cj + dk) = \begin{pmatrix} a - bi & c - di \\ -c - di & a + bi \end{pmatrix}.$$

That is, we have an algebra isomorphism:

$$\mathbb{H} \cong \mathbb{R}[i, j]/(i^2 = j^2 = -1, ij = -ji = k) \xrightarrow{\varphi} \mathbf{M}_2(\mathbb{C}) \cong \mathbb{R}[i, j]/(i^2 = j^2 = 1, ij = -ji = k).$$

Remark 1.1.6.1. *From the above formula, it is easy to see that*

$$\begin{pmatrix} a - bi & c - di \\ -c - di & a + bi \end{pmatrix} \cdot \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \cdot \begin{pmatrix} a - bi & c - di \\ -c - di & a + bi \end{pmatrix}.$$

So, we deduce that an action of $\mathbb{C}$ on $\varphi(\mathbb{H})$ commutes. That is,

$$i \cdot \varphi(a + bi + cj + dk) = \varphi(a + bi + cj + dk) \cdot i.$$

Remark 1.1.6.2. *We note that there is another representation $\psi : \mathbb{H} \to \mathbf{End}_{\mathbb{C}}(\mathbb{C}^2)$, given by*

$$\begin{aligned}
\psi(1) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\psi(i) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
\psi(j) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\psi(k) &= \psi(i)\psi(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
\end{aligned}$$

It is easy to see that for $a, b, c, d \in \mathbb{R}$,

$$\psi((\frac{a+d}{2}) + (\frac{a-d}{2})i + (\frac{b+c}{2})j + (\frac{b-c}{2})k) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

From this, we conclude that there is an algebra isomorphism between $\mathbb{H}$ and $\mathbf{M}_2(\mathbb{R})$. In this description, we let the algebra structure $\mathbf{M}_2(\mathbb{R})$ induced from $\mathbb{H}$ be $\mathfrak{M}_2(\mathbb{R})$, i.e.,

$$\mathbb{H} \cong \mathfrak{M}_2(\mathbb{R}) \cong \mathbb{R}[i, j]/(i^2 = j^2 = 1, ij = -ji = k).$$

Note that in this correspondence, we also have

$$\mathbb{H}_0 \cong \mathfrak{M}_2(\mathbb{R})_0 \stackrel{def}{=} \{\textsf{\textit{traceless elements of}}\ \mathfrak{M}_2(\mathbb{R})\}.$$

If we think the conjugation α of $\mathbb{H}$ as an involution of $\mathfrak{M}_2(\mathbb{R})$, we have

$$\alpha(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Also, we can define a norm on $\mathfrak{M}_2(\mathbb{R})$, using the formula $||v||^2 = v\alpha(v)$ for $v \in \mathbb{H}$. That is, for $A \in \mathfrak{M}_2(\mathbb{R})$, we define

$$N(A) \stackrel{def}{=} A\alpha(A) = (ad - bc)I = \mathbf{det}(A)I.$$

1.1.7 Action preserving a Hermitian structure on $\mathbb{C}^2$.

A hermitian structure on $\mathbb{C}^2$ is a positive definite, sesquilinear, and $\mathbb{R}$-bilinear inner product on $\mathbb{C}^2$. That is for $u, v \in \mathbb{C}^2$,
1. $\langle u, u\rangle > 0$ if $u \neq 0$.
2. $\langle u, v\rangle = \overline{\langle v, u\rangle}$.
3. $\langle \lambda(u_1 + u_2), \mu v\rangle = \lambda\overline{\mu}(\langle u_1, v\rangle + \langle u_2, v\rangle)$.

The usual Hermitian structure on $\mathbb{C}^n$ is given by

$$\langle z, w\rangle = z^T \cdot \overline{w} = \begin{pmatrix} z_i, \dots, z_n \end{pmatrix} \cdot \begin{pmatrix} \overline{w_1} \\ \vdots \\ \overline{w_n} \end{pmatrix} = \sum_{i=1}^{n} z_i \overline{w_i} \in \mathbb{C} \text{ for } z, w \in \mathbb{C}^n.$$

If we let $z_j = x_j + iy_j$ and $w_j = u_j + iv_j$, we also have $Re\langle z, w\rangle = x \cdot u + y \cdot v$.

Exercise 1.1.7.1. *Construct the Hermitian structure on $\mathbb{C}^2$ corresponding to the quaternions $\mathbb{H}$.*

Proof. Let $z \in \mathbb{C}^2$. Since we identify $\mathbb{C}^2$ with $\mathbb{H}$ as vector spaces, we can write

$$z = a + bj \text{ where } a, b \in \mathbb{C}.$$

Define for $z_1, z_2 \in \mathbb{C}^2$,

$$\langle z_1, z_2\rangle = \langle a_1 + b_1 j, a_2 + b_2 j\rangle = a_1 \cdot \overline{a_2} + b_1 \cdot \overline{b_2}.$$

It is easy to see that $\langle , \rangle \in \mathbb{C}$ and satisfies the required properties. □

By definition,

$$\mathbf{U}(n) = \{A \in \mathbf{GL}(n, \mathbb{C}) \mid A \text{ preserves } \langle , \rangle \text{ the usual Hermitian structure}\}.$$

Since for $A \in \mathbf{U}(n)$ and for all $z, w \in \mathbb{C}^n$,

$$\langle Az, Aw\rangle = (Az)^T(\overline{Aw}) \stackrel{def}{=} z^T\overline{w} = \langle z, w\rangle,$$

we conclude that $A^T\overline{A} = I$. That is,

$$\mathbf{U}(n) = \{A \in \mathbf{GL}(n, \mathbb{C}) \mid A^T\overline{A} = I\}.$$

From this, we also have $A = (\overline{A}^T)^{-1}$ and ϕ is an involution, i.e., automorphism of (real) Lie group of order two:

$$\mathbf{GL}(n, \mathbb{C}) \to \mathbf{GL}(n, \mathbb{C}) \text{ given by } \phi(A) \to (\overline{A}^T)^{-1}.$$

Note that ϕ is smooth but not complex analytic.

Let G be a group and $\phi \in \mathbf{Aut}(G)$.

$$\mathbf{Fix}(\psi) = \{x \in G \mid \psi(x) = x\}.$$

By the above discussion, we have $\mathbf{Fix}(\phi) = \mathbf{U}(n)$.

Exercise 1.1.7.2. *Prove that* $\mathbf{U}(n)$ *is compact.*

Proof. From the relation $A^T\overline{A} = I$,

$$\sum_{j=1}^{n} |a_{ij}|^2 = 1 \text{ where } (a_{ij})_{n\times n} = A \in \mathbf{U}(U) \text{ and for } i = 1, \ldots, n.$$

So, $\mathbf{U}(n)$ is closed and bounded in $\mathbb{C}^{n^2}$. ☐

1.1.8 Revisit to the Matrix algebra of $\mathbf{M}_2(\mathbb{R})$.

In this section, we revisit the material, which we discussed in Remark 1.1.6.2, so that it would give you more rigorous and clearer views. We note that from the discussion in Remark 1.1.6.2, we have $\mathfrak{M}_2(\mathbb{R}) \cong \mathbb{R}[I, J]/(I^2 = J^2 = \mathbf{1}, IJ + JI = \mathbf{0})$, where

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$I = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$IJ = K = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We also remind you that $\alpha : \mathfrak{M}_2(\mathbb{R}) \to \mathfrak{M}_2(\mathbb{R})$ is given by

$$I \mapsto -J, J \mapsto -J, \mathbf{1} \mapsto \mathbf{1}.$$

Observe that $\alpha(K) = \alpha(IJ) = \alpha(J)\alpha(I) = (JI) = -K$. So, we have

$$\alpha(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The norm N on $\mathfrak{M}_2(\mathbb{R})$ is defined by

$$N(A) \overset{def}{=} A\alpha(A) = (ad - bc)\mathbf{1} = \mathbf{det}(A)\mathbf{1}.$$

Using α, we also had for $A \in \mathfrak{M}_2(\mathbb{R})$,

$$A + \alpha(A) = \mathbf{tr}(A)\mathbf{1}.$$

In this correspondence, we had a subalgebra $\mathfrak{M}_2(\mathbb{R})_0$ of $\mathfrak{M}_2(\mathbb{R})$, i.e.,

$$\mathbb{H}_0 \cong \mathfrak{M}_2(\mathbb{R})_0 \overset{def}{=} \{\textsf{traceless elements of } \mathfrak{M}_2(\mathbb{R})\}.$$

Generally, $\mathfrak{sl}_2(\mathbb{R})$ is often written as $\mathfrak{M}_2(\mathbb{R})_0$. So, from now on, we will freely exchange two notations.

Now, we can think $\mathbb{H}$ has the algebra induced by $\mathfrak{M}_2(\mathbb{R})$. That is, $\mathbb{H}$ has the structure $\mathbb{R}[i,j]/(i^2 = j^2 = 1, ij + ji = 0)$ rather than $\mathbb{R}[i,j]/(i^2 = j^2 = -1, ij + ji = 0)$. So, by the correspondence

$$a + bi + cj + dk \longleftrightarrow \begin{pmatrix} a+b & c+d \\ c-d & a-b \end{pmatrix},$$

we have

$$\mathbb{H}_1 = \{v \in \mathbb{H} \mid a^2 - b^2 - c^2 + d^2 = 1\} \longleftrightarrow \mathbf{SL}(2,\mathbb{R}) = \{A \in \mathfrak{M}_2(\mathbb{R}) \mid \mathbf{det}(A) = 1\}.$$

Note that

$$\mathbb{H}_0 \longleftrightarrow \mathfrak{sl}_2(\mathbb{R}) = \{\textsf{traceless elements of } \mathfrak{M}_2(\mathbb{R})\} = \{\begin{pmatrix} a & b \\ c & -a \end{pmatrix} \mid a,b,c,d \in \mathbb{R}\}.$$

Since each generator of $\mathfrak{M}_2(\mathbb{R})$ consists of integer entries, we can define a $\mathbb{Z}$-algebra, i.e.,

$$\mathfrak{M}_2(\mathbb{Z}) \cong \mathbb{Z}[I,J]/(I^2 = J^2 = \mathbf{1}, IJ + JI = \mathbf{0}).$$

So, in this description, we have successive inclusions of algebras:

$$\mathbf{SL}(2,\mathbb{Z}) \subset \mathbf{SL}(2,\mathbb{R}) \subset \mathfrak{M}_2(\mathbb{R}).$$

The following example will exhibit why this description is so useful sometimes:

Example 1.1.8.1. *Let T^2 be a torus and $\mathbf{Homeo}(T^2)_+$ is a group of orientation preserving homeomorphisms of T^2. Consider*

$$\mathbf{Homeo}(T^2)_+ \overset{\varphi}{\to} \mathbf{Aut}(H_1(T^2)) \text{ by } \varphi(f) = f_*.$$

Since a homeomorphism f induces an isomorphism f_* between $H_1(T^2) \cong \mathbb{Z}^2$ and $H_1(T^2) \cong \mathbb{Z}^2$, φ is well-defined. From $H_1(T^2) \cong \mathbb{Z}^2$, it is not hard to see that

$$\mathbf{Aut}(H_1(T^2)) \cong \mathbf{SL}(2,\mathbb{Z}), \text{ which is called the modular group.}$$

However, since $\mathbb{Z}$ is not a field, sometimes we can not get quite satisfactory information from the situation. So, it would be more useful ways of studying this situation that we think $\mathbf{SL}(2,\mathbb{Z})$ as an imbedding space of $\mathbf{SL}(2,\mathbb{Z})$. In general, a Lie group can be continuously approximated by a discrete Lie group. As a reference,

$$\mathbf{Homeo}(T^2)_+ \cong \mathbf{SL}(2,\mathbb{Z}) \times H \text{ where } h = \{f \in \mathbf{Homeo}(T^2) \mid f \simeq id\}.$$

Also, note that H is called a contractible group.

Remark 1.1.8.1. *Since there exists a Borel measure invariant under both left and right multiplications on* $\mathbf{SL}(2,\mathbb{R})$*, there exists an* $\mathbf{SL}(2,\mathbb{R})$*-invariant measure on a symmetric space* $\frac{\mathbf{SL}(2,\mathbb{R})}{\mathbf{SL}(2,\mathbb{Z})}$ *under the left-multiplication by* $\mathbf{SL}(2,\mathbb{R})$ *on* $\frac{\mathbf{SL}(2,\mathbb{R})}{\mathbf{SL}(2,\mathbb{Z})}$*. If the measure is finite, i.e.,*

$$\mu(\frac{\mathbf{SL}(2,\mathbb{R})}{\mathbf{SL}(2,\mathbb{Z})}) < \infty,$$

it is called a Haar measure on $\frac{\mathbf{SL}(2,\mathbb{R})}{\mathbf{SL}(2,\mathbb{Z})}$.

Exercise 1.1.8.1. *Compute* $\mathbb{H}_1(\mathbb{Z}) = \{h \in \mathbb{H}(\mathbb{Z}) \mid ||h|| = 1\}$.

Proof. First, observe that $\mathbb{H}_1(\mathbb{Z})$ is a discrete space inside a compact space

$$\mathbb{H}_1(\mathbb{R}) \stackrel{def}{=} \mathbb{H}_1 = \{h \in \mathbb{H} \mid ||h|| = 1\} \cong S^3.$$

So, it must consist of a finite number of points. Clearly,

$$\{\pm\mathbf{1}, \pm I, \pm J, \pm K\} \subset \mathbb{H}_1(\mathbb{Z}).$$

Since $\{\pm\mathbf{1}, \pm I, \pm J, \pm K\}$ is the set of all the generators of $\mathbb{H}_1$ and $h \in \mathbb{H}_1(\mathbb{Z})$ must satisfy $||h|| = 1$, we have

$$\{\pm\mathbf{1}, \pm I, \pm J, \pm K\} = \mathbb{H}_1(\mathbb{Z}).$$

□

Recall that in Exercise 1.1.3.1, $\mathbb{H}_1$ acts on $\mathbb{H}_0$ by conjugations in unimodular and orthogonal ways. It is clear that $\mathbb{H}^\times$ acts on $\mathbb{H}_0 = \mathbb{R}^3 \cong \mathfrak{sl}(2,\mathbb{R})$ by conjugations in nonsingular ways, considering the proof of Exercise 1.1.3.1. That is, it is the usual action of an element in $\mathbf{GL}(3,\mathbb{R})$ on $\mathbb{R}^3$.

Observation 1.1.8.1. *We know that* $\mathbb{H}_0 \cong \mathbb{R}^3$ *as vector spaces. Here,* $\mathbb{R}^3$ *is a vector space generated by* $\{i, j, k\}$*. Since we show that* $\mathbb{H}_0 \cong \mathfrak{sl}(2,\mathbb{R})$*, we can give* $\mathbb{R}^3$ *the algebra structure induced by* $\mathfrak{sl}(2,\mathbb{R})$*. That is, we have the following relations in* $\mathbb{R}^3$*:*

$$i^2 = j^2 = 1, ij + ji = 0, k^2 = -1.$$

More precisely, we defined a multiplicative structure on $\mathbb{R}^3$*: For* $v_1, v_2 \in \mathbb{R}^3$*,*

$$v_1 v_2 \in \mathbb{R}^3 \cong \mathfrak{sl}(2,\mathbb{R}).$$

So, we can define a new inner product $\langle , \rangle$ *on* $\mathbb{R}^3$*, i.e., for* $v_1, v_2 \in \mathbb{R}^3$*,*

$$\langle v_1, v_2 \rangle = Re(v_1 v_2).$$

Let

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

One moment of thinking gives you that the new inner product on $\mathbb{R}^3$ can be written as the usual quadratic forms: For $v_1, v_2 \in \mathbb{R}^3$,

$$\langle v_1, v_2 \rangle = v_1^T \cdot B \cdot v_2.$$

Let

$$G(B) = \{A \in \mathbf{M}_3(\mathbb{R}) \mid A^T BA = B\}.$$

Obviously, $G(B)$ is a subgroup of $\mathbf{M}_3(\mathbb{R})$. Since $\mathbf{det}(A^T BA) = -(\mathbf{det}(A))^2 = -1$ for $A \in G(B)$, we have $G(B) \subset \mathbf{GL}(3, \mathbb{R})$. Moreover, from the following consideration;

$$\langle v_1, v_2\rangle = v_1^T \cdot B \cdot v_2 = v_1^T \cdot (A^T BA) \cdot v_2 = \langle Av_1, Av_2\rangle \text{ for } A \in G(B),$$

we have $G(B)$ preserves the orthonormal relations with respect to $\langle , \rangle$. So, we call $G(B)$ an indefinite orthonormal group, denoted by $\mathbf{O}(2, 1)$. In general,

$$B = \begin{pmatrix} \mathbf{1}_p & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_q \end{pmatrix} \text{ where } \mathbf{1}_p \text{ is a } p \times p \text{ identity matrix.}$$

Also we have $G(B) \overset{def}{=} \mathbf{O}(p, q) \subset \mathbf{GL}(n, \mathbb{R})$ where $p + q = n$.

Definition 1.1.8.1 (Local Lie group isomorphisms). *If G, H are Lie groups, a* ***local Lie group isomorphism*** *is a map $f : G \to H$ such that f is a homomorphism of groups and f is a local diffeomorphism.*

Example 1.1.8.2. *Exercise 1.1.3.1 show that $\mathbb{H}_1 \cong S^3 \to \mathbf{SO}(3)$ is a local isomorphism and surjective as a group homomorphism. Note that the algebra of $\mathbb{H}_1$ is given by $\mathbb{R}[i, j]/(i^2 = j^2 = -1, ij + ji = 0)$.*

In general, a local isomorphism is not injective nor surjective.

Example 1.1.8.3. $\mathbf{SL}(2, \mathbb{R}) \overset{f}{\hookrightarrow} \mathbf{O}(2, 1)$ *is a local isomorphism.*

Proof. Let the algebra of $\mathbb{H}$ be given by $\mathbb{R}[i, j]/(i^2 = j^2 = 1, ij + ji = 0)$. A similar proof of Exercise 1.1.3.1 shows that the action of $\mathbb{H}^\times$ on $\mathbb{H}_0$ corresponds to the conjugation action of $\mathbf{GL}(3, \mathbb{R})$ on $\mathfrak{sl}(2, \mathbb{R})$. Moreover, it gives that the action of $\mathbb{H}_1$ on $\mathbb{H}_0$ corresponds to the conjugation action of $\mathbf{SL}(2, \mathbb{R})$ on $\mathfrak{sl}(2, \mathbb{R})$. So, we have $\mathbf{SL}(2, \mathbb{R}) \subset \mathbf{GL}(3, \mathbb{R})$. Let $u = a + bi + cj + dk \in \mathbb{H}_1$. So, $a^2 - b^2 - c^2 + d^2 = 1$. For $v_1 = \alpha_1 i + \beta_1 j + \gamma_1 k, v_2 = \alpha_2 i + \beta_2 j + \gamma_2 k \in \mathbb{H}_0$, we have

$$\langle uv_1u^{-1}, uv_2u^{-1}\rangle = Re(uv_1v_2u^{-1}) = (a^2 - b^2 - c^2 + d^2)Re(v_1v_2) = \langle v_1, v_2\rangle.$$

That is, $\mathbf{SL}(2, \mathbb{R}) \subset \mathbf{O}(2, 1)$.

Note that

$$\begin{aligned} Re\big(uv_1v_2u^{-1}\big) &= Re((a + Im(u))(Re(v_1v_2) + Im(v_1v_2))(a - Im(u))) \\ &= Re\big((a + Im(u))(Re(v_1v_2))(a - Im(u)) + (a + Im(u))(Im(v_1v_2))(a - Im(u))\big) \\ &= Re((a^2 - b^2 - c^2 + d^2)Re(v_1v_2) + a^2 Im(v_1v_2) - aIm(v_1v_2)Im(u) \\ &+ aIm(u)Im(v_1v_2) - Im(u)Im(v_1v_2)Im(u)) \\ &= (a^2 - b^2 - c^2 + d^2)Re(v_1v_2) - Re(aIm(v_1v_2)Im(u)) + aRe(Im(u)Im(v_1v_2)) \\ &= (a^2 - b^2 - c^2 + d^2)Re(v_1v_2) \end{aligned}.$$

The local isomorphism is not surjective. □

Even though $\mathbf{SO}(2) \cong S^1$ is connected, we have the followings:

Exercise 1.1.8.2. *Prove that* $\mathbf{SO}(1,1)$ *is not connected.*

Proof. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SO}(1,1) = \{A \in \mathbf{M}_3(\mathbb{R}) \mid A^T BA = B\}.$$

Since $A^T BA = B$, we have four equations:

$$a^2 - c^2 = 1, b^2 - d^2 = -1, ab = cd, \text{and } ad - bc = 1.$$

So, if $d = 0$, we have a contradiction. So, we have $d \neq 0$. Since $d \neq 0$, we must have $b = c$. If $b = c = 0$, then $a = d = 1$. If $b = c \neq 0$, then $a = d$. Hence,

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \text{ where } a^2 - b^2 = 1.$$

That is, $\mathbf{SO}(1,1)$ is not connected. □

1.1.9 Exponential Map

Recall that $\mathbb{H} = \mathbb{R} \oplus \mathbb{H}_0$. So if $w \in \mathbb{H}$, then we can write $w = a \cdot 1 + v$ for $a \in \mathbb{R}$ and $v \in \mathbb{H}_0$. From this, we define

$$\exp(w) = \exp(a \cdot 1 + v) = e^a \exp(v) = e^a \big(\cos(||v||) + \sin(||v||)\frac{v}{||v||}\big).$$

Exercise 1.1.9.1. *Show that for* $q = a + bi + cj + dk = a + v \in \mathbb{H}$, *we have*

$$||\exp(q)|| = e^{\mathbf{tr}(q)}.$$

Proof. By identifying $\mathbb{H} \cong \mathfrak{M}_2(\mathbb{R})$, we have

$$q = a + bi + cj + dk \to \begin{pmatrix} a+b & c+d \\ c-d & a-b \end{pmatrix}$$

So, $\mathbf{tr}(q) = 2a$. In the other hand,

$$\begin{aligned} ||\exp(q)||^2 &= \exp(q) \cdot \alpha(\exp(q)) \\ &= e^a \big(\cos(||v||) + \sin(||v||)\frac{v}{||v||}\big) \cdot e^a \big(\cos(||v||) - \sin(||v||)\frac{v}{||v||}\big) \\ &= e^{2a}\big(\cos^2(||v||) + \sin^2(||v||)\big) = e^{2a} = e^{2\mathbf{tr}(q)}. \end{aligned}$$

□

From Exercise 1.1.9.1, we deduce that traceless quaternions correspond to unit numbers. Since $\mathbb{H}_0 \cong \mathfrak{sl}(2,\mathbb{R}) = \{\text{ traceless elements of } \mathfrak{M}_2(\mathbb{R})\}$, we have a well-defined map $\exp : \mathbb{H}_0 \to \mathbb{H}_1$.

Exercise 1.1.9.2. $\exp : \mathbb{H}_0 \to \mathbb{H}_1 \cong S^3$ *is surjective, but not injective.*

Proof. Since $\mathbb{H}_0$ is not compact and $\mathbb{H}_1 \cong S^3$ is compact, obviously it is not injective.

Let $u = a + bi + cj + dk \in \mathbb{H}_1$. So, $a^2 + b^2 + c^2 + d^2 = 1$.

If $a = \pm 1$, $\exp(2\pi i) = 1$ and $\exp(\pi i) = -1$. So, we can assume $a \neq 1$. Let

$$v = \frac{\cos^{-1} a}{\sqrt{1 - a^2}}(bi + cj + dk) \in \mathbb{H}_0.$$

Note that $||v|| = \cos^{-1} a$ and $\cos^2(\cos^{-1} a) + \sin^2(\cos^{-1} a) = a^2 + \sin^2(\cos^{-1} a) = 1$. So,

$$\exp(v) = \cos(\cos^{-1} a) + \sin(\cos^{-1} a)\frac{bi + cj + dk}{\sqrt{1 - a^2}} = a + bi + cj + dk.$$

□

Exercise 1.1.9.3. *Show that* $\exp : \mathfrak{sl}(2, \mathbb{R}) \to \mathbf{SL}(2, \mathbb{R})$ *is not surjective.*

Proof. First note that $\exp(A) \overset{def}{=} \sum \frac{1}{k!} A^k$. It is easy to see that it converges. By that fact $\mathbf{det}(\exp(A)) = e^{\mathbf{tr}(A)}$, we conclude that exp is well-defined. Consider

$$A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \in \mathbf{SL}(2, \mathbb{R}).$$

Suppose that $\exp(X) = A$ for $X \in \mathfrak{sl}(2, \mathbb{R})$. Since $X \in \mathfrak{sl}(2, \mathbb{R})$,

$$X = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}.$$

If $a^2 - bc = 0$, then $X^2 = 0$. So,

$$\exp(X) = \mathbf{1} + X = \begin{pmatrix} 1 + a & b \\ c & 1 - a \end{pmatrix}.$$

Hence, $a^2 - bc = 0$ implies that $a = 0, c = 0$, which is a contradiction.

If $a^2 - bc \neq 0$, then $X^2 = (a^2 + bc)\mathbf{1}$. By the Jordan canonical form, we can assume

$$PXP^{-1} = \begin{pmatrix} \lambda & 1 \\ 0 & -\lambda \end{pmatrix} \text{ or } PXP^{-1} = \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} \text{ where } P \in \mathbf{GL}(2, \mathbb{C}).$$

However, $X^2 = (a^2 + bc)\mathbf{1}$ implies that we must have

$$PXP^{-1} = \begin{pmatrix} \sqrt{a^2 + bc} & 0 \\ 0 & -\sqrt{a^2 + bc} \end{pmatrix}.$$

Since

$$\begin{aligned} P \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} P^{-1} &= PAP^{-1} = \exp(PXP^{-1}) \\ &= \exp(\begin{pmatrix} \sqrt{a^2 + bc} & 0 \\ 0 & -\sqrt{a^2 + bc} \end{pmatrix}) = \begin{pmatrix} e^{\sqrt{a^2+bc}} & 0 \\ 0 & e^{-\sqrt{a^2+bc}} \end{pmatrix}, \end{aligned}$$

we have a contradiction. So, it is not surjective. □

Question 1.1.9.1. *What about* $\exp : \mathfrak{sl}(2, \mathbb{C}) \to \mathbf{SL}(2, \mathbb{C})$*?*

Proof. The same proof of Exercise 1.1.9.3 works here, So, it is not surjective. □

1.2 Various aspects of Lie groups

1.2.1 Revisit to the exponential map

Definition 1.2.1.1. *A Lie algebra* $\mathfrak{g}$ *over* $\mathbb{R}$ *is a real* ***vector space*** $\mathfrak{g}$ *together with a skew-symmetric bilinear operator*

$$[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \text{ which satisfies Jacobi identity.}$$

First of all, we define a map exp from $T(\mathbb{R}) \cong \mathbb{R}$ as a Lie algebra to a Lie algebra $\mathfrak{g}$ of a Lie group G, i.e., pick $X \in \mathfrak{g}$ and define

$$\exp : \mathbb{R} \to \mathfrak{g} \text{ by } \lambda \frac{d}{dt} \to \lambda X.$$

Then, since $\mathbb{R}$ is simply connected, there is a unique one-parameter subgroup such that

$$t \to \exp(tX) \in G.$$

Note that this map only depends on a vector field, which we chose. So, we will denote $\exp(tX)$ by $\exp_X(t)$. From this one-parameter subgroup, we define the exponential map exp from the Lie algebra $\mathfrak{g}$ to a Lie group G

$$\exp : \mathfrak{g} \to G \text{ by setting } \exp(X) = \exp_X(1).$$

Take G to be a group of matrixes. We define

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Note that later we will see that two constructions are the same. If $A = (a_{ij})_{n\times n}$, let

$$||A|| = \sum_{i,j} |a_{ij}|.$$

Clearly, we have $||AB|| \le ||A|| ||B||$. Also, there exists $\delta > 0$ such that $||A|| < \delta$. Since

$$|| \exp(A)|| \le \sum_{n=0}^{\infty} \frac{||A^n||}{n!} \le \sum_{n=0}^{\infty} \frac{||A||^n}{n!} < \sum_{n=0}^{\infty} \frac{\delta^n}{n!} \to e^{\delta} \text{ as } n \to \infty,$$

we can deduce that all entries of $\exp(A)$ converge uniformly and absolutely. There are useful formulas:
1. If X, Y commute, then $\exp(X + Y) = \exp(X)\exp(Y)$.
2. $\exp(X)\exp(-X) = \mathbf{1}$. So, The inverse matrix of $\exp(X)$ is $\exp(-X)$.
3. $\mathbf{det}(\exp(X)) = \exp(\mathbf{tr}(X))$. Note that we already proved this in terms of quaternions, i.e., this is true for $\mathbf{M}_2(\mathbb{C})$. For generality, we prove this for $\mathbf{M}_n(\mathbb{C})$.

Proof. The Jordan canonical form theorem tells us that for any matrix A over $\mathbb{C}$, we can find P such that $P^{-1}AP$ can be expressed as a block diagonal matrix with Jordan blocks along the diagonal. Moreover, it is unique up to a permutation. So, it suffices to prove the claim for one Jordan block. Since $\mathbf{det}(P^{-1}AP) = \mathbf{det}(A)$ and $\mathbf{tr}(P^{-1}AP) = \mathbf{tr}(A)$, without loss of generality, we can assume $A = \lambda I_n + N$ where I_n is an identity matrix and N is a matrix with 1's on the first superdiagonal and zeros elsewhere. So, $\mathbf{tr}(A) = n\lambda$. Hence, $\exp(\mathbf{tr}(A)) = e^{n\lambda}$. In the other hands,

$$\exp(A) = \exp(\lambda I_n + N) = e^{\lambda} I_n \cdot \exp(N).$$

It is easy to see that $\exp(N)$ is a unipotent matrix, i.e., an upper diagonal matrix with 1's along the diagonal. Hence, $\mathbf{det}(\exp(A)) = e^{n\lambda}$. $\square$

Example 1.2.1.1. $\mathbb{H} \overset{\exp}{\to} \mathbb{H}^{\times}$ *is surjective but not injective.*

Proof. Obviously, $\exp(2\pi i) = \exp(2\pi j) = 1$. So, it is not injective. We know that $\mathbb{H}_0 \overset{\exp}{\to} \mathbb{H}_1$ is surjective. That is, given $v \in \mathbb{H}_0$, there exists $u \in \mathbb{H}_1$ such that $\exp(v) = u$. Since for $w \in \mathbb{H}^{\times} = \mathbb{R}^4/\{0\}$ there exists $u \in \mathbb{H}_1 \cong S^3$ such that $w = ru$ for some $r \in \mathbb{R}$, we have

$$\exp(\log r + v) = e^{\log r} \cdot \exp(v) = ru = w.$$

So, it is surjective. $\square$

Example 1.2.1.2. $\mathfrak{gl}(2, \mathbb{C}) \cong \mathbf{M}_2(\mathbb{C}) \overset{\exp}{\to} \mathbf{M}_2(\mathbb{C})^{\times} \overset{def}{=} \mathbf{GL}(2, \mathbb{C})$ *is surjective.*

Proof. Notice that by formula 2, $\exp(X)$ is invertible. Let $A \in \mathbf{M}_2(\mathbb{C})^{\times}$. By the Jordan canonical form theorem, there exists $P \in \mathbf{GL}(2, \mathbb{C})$ such that

$$PAP^{-1} = \begin{pmatrix} \lambda & 0 \\ 0 & \omega \end{pmatrix} \text{ or } PAP^{-1} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Suppose that

$$PAP^{-1} = \begin{pmatrix} \lambda & 0 \\ 0 & \omega \end{pmatrix}.$$

Let

$$X = \begin{pmatrix} \log \lambda & 0 \\ 0 & \log \omega \end{pmatrix}.$$

We have

$$\exp(P^{-1}XP) = P^{-1} \begin{pmatrix} e^{\log \lambda} & 0 \\ 0 & e^{\log \omega} \end{pmatrix} P = A.$$

Suppose that

$$PAP^{-1} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Let

$$Y = \begin{pmatrix} 0 & \lambda^{-1} \\ 0 & 0 \end{pmatrix}.$$

We have

$$\exp(P^{-1}YP + (\log\lambda)\mathbf{1}) = P^{-1}(\exp(Y))P \cdot \begin{pmatrix} e^{\log\lambda} & 0 \\ 0 & e^{\log\omega} \end{pmatrix}$$
$$= P^{-1}\begin{pmatrix} 1 & \lambda^{-1} \\ 0 & 1 \end{pmatrix} P \cdot \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = P^{-1}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} P = A.$$

So, it is surjective. □

Example 1.2.1.3. $\mathfrak{g} = i\mathbb{R} \to S^1 \subset \mathbb{C}$ *given by* $i\theta \to \exp(i\theta)$ *is surjective.*

Example 1.2.1.4. *In Exercise 1.1.9.2, we show that* $\mathfrak{g} = \mathbb{H}_0 \overset{\exp}{\to} \mathbb{H}_1 \cong S^3$ *is surjective.*

So far, we have seen that $\mathfrak{g} \overset{\exp}{\to} G$ is surjective. However, there is an exp, which is not surjective. From $\mathbf{det}(\exp(A)) = \exp(\mathbf{tr}(A)$, we have the followings:

Example 1.2.1.5. *In Exercise 1.1.9.3, we have seen that* $\exp : \mathfrak{sl}(2,\mathbb{R}) \to \mathbf{SL}(2,\mathbb{R})$ *is not surjective. By considering the proof of Exercise 1.1.9.3, we have the followings: If* $\lambda \neq 0$*, then*

$$\begin{pmatrix} -1 & \lambda \\ 0 & -1 \end{pmatrix} \notin \exp(\mathfrak{sl}(2,\mathbb{R})).$$

Now we give an analytic definition of Lie algebra.

Let G be a Lie group contained in $\mathbf{M}_n(\mathbb{R})$. The Lie algebra $\mathfrak{g}$ of G is a tangent space at an identity $\mathbf{1}$. That is,

$$\mathfrak{g} \overset{def}{=} T_{\mathbf{1}}(G) \subset T_{\mathbf{1}}(\mathbf{M}_n(\mathbb{R})).$$

A left invariant vector field X is uniquely determined by the evaluation at $\mathbf{1}$, i.e., $X_{\mathbf{1}}$. So, from now on we denote $X_{\mathbf{1}}$ by X. Look at a smooth path φ_t in G starting at $\mathbf{1}$, i.e.,

$$\varphi_t \in G \text{ and } \varphi_0 = \mathbf{1}.$$

Let $X = \frac{d\varphi_t}{dt}|_{t=0} \in \mathfrak{g} \overset{def}{=} T_{\mathbf{1}}(G)$. Define a map

$$\exp : \mathfrak{g} \to G \text{ by } tX \to \varphi_t.$$

This is the exactly same description which we mentioned earlier.

Example 1.2.1.6. *Let* $\varphi : \mathbb{R} \to \mathbb{H}$ *be a one-parameter subgroup with* $\varphi_0 = 1$ *and* $\dot{\varphi_0} = X$*. Suppose that* $||\varphi_t||^2 = 1$*. So,* $\varphi_t\alpha(\varphi_t) = 1$*. Differentiate* $\varphi_t\alpha(\varphi_t) = 1$ *with respect to* t *using linearity of* α*;*

$$\dot{\varphi_t}\alpha(\varphi_t) + \varphi_t\alpha(\dot{\varphi_t}) = 0.$$

Set $t = 0$*. We have*

$$\dot{\varphi_0}\alpha(\varphi_0) + \varphi_0\alpha(\dot{\varphi_0}) = X + \alpha(X) = 0, \textit{ which shows that } X \textit{ is necessarily traceless.}$$

That is, the equation $||\varphi_t||^2 = 1$ *defines the exponential map* $\exp(tX) = \varphi_t$*, i.e.,*

$$\mathbb{H}_0 \overset{\exp}{\to} \mathbb{H}_1.$$

1.2.2 Symplectic group

Let V be a vector space over $\mathbb{R}$. Then a bilinear form $V \times V \xrightarrow{\beta} \mathbb{R}$ is skew-symmetric if

$$\beta(v_1, v_2) = -\beta(v_2, v_1).$$

Since there is a unique matrix such that

$$\beta(v_1, v_2) = v_1^T B v_2,$$

we also have $B = (\beta_{ij})_{n\times n}$ is a skew-symmetric matrix if $B = -B^T$ where $\beta_{ij} = \beta(e_i, e_j)$ in terms of a basis $\{e_1, \dots, e_n\}$. So, we have three correspondences:

$$\Lambda^2 V^* = \{\text{skew-symmetric bilinear forms on } V \text{ of dimension } n\}$$
$$\Updownarrow$$
$$\{\text{skew-symmetric } n \times n \text{ matirxs}\}$$
$$\Updownarrow$$
$$\mathfrak{o}(n) \text{ the Lie algebra of } \mathbf{O}(n).$$

We only need some clarification of the second up-and-down arrow. An element of the Lie algebra $\mathfrak{g}$ of a Lie group G is uniquely determined by one-parameter subgroup. With keeping this in your mind, we have;

Proof. Let $\varphi : \mathbb{R} \to \mathbf{M}_n(\mathbb{R})$ be a one-parameter subgroup with $\varphi_0 = \mathbf{1}$ and $\dot{\varphi_0} = X$. Suppose that $\varphi_t \cdot \varphi_t^T = \mathbf{1}$. Differentiate $\varphi_t \cdot \varphi_t^T = \mathbf{1}$ with respect to t using linearity of matrix transpose;

$$\dot{\varphi_t} \cdot \varphi_t^T + \varphi_t \cdot (\dot{\varphi_t})^T = 0.$$

Set $t = 0$. We have

$$\dot{\varphi_0} \cdot \varphi_0^T + \varphi_0 \cdot (\dot{\varphi_0})^T = X + X^T = 0, \text{ which shows that } X \text{ is skew-symmetric.}$$

That is, the Lie algebra $\mathfrak{o}(n)$ of a Lie group $\mathbf{O}(n) = \{X \in \mathbf{M}_n(\mathbb{R}) \mid X \cdot X^T = \mathbf{1}\}$ is a set of skew-symmetric matrixes. □

Remark 1.2.2.1. *Another interesting relation is the following:*

$$\{\textit{bilinear forms on } \mathfrak{g} = T_1(G)\}$$
$$\Updownarrow$$
$$\{\textit{left-invariant metrics on } G\}$$
$$\Updownarrow$$
$$\{\ \textit{right-invariant metrics on } G\ \}.$$

Note that this doesn't imply that there exists a bi-invariant metric on G. For an example, we give $\mathbf{SL}(2, \mathbb{R})$.

Let $V = \mathbb{R}^2$ and β be a skew-symmetric bilinear form. Since

$$\beta(e_1, e_1) = -\beta(e_1, e_1), \beta(e_1, e_2) = -\beta(e_2, e_1), \text{ and } \beta(e_2, e_2) = -\beta(e_2, e_2),$$

the correspondence matrix B of β is

$$B = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} \text{ where } a \in \mathbb{R}.$$

So, we have

$$\begin{array}{c} \Lambda^2 V^* = \{\text{skew-symmetric bilinear forms on } V \text{ of dimension } 2\} \\ \Updownarrow \\ \{aJ \mid J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } a \in \mathbb{R}\} \end{array}.$$

That is, J spans $\Lambda^2 V^*$. Since $A^T JA = (\mathbf{det} A)J$ is skew-symmetric, we can define an action of $\mathbf{GL}(2, \mathbb{R}) \cong \mathbf{Aut}(V)$ on $\Lambda^2 V^*$ by the following way:

$$A : J \to A^T JA.$$

We can observe two interesting properties of J:
1. $J^2 = -\mathbf{1}$.
2. $J^{-1} = -J$.
As consequences of these, we have
1. $A(-J)A^T J = (\mathbf{det} A)\mathbf{1} \Longrightarrow \alpha(A) = (-J)A^T J = J^{-1} A^T J$ from $A\alpha(A) = (\mathbf{det} A)\mathbf{1}$.
2.

$$A^{-1} = \frac{1}{(\mathbf{det} A)}(-J)A^T J \text{ for } A \in \mathbf{GL}(2, \mathbb{R}).$$

In general, we let

$$J_{2n} = \begin{pmatrix} 0 & -\mathbf{1}_n \\ \mathbf{1}_n & 0 \end{pmatrix} \text{ where } \mathbf{1}_n \text{ is an } n \times n \text{ identity matrix.}$$

We define a real symplectic group to be

$$\mathbf{Sp}(n, \mathbb{R}) = \{A \in \mathbf{M}_{2n}(\mathbb{R}) \mid A^T J_{2n} A = J_{2n}\}.$$

Likewise, we have

$$\mathbf{Sp}(n, \mathbb{C}) = \{A \in \mathbf{M}_{2n}(\mathbb{C}) \mid A^T J_{2n} A = J_{2n}\}.$$

From $A^T JA = (\mathbf{det} A)J$, we can see easily

$$\mathbf{SL}(2, \mathbb{R}) = \mathbf{Sp}(1, \mathbb{R}).$$

We define

$$\mathbf{1}_{p,q} = \mathbf{1}_p \oplus -\mathbf{1}_q = \begin{pmatrix} \mathbf{1}_p & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_q \end{pmatrix}.$$

From this,

$$\mathbf{U}(p, q) \overset{def}{=} \{A \in \mathbf{M}_n(\mathbb{C}) \mid \overline{A}^T \cdot \mathbf{1}_{p,q} \cdot A = \mathbf{1}_{p,q} \text{ where } p + q = n\}.$$

So,

$$\mathbf{U}(n) = \mathbf{U}(n, 0).$$

If $p + q = 2n$, then

$$\mathbf{Sp}(p, q) \overset{def}{=} \mathbf{Sp}(n, \mathbb{C}) \cap \mathbf{U}(p, q).$$

Exercise 1.2.2.1. *Show that*

$$\mathbf{SL}(2,\mathbb{R}) \cong \mathbf{SU}(1,1).$$

Proof. Pick $C \in \mathbf{SL}(2,\mathbb{C})$ such that

$$\overline{C}^T \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot C = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Of course,

$$C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}.$$

Define

$$\mathbf{SL}(2,\mathbb{R}) \xrightarrow{f} \mathbf{SU}(1,1) \text{ by } A \mapsto CAC^{-1}.$$

If f is well-defined, f is necessarily an isomorphism. Clearly, $\mathbf{det}(CAC^{-1}) = 1$. Note that

$$A^T \cdot \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cdot A = iA^TJA = iJ = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ for } A \in \mathbf{SL}(2,\mathbb{R}) = \mathbf{Sp}(1,\mathbb{R}).$$

So, we have

$$\begin{aligned} \overline{CAC^{-1}}^T \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot CAC^{-1} &= \overline{C^{-1}}^T A^T \overline{C}^T \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot CAC^{-1} = \overline{C^{-1}}^T A^T \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} AC^{-1} \\ &= \overline{C^{-1}}^T \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cdot C^{-1} = \overline{C^{-1}}^T \overline{C}^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} CC^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

□

A moment later, we ask you to prove the following identification:

$$\mathbf{GL}(n,\mathbb{H}) \subset \mathbf{GL}(2n,\mathbb{C}).$$

Before proving this, we give an example.

Example 1.2.2.1. *We have*

$$\mathbf{M}_n(\mathbb{C}) \subset \mathbf{M}_{2n}(\mathbb{R}).$$

Proof. The main point here is that we can identify a $\mathbb{R}$-linear homomorphism with a $\mathbb{C}$-linear homomorphism and vice versa. Let $z_{ij} = x_{ij} + \sqrt{-1}y_{ij}$ where $x_{ij}, y_{ij} \in \mathbb{R}$. We have

$$\begin{pmatrix} z_{11} & \dots & z_{1n} \\ \vdots & \dots & \vdots \\ z_{n1} & \dots & z_{nn} \end{pmatrix} \mapsto \begin{pmatrix} \begin{pmatrix} x_{11} & -y_{11} \\ y_{11} & x_{11} \end{pmatrix} & \dots & \begin{pmatrix} x_{1n} & -y_{1n} \\ y_{1n} & x_{1n} \end{pmatrix} \\ \vdots & \dots & \vdots \\ \begin{pmatrix} x_{n1} & -y_{n1} \\ y_{n1} & x_{n1} \end{pmatrix} & \dots & \begin{pmatrix} x_{nn} & -y_{nn} \\ y_{nn} & x_{nn} \end{pmatrix} \end{pmatrix}$$

□

Exercise 1.2.2.2. *Considering $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}j \cong \mathbb{C}^2$, construct the following inclusion*

$$\mathbf{GL}(n, \mathbb{H}) \subset \mathbf{GL}(2n, \mathbb{C}).$$

Proof. The proof of this example is almost same as the above proof except one thing. The difference occurs because j and $\mathbb{C}$ does not commute as opposed to the fact $\sqrt{-1}$ and $\mathbb{R}$ commute. We will show that we can identify a $\mathbb{C}$-linear homomorphism with a $\mathbb{H}$-linear homomorphism and vice versa. That is,

$$\mathbf{M}_n(\mathbb{H}) \subset \mathbf{M}_{2n}(\mathbb{C}).$$

Let $u \in \mathbb{H}$. So, $u = v + wj$ where $v, w \in \mathbb{C}$. Suppose that there exists an inclusion

$$v + wj \in \mathbb{H} = \mathbf{M}_1(\mathbb{H}) \to \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{M}_2(\mathbb{C}).$$

Note that

$$(\alpha + \beta j) \cdot (v + wj) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

First, we have

$$(v + wj) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Longrightarrow a = v, c = w.$$

Second, consider

$$ju = j(v + wj) = jv + jwj = \overline{v}j - \overline{w}.$$

That is,

$$-\overline{w} + \overline{v}j = j(v + wj) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Longrightarrow b = -\overline{w}, d = \overline{v}.$$

So, by defining an inclusion

$$v + wj \in \mathbb{H} = \mathbf{M}_1(\mathbb{H}) \to \begin{pmatrix} v & -\overline{w} \\ w & \overline{v} \end{pmatrix} \in \mathbf{M}_2(\mathbb{C}),$$

We have an $\mathbb{H}$-linear matrix in $\mathbf{M}_2(\mathbb{C})$. It is easy to see that

$$\mathbf{M}_1(\mathbb{H}) = \{A \in \mathbf{M}_2(\mathbb{C}) \mid AJ_2 = J_2\overline{A}\}.$$

In general, we identify $(u_1, \ldots, u_n) \in \mathbb{H}^n$ for $u_i = v_i + w_i j$ with

$$(v_1, \ldots, v_n, w_1, \ldots, w_n) \in \mathbb{C}^{2n}.$$

Let $u_{ik} = v_{ik} + w_{ik}j$ where $v_{ij}, w_{ij} \in \mathbb{C}$. Define

$$\begin{pmatrix} u_{11} & \ldots & u_{1n} \\ \vdots & \ldots & \vdots \\ u_{n1} & \ldots & u_{nn} \end{pmatrix} \mapsto \begin{pmatrix} v_{11} & \ldots & v_{1n} & -\overline{w_{11}} & \ldots & -\overline{w_{1n}} \\ \vdots & & \vdots & \vdots & & \vdots \\ v_{n1} & \ldots & v_{nn} & -\overline{w_{n1}} & \ldots & -\overline{w_{nn}} \\ w_{11} & \ldots & w_{1n} & \overline{v_{11}} & \ldots & \overline{v_{1n}} \\ \vdots & & \vdots & \vdots & & \vdots \\ w_{n1} & \ldots & w_{nn} & \overline{v_{n1}} & \ldots & \overline{v_{nn}} \end{pmatrix}$$

It is easy to see that

$$\mathbf{M}_n(\mathbb{H}) = \{A \in \mathbf{M}_{2n}(\mathbb{C}) \mid AJ_{2n} = J_{2n}\overline{A}\}.$$

Of course, in this description, we also have

$$\mathbf{GL}(n, \mathbb{H}) = \{A \in \mathbf{GL}(2n, \mathbb{C}) \mid AJ_{2n} = J_{2n}\overline{A}\}.$$

□

Example 1.2.2.2.

$$\mathbf{Sp}(n, 1) \Longleftrightarrow \mathbf{Isom}(\ \textit{quaternion hyperbolic } n\textit{-space}\ ).$$

$$\mathbf{U}(n, 1) \Longleftrightarrow \mathbf{Isom}(\ \textit{complex hyperbolic } n\textit{-space}\ ).$$

Exercise 1.2.2.3. *Define* $\mathbf{SL}(n, \mathbb{H})$.

Proof. By the previous hard work, it is easy to see that

$$\mathbf{SL}(n, \mathbb{H}) = \{A \in \mathbf{GL}(2n, \mathbb{C}) \mid AJ_{2n} = J_{2n}\overline{A} \text{ and } \mathbf{det}(A) = 1\}.$$

□

1.2.3 Transformation groups

Let G be a group and X be a space. Define

$$G \times X \stackrel{\alpha}{\to} X \text{ by } (g, x) \to gx.$$

We call the image of $\alpha(G \times \{x\}) = Gx$ the orbit of x.

G acts **transitively** on X if and only if $Gx = X$ for some x. In turns, equivalently, $Gx = X$ for some x if and only if $Gx = X$ for all x. In this case, we say X is a **homogeneous space** of G.

We call $\mathbf{Stab}(x, G) = G^x = \{g \in G \mid gx = x\}$ the isotropy group of x.

G acts **freely** on X if for all $x \in X$, $G^x = \{1\}$. Equivalently, $gx = x$ implies $g = 1$.

G acts **effectively** (**faithfully**, injectively) on X if $G \to \mathbf{Aut}(X)$ is injective. Equivalently, if $gx = x$ for all $x \in X$, then $g = 1$.

We call $G \stackrel{e_x}{\to} X$ the evaluation of x given by $e_x(g) = gx$. The image of e_x is the orbit of x.

We say f is G-equivariant with respect to left-multiplication on G and (left) action of G on X if the diagram commute: $f(l_g(u)) = g \cdot f(u)$, i.e.,

$$\begin{array}{ccc} G & \stackrel{f}{\longrightarrow} & X \\ {\scriptstyle l_g}\downarrow & & \downarrow{\scriptstyle g} \\ G & \stackrel{f}{\longrightarrow} & X \end{array}$$

Example 1.2.3.1. *Since* $e_x(l_g(u)) = gux = g \cdot e_x(u)$, e_x *defines* G*-equivariant with respect to left-multiplication on* G *and (left) action of* G *on* X*, i.e.,*

$$\begin{array}{ccc} G & \stackrel{e_x}{\longrightarrow} & X \\ {\scriptstyle l_g}\downarrow & & \downarrow{\scriptstyle g} \\ G & \stackrel{e_x}{\longrightarrow} & X \end{array}$$

It is easy to see that $G/G^x \xrightarrow{F} Gx$ by $F(gG^x) = gx$ is a well-defined bijective map. Since if $h \in G^x$ then $(gh)x = g(hx) = gx$, it is well-defined. Bijectivity is obvious.

G acts **transitively** on X if and only if $G/G^x \xrightarrow{\cong} X$.

G acts **freely** on X if and only if $G \xrightarrow{\cong} Gx$.

G acts **freely transitively** ($\overset{def}{=}$ **simply transitively**) on X if and only if $G \underset{\cong}{\xrightarrow{e_x}} X$.

1.2.4 Projective spaces

An affine space X is a space with a simply transitive action of a vector group, i.e., a vector space under addition.

Let k be a field and V be an n-dimensional vector space over k. Since $\mathbf{GL}(V) = \mathbf{GL}(n,k)$ acts linearly on V, a linear action preserves a line, i.e., one-dimensional subvector space of V. So, we can define an action of $\mathbf{GL}(V)$ on the set of lines in V, i.e., $\mathbb{P}(V) = \mathbb{P}^{n-1}(k) = \{$ lines in $V\}$. So, in $\mathbb{P}^{n-1}(k)$ we have

$[x_1, \dots, x_n] \backsim \lambda[x_1, \dots, x_n]$ for $\lambda \in k^\times$, which is called a homogeneous coordinate on $\mathbb{P}^{n-1}(k)$.

Note that $\mathbf{GL}(V)$ acts transitively on $\mathbb{P}(V) = \mathbb{P}^{n-1}(k) = \{$ lines in $V\}$ and in $k^n/\{0\} \to \mathbb{P}^{n-1}(k)$, the fibers are orbits of $k^\times$ acting by scalar multiplications.

Example 1.2.4.1. *The isotropy subgroup of*

$$l = \begin{pmatrix} * \\ 0 \\ \vdots \\ 0 \end{pmatrix} = k \oplus 0 \oplus \cdots \oplus 0 \subset k^n \textit{ is } \left(\begin{array}{c|c} * & * \\ \hline 0 & \\ \vdots & * \\ 0 & \end{array}\right) \in \mathbf{GL}(n,k).$$

Example 1.2.4.2. *It is easy to see that the slope of a line through the origin in $\mathbb{R}^2$ except a y-axis parameterize an element of $\mathbb{P}(\mathbb{R})$. Moreover, in the case of y-axis, by replacing $m = \frac{y}{x}$ by $m' = \frac{x}{y}$, we can parameterize all the elements of $\mathbb{P}(\mathbb{R})$ by slopes.*

By a homeomorphism ψ_j for $j = 1, \dots, n$,

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{A}_j = \{(x_1, \dots, x_n) \mid x_j \neq 0\} \xrightarrow{\psi_j} \begin{pmatrix} \frac{x_1}{x_j} \\ \vdots \\ \frac{x_{j-1}}{x_j} \\ \frac{x_{j+1}}{x_j} \\ \vdots \\ \frac{x_n}{x_j} \end{pmatrix} \in k^{n-1} \subset \mathbb{P}^{n-1}(k) \xrightarrow{\psi_j^{-1}} \begin{pmatrix} x_1 \\ \vdots \\ x_{j-1} \\ 1 \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix},$$

we deduce that $\mathbb{P}^{n-1}(k)$ can be covered by n coordinate-patches, which of each is an affine space k^{n-1}.

Let $\begin{bmatrix} z \\ w \end{bmatrix}$ be homogeneous coordinates of $\mathbb{CP}^1$. By letting

$$\begin{cases} \text{if } w \neq 0, & \frac{z}{w} \in \mathbb{C} \\ \text{if } w = 0, & \frac{z}{w} \overset{def}{=} \infty \end{cases}, \text{ we have } \mathbb{CP}^1 = \mathbb{C} \cup \{\infty\}.$$

In general, $\mathbb{P}^1(k) = k \cup \{\infty\}$. In this description, $\mathbf{GL}(2,\mathbb{C})$ acts on $\mathbb{CP}^1$ in the following ways:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} = \begin{bmatrix} az+b \\ cz+d \end{bmatrix} \backsim \begin{bmatrix} \frac{az+b}{cz+d} \\ 1 \end{bmatrix} \text{ or } \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ depending on whether or not } cz+d=0.$$

Clearly, this action is not effective, i.e.,

$$\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} = \begin{bmatrix} z \\ 1 \end{bmatrix} \text{ for } \lambda \in \mathbb{C}^\times.$$

So, we make this action effective by making a quotient;

$$\frac{\mathbf{GL}(2,\mathbb{C})}{\{\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mid \lambda \in \mathbb{C}^\times\}} \overset{def}{=} \mathbf{PGL}(2,\mathbb{C}).$$

In general,

$$\frac{\mathbf{GL}(n,k)}{\{\lambda \mathbf{1}_n \mid \mathbf{1}_n \text{ is an } n\times n \text{ identity matrix and } \lambda \in k^\times\}} \overset{def}{=} \mathbf{PGL}(n,k).$$

Remark 1.2.4.1.

$$\frac{\mathbf{SL}(2,\mathbb{R})}{\{\pm \mathbf{1}_2\}} = \mathbf{PSL}(2,\mathbb{C}) \text{ and } \frac{\mathbf{O}(n)}{\{\lambda \mathbf{1}_n \mid \lambda \in \mathbb{R}^\times\}} = \mathbf{PO}(n).$$

If $G \subset \mathbf{GL}(n,\mathbb{R})$, then

$$\mathbf{P}(G) = \frac{G}{G \cap \{\lambda \mathbf{1}_n \mid \lambda \in \mathbb{R}^\times\}}.$$

Remark 1.2.4.2. *The universal covering group of $\mathbf{PSL}(2,\mathbb{R})$ is a non-linear group. So, it can not be represented by matrixes.*
A Heisenberg group H is a nilpotent Lie group and $H/\mathbb{Z}$ is a nonlinear group.

1.2.5 Fibrations

Definition 1.2.5.1 (Fiber bundle). *We say $M \xrightarrow{\pi} N$ is a fiber bundle over N with fiber F if for each $x \in N$, there exists a neighborhood $U_x \subset N$ and a map $\pi^{-1}(U_x) \xrightarrow{\psi} F$ such that $\pi \times \psi$ is a diffeomorphism, i.e.*

$$\begin{array}{ccccc} U \times F & \underset{\pi\times\psi}{\xrightarrow{\cong}} & \pi^{-1}(U_x) & \hookrightarrow & M \\ & & {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ & & U_x & \longrightarrow & N \end{array}$$

Definition 1.2.5.2 (Covering homotopy property). *Let $p_1(y) = (y,0)$ and $I = [0,1]$. We say a map $\pi : M \to N$ satisfies the covering homotopy property if given $f : Y \to M$ and $f_t : Y \times I \to N$ satisfying $f_0 = \pi \circ \widetilde{f}_0 : Y \times I \to N$, there exists a homotopy $\widetilde{f}_t$ such that $f_t = \pi \circ \widetilde{f}_t$.*

$$\begin{array}{ccc} Y & \xrightarrow{f} & M \\ {\scriptstyle p_1}\downarrow & \overset{\widetilde{f}_t}{\nearrow} & \downarrow{\scriptstyle\pi} \\ Y \times I & \xrightarrow{f_t} & N \end{array}$$

Definition 1.2.5.3 (Fibration). *We say a map $\pi : M \to N$ is a fibration if it satisfies the covering homotopy property.*

Note that π is a trivial fibration if there exists $\psi : M \to F$ such that $\pi \times \psi : M \to N \times F$ is a diffeomorphism.

Remark 1.2.5.1. *Note that fiber bundles over Hausdorff and paracompact spaces are fibrations. See "Algebraic Topology" by Spanier. Since our base spaces are manifolds, from now on we can interchange two terminologies freely.*

Given a map $\pi : M \to N$ and $f : A \to N$, by defining a space

$$f^*M = \{(a, m) \in A \times M \mid f(a) = \pi(m)\},$$

we have a commutative diagram: For $p_1(a, m) = a$ and $p_2(a, m) = m$,

$$\begin{array}{ccc} f^*M & \xrightarrow{p_2} & M \\ {\scriptstyle p_1}\downarrow & & \downarrow{\scriptstyle \pi} \\ A & \xrightarrow{f} & N \end{array}$$

Now suppose that $\pi : M \to N$ is a fibration and we are given the following diagram: For $g_0(y) = p_1 \circ \widetilde{g_0}(y)$ where $\widetilde{g_0} : Y \times \{0\} \to A$,

$$\begin{array}{ccccc} Y & \xrightarrow{g} & f^*M & \xrightarrow{p_2} & M \\ {\scriptstyle p_1'}\downarrow & & {\scriptstyle p_1}\downarrow & & \downarrow{\scriptstyle \pi} \\ Y \times I & \xrightarrow{g_t} & A & \xrightarrow{f} & N \end{array}$$

Since $\pi : M \to N$ is a fibration and $f \circ g_0(y) = \pi \circ \widetilde{f \circ g_0}(y)$ where $\widetilde{g_0} : Y \times \{0\} \to N$, we have a homotopy $\widetilde{f \circ g_t}$

$$\begin{array}{ccc} Y & \xrightarrow{p_2 \circ g} & M \\ {\scriptstyle p_1'}\downarrow & \overset{\widetilde{f \circ g_t}}{\nearrow} & \downarrow{\scriptstyle \pi} \\ Y \times I & \xrightarrow{f \circ g_t} & N \end{array}$$

Since $f \circ g_t(y) = \pi \circ \widetilde{f \circ g_t}(y)$ by the definition $f^*M = \{(a, m) \in A \times M \mid f(a) = \pi(m)\}$, we can have a homotopy

$$g_t : Y \times I \to f^*M \text{ by } \widetilde{g_t}(x) = (g_t(y), \widetilde{f \circ g_t}(y)).$$

That is, we have a **pullback or induced** fibration $p_1 : f^*M \to A$, i.e.,

$$\begin{array}{ccc} Y & \xrightarrow{g} & f^*M \\ {\scriptstyle p_1'}\downarrow & \overset{\widetilde{g_t}}{\nearrow} & \downarrow{\scriptstyle p_1} \\ Y \times I & \xrightarrow{g_t} & A \end{array}$$

Example 1.2.5.1. *Suppose that $\pi : M \to N$ is a fibration. Given $F : A\times[1,2] \to N$, we have F^*M is a fibration over $A \times [1,2]$. Since there are the canonical inclusions $i_a : \{a\} \times [1,2] \to A \times [1,2]$ for $a \in A$, we also deduce that $(i_a^* \circ F^*)M = (i_a \circ F)^*M$ are fibrations over $\{a\} \times [1,2]$.*

One of useful consequences in fibrations is the following:

Theorem 1.2.5.1. *There is a bundle isomorphism between a fiber bundle over any contractible space and a trivial bundle.*

Proof. See "The Topology of Fiber Bundles" by Steenrod. □

An interesting fact about fibration is that some smooth map between two manifolds is a local fibration:

Theorem 1.2.5.2 (Local submersion theorem). *Let $f : M^m \to N^n$ be a smooth map for $m \geq n$. If $df_x : T_xM \to T_{f(x)}N$ is surjective where $x \in M$, then there exists a neighborhood U of $x \in M$ such that $f : U \to f(U)$ is a fibration.*

We know that $\mathbf{GL}(2,\mathbb{C})$ acts on $\mathbb{CP}^1$ by linear fractional transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} = \begin{bmatrix} az+b \\ cz+d \end{bmatrix} \backsim \begin{bmatrix} \frac{az+b}{cz+d} \\ 1 \end{bmatrix} \text{ or } \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ depending on whether or not } cz+d=0.$$

So does $\mathbf{SL}(2,\mathbb{R})$. Let $U = \{z \in \mathbb{C} \mid \Im(z) > 0\}$. For $x+iy \in U$, we have

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e^{\frac{1}{2}\log y} & 0 \\ 0 & e^{-\frac{1}{2}\log y} \end{pmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{x+iy}{\sqrt{y}} \\ \frac{1}{\sqrt{y}} \end{bmatrix}.$$

That is, $\mathbf{SL}(2,\mathbb{R})$ acts transitively on $U = \{z \in \mathbb{C} \mid \Im(z) > 0\}$. Now, we show that the isotropy subgroup of i is $\mathbf{SO}(2)$:

Exercise 1.2.5.1. *Let $A \in \mathbf{SL}(2,\mathbb{R})$.*

Show that $A(i) = i$ if and only if $A \in \mathbf{SO}(2) = \{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mid \theta \in \mathbb{R}\} \cong S^1$.

Proof.

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \cos\theta & \sin\theta \end{pmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i(i\sin\theta + \cos\theta) \\ i\sin\theta + \cos\theta \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Suppose that for $\lambda \in \mathbb{C}^\times$,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i\lambda \\ \lambda \end{bmatrix}.$$

We have $a = d$ and $b = -c$. Since $ad - bc = 1$, we conclude that $A \in \mathbf{SO}(2)$. □

Finally, we conclude that $\mathbf{SL}(2,\mathbb{R})$ acts transitively on $U = \{z \in \mathbb{C} \mid \Im(z) > 0\}$ with isotopy subgroup $\mathbf{SO}(2)$. So, $\mathbf{SL}(2,\mathbb{R})/\mathbf{SO}(2) \cong U \cong \mathbb{R}^2$ is a homogeneous space of $\mathbf{SL}(2,\mathbb{R})$. In this description, we have:

Example 1.2.5.2.

$$\mathbf{SL}(2,\mathbb{R}) \xrightarrow[\cong]{diffeo} \mathbb{R}^2 \times S^1.$$

Proof. Since $\mathbf{SL}(2,\mathbb{R}) \to \mathbf{SL}(2,\mathbb{R})/\mathbf{SO}(2)$ is smooth and everywhere submersion, by the local submersion theorem, we conclude that the following diagram is a fibration with a fiber $\mathbf{SO}(2)$:

$$\begin{array}{ccc} \mathbf{SO}(2) & \hookrightarrow & \mathbf{SL}(2,\mathbb{R}) \\ & & \downarrow \\ & & \mathbf{SL}(2,\mathbb{R})/\mathbf{SO}(2). \end{array}$$

Since $\mathbf{SL}(2,\mathbb{R})/\mathbf{SO}(2) \cong U \cong \mathbb{R}^2$ is contractible, by Theorem 1.2.5.1, $\mathbf{SL}(2,\mathbb{R})$ is diffeomorphic to $U \times \mathbf{SO}(2) \cong \mathbb{R}^2 \times S^1$. □

Now, we show a covering space of $\mathbf{SL}(2,\mathbb{R})$:

Example 1.2.5.3. $(\mathbb{R}^2/\{0\}) \times \mathbb{R}^1$ *is a trivial covering space of* $\mathbf{SL}(2,\mathbb{R})$*.*

Proof. Note that $\mathbf{SL}(2,\mathbb{R})$ acts transitively on $\mathbb{R}^2/\{0\}$, i.e., for $(x,y) \in \mathbb{R}^2/\{0\}$,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & b \\ y & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

It is easy to see that the isotropy subgroup of $(1,0)$ is

$$\{\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \mid r \in \mathbb{R}\} \cong \mathbb{R}.$$

Since $\mathbf{SL}(2,\mathbb{R}) \to \mathbf{SL}(2,\mathbb{R})/\{\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \mid r \in \mathbb{R}\} \cong \mathbb{R}^2/\{0\}$ is smooth and everywhere submersion, by the local submersion theorem, we conclude that the following diagram is a fibration with a fiber $\{\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \mid r \in \mathbb{R}\}$:

$$\begin{array}{ccc} \mathbb{R} \cong \{\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \mid r \in \mathbb{R}\} & \hookrightarrow & \mathbf{SL}(2,\mathbb{R}) \\ & & \downarrow \\ & & \mathbf{SL}(2,\mathbb{R})/\{\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \mid r \in \mathbb{R}\} \cong \mathbb{R}^2/\{0\}. \end{array}$$

Here, unfortunately, $\mathbf{SL}(2,\mathbb{R})/\{\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \mid r \in \mathbb{R}\} \cong \mathbb{R}^2/\{0\}$ is not contractible. So, **we can not use Theorem 1.2.5.1. However, what we do know is that** $\mathbb{R}^2/\{0\}$ **is locally contractible. So, we conclude that** $\mathbf{SL}(2,\mathbb{R})$ is **locally** diffeomorphic to $\mathbb{R}^2/\{0\} \times \mathbb{R}^1$. That is, $(\mathbb{R}^2/\{0\}) \times \mathbb{R}^1$ is a covering space of $\mathbf{SL}(2,\mathbb{R})$, which is trivial (=connected). □

Now, we show a nontrivial covering space of $\mathbf{SL}(2,\mathbb{R})$:

Example 1.2.5.4. $\mathbb{RP}^1 \times \mathbb{R}^2 \times \{1,2\}$ *is a nontrivial covering space of* $\mathbf{SL}(2,\mathbb{R})$*.*

Proof. Since $\mathbf{SO}(2) \subset \mathbf{SL}(2,\mathbb{R})$, $\mathbf{SL}(2,\mathbb{R})$ acts transitively on $\mathbb{RP}^1$. Using homogeneous coordinates of $\mathbb{RP}^1$, it is easy to see that the isotropy subgroup of $(1,0)$ is

$$\{\begin{pmatrix} a & r \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{R}^\times \text{ and } r \in \mathbb{R}\} \cong \mathbb{R}^\times \times \mathbb{R} \cong \mathbb{R} \times \{1,2\} \times \mathbb{R}.$$

Since $\mathbf{SL}(2,\mathbb{R}) \to \mathbf{SL}(2,\mathbb{R})/\{\begin{pmatrix} a & r \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{R}^\times \text{ and } r \in \mathbb{R}\} \cong \mathbb{RP}^1$ is smooth and everywhere submersion, by the local submersion theorem, we conclude that the following diagram is a fibration with a fiber $\{\begin{pmatrix} a & r \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{R}^\times \text{ and } r \in \mathbb{R}\} \cong \mathbb{R}^\times \times \mathbb{R} \cong \mathbb{R} \times \{1,2\} \times \mathbb{R}$:

$$\begin{array}{ccc} \mathbb{R}^2 \times \{1,2\} \cong \{\begin{pmatrix} a & r \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{R}^\times \text{ and } r \in \mathbb{R}\} & \hookrightarrow & \mathbf{SL}(2,\mathbb{R}) \\ & & \downarrow \\ & & \mathbb{RP}^1. \end{array}$$

Again, since $\mathbf{SL}(2,\mathbb{R})/\{\begin{pmatrix} a & r \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{R}^\times \text{ and } r \in \mathbb{R}\} \cong \mathbb{RP}^1$ is locally contractible, $\mathbf{SL}(2,\mathbb{R})$ is **locally** diffeomorphic to $\mathbb{R}^2 \times \{1,2\} \times \mathbb{PR}^1$. That is, $\mathbb{R}^2 \times \{1,2\} \times \mathbb{RR}^1$ is a nontrivial (=non connected) covering space of $\mathbf{SL}(2,\mathbb{R})$. □

Exercise 1.2.5.2. *Express the Möbius band $\mathcal{M}$ as a homogeneous space of* $\mathbf{SO}(2,1)$.

Proof. It is a well-known fact that the Möbius band is homeomorphic to the punctured projective plane. Think $\mathbb{RP}^2$ as a disk with an identification with diametrically opposite points of the boundary. So, the $\mathbb{RP}^2$ without a small closed disk at the origin is the Möbius band. If we give $\mathbb{RP}^2$ homogeneous coordinates from $\mathbb{R}^3$, then we can give homogeneous coordinates each point of the Möbius band. So, it is intuitively clear that the Möbius band is given by with respect to homogeneous coordinates $[x_0, x_1, x_2]$:

$$\mathcal{M} = \{v = [x_0, x_1, x_2] \mid v^t \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} v < 0\}.$$

Since

$$\mathbf{SO}(2,1) = \{A \in \mathbf{M}_2(\mathbb{R}) \mid A^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ and } \det A = 1\},$$

we have a well-defined action of $\mathbf{SO}(2,1)$ on $\mathcal{M}$. Actually, the inner product given by

$$w \circ v = w^t \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} v \text{ is called the Lorentzian inner product.}$$

If $v \in \mathbb{R}^3$, using the usual linear algebraic manipulations, we can find a Lorentzian orthonormal basis $\{v, w_1, w_2\}$ with respect to the Lorentzian inner product. Since $(0,0,1), (1,0,0), (0,1,0)$ is Lorentzian orthonormal, regarding v, w_1, w_2 as $n \times 1$ column vectors, i.e., $(w_1, w_2, v) \in \mathbf{O}(2,1)$, we have

$$v = (w_1, w_2, v) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

That is, $\mathbf{O}(2,1)$ acts transitively on $\mathbb{R}^3$. Since we are using homogeneous coordinates, that action of $\mathbf{SO}(2,1)$ is thee same as the action of $\mathbf{O}(2,1)$. so, the transitivity of $\mathbf{O}(2,1)$ and $[0,0,1] \in \mathcal{M}$ implies the transitivity of $\mathbf{SO}(2,1)$ on $\mathcal{M}$. So, we conclude that Möbius band $\mathcal{M}$ is a homogeneous space of $\mathbf{SO}(2,1)$. □

1.2.6 Projective transformations as isometries

Example 1.2.6.1.

$$\mathbf{PSL}(2,\mathbb{R}) \xrightarrow[\cong]{diffeo} \mathbb{R}^2 \times S^1.$$

Proof. It is easy to see that $\mathbf{PSL}(2,\mathbb{R})$ acts transitively on $U = \{z \in \mathbb{C} \mid \Im(z) > 0\}$ with isotopy subgroup $\mathbf{PSO}(2)$. So, $\mathbf{PSL}(2,\mathbb{R})/\mathbf{PSO}(2) \cong U \cong \mathbb{R}^2$ is a homogeneous space of $\mathbf{PSL}(2,\mathbb{R})$. By the same argument in the proof of Exercise 1.2.5.2, we have

$$\mathbf{PSL}(2,\mathbb{R}) \xrightarrow[\cong]{diffeo} \mathbb{R}^2 \times S^1.$$

□

Recall that the upper half plane $\mathcal{U} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$ can be equipped with a Riemannian metric

$$g = \frac{|dz|}{\Im(z)}.$$

The following definition is due to É. Cartan:

Definition 1.2.6.1 (symmetric spaces). *We say M is a symmetric space if for all $p \in M$ there exists an isometry $s_p : M \to M$ such that*

$$s_p(p) = p \text{ and } (ds_p)_p = -id.$$

Example 1.2.6.2. $\mathcal{U} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$ *with* $g = \frac{|dz|}{\Im(z)}$ *is a symmetric space.*

Proof. First of all, note that $g = \frac{|dz|}{\Im(z)}$ implies that if $(ds_p)_p = -id$, then s_p is necessarily an isometry. Let $s_i(z) = -\frac{1}{z}$. So, $(ds_i)_i = \frac{1}{z^2}|_{z=i} = -1$. Hence, we have

$$s_i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } (ds_i)_i = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

So, we have $s_i(i) = i$. Since we know that $\mathbf{PSL}(2,\mathbb{R})$ acts transitively on $\mathcal{U}$, there is $A \in \mathbf{PSL}(2,\mathbb{R})$ such that $A(i) = p$ for a given $p \in \mathcal{U}$. So, letting $s_p = A \cdot s_i \cdot A^{-1}$, we have

$$s_p(p) = A \cdot s_i \cdot A^{-1}(p) = A \cdot s_i(i) = A(i) = p.$$

Now, making an identification $A \longleftrightarrow f(z) = \frac{az+b}{cz+d}$, we have $s_p(z) = (f \circ (-\frac{1}{z}) \circ f^{-1})(z) = f(-\frac{1}{f^{-1}(z)})$. From $s_p(p) = p$, we have

$$f^{-1}(p) = -\frac{1}{f^{-1}(p)}.$$

Since $f'(f^{-1}(z)) \cdot (f^{-1})'(z) = 1$, we have

$$(ds_p)_p = f'(-\frac{1}{f^{-1}(p)}) \cdot \frac{(f^{-1})'(p)}{(f^{-1}(p))^2} = f'(f^{-1}(p)) \cdot \frac{(f^{-1})'(p)}{(f^{-1}(p))^2} = \frac{1}{i^2} = -1.$$

□

Remark 1.2.6.1. *If X is a positively (or negatively) curved symmetric space, the dual space X^* is a negatively (or positively) curved symmetric space.*

Recall that H is Hermitian if $H = \overline{H}^T$.

Definition 1.2.6.2. *$\mathbb{C}^{1,1}$ is a $\mathbb{C}$-vector space $\mathbb{C}^2$ with a Hermitian form defined by*

$$\langle z, w\rangle = z_1\overline{w_1} - z_2\overline{w_2} = \overline{w}^T Hz \text{ for } z = (z_1, z_2), w = (w_1, w_2) \in \mathbb{C}^2 \text{ and } H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Let

$$\begin{aligned} \mathbf{U}(H) &\overset{def}{=} \{A \in \mathbf{GL}(2,\mathbb{C}) \mid \langle Av, Aw\rangle = \langle v, w\rangle \text{ for all } v, w \in \mathbb{C}^{1,1}\} \\ &= \{A \in \mathbf{GL}(2,\mathbb{C}) \mid \overline{A}^T HA = H\}. \end{aligned}$$

Note that in the notation above Exercise 1.2.2.1, we have $\mathbf{U}(1,1) = \mathbf{U}(H)$.

Clearly, $\mathbf{U}(1,1)$ acts on

$$X = \{v \in \mathbb{C}^{1,1} \mid \langle v, v\rangle < 0\} = \{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 - |z_2|^2 < 0\}.$$

This implies the following:

Example 1.2.6.3. $\mathbf{U}(1,1)$ *preserves a unit disk $D_{\mathbb{C}} = \{z \in \mathbb{C} \mid |z| < 1\}$.*

Proof. Notice that $|z_1|^2 - |z_2|^2 < 0$ implies that $(z_1, 0) \notin X$. Since

$$\frac{|z_1|^2}{|z_2|^2} = |z|^2 < 1 \Longleftrightarrow |z_1|^2 - |z_2|^2 < 0,$$

we have a well-defined holomorphic map f from X to $D_{\mathbb{C}}$ by $f(z_1, z_2) = \frac{z_1}{z_2}$. So, we can define an action of $A \in \mathbf{U}(1,1)$ on $z \in D_{\mathbb{C}}$ by

$$A \cdot z = f(A \begin{pmatrix} z \\ 1 \end{pmatrix}).$$

□

Note that from $\overline{A}^T HA = H$, we deduce that $\mathbf{U}(1,1)$ is a zero locus in $\mathbb{R}^8$ by four transversal equations. So, $\dim_{\mathbb{R}} \mathbf{U}(1,1) = 4$. Also, notice that

$$\begin{array}{ccccc} \mathbf{U}(1) & \hookrightarrow & \mathbf{U}(1) \times \mathbf{U}(1) & \hookrightarrow & \mathbf{U}(1,1) \\ e^{i\theta} & \to & (e^{i\theta}, e^{i\theta}) & & \\ & & (e^{i\theta}, e^{i\psi}) & \to & \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\psi} \end{pmatrix} \end{array}.$$

Example 1.2.6.4. $\mathbf{PSU}(1,1) \overset{def}{=} \mathbf{SU}(1,1)/\{\pm \mathbf{1}_2\} \cong \mathbf{Aut}(D_{\mathbb{C}})$.

Proof. We are going to show that $\mathbf{PSU}(1,1)$ acts effectively on $D_{\mathbb{C}}$, i.e., $\mathbf{PSU}(1,1) \to \mathbf{Aut}(D_{\mathbb{C}})$ is injective. For surjectivity, we need the Schwarz's lemma in complex analysis. So, we leave it to the readers.

Since $\det A \neq 0$ for $A \in \mathbf{U}(1,1)$, A is an automorphism of X. Also, it is easy to see that $Av = v$ for all $v \in X$ implies that $A = \mathbf{I}_2$. $\mathbf{U}(1,1)$ acts effectively on X. If we closely look at the identification in Example 1.2.6.3, we can see that A and $-A$ define the same action on $D_{\mathbb{C}}$. So, we conclude that $\mathbf{PU}(1,1)$ acts effectively on $D_{\mathbb{C}}$. Hence, $\mathbf{PSU}(1,1)$ acts effectively on $D_{\mathbb{C}}$. ☐

Note that we showed that we can identify $D_{\mathbb{C}}$ as

$$X = \{v = (z_1, z_2) \in \mathbb{C}^{1,1} \mid \overline{v}^T \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} v < 0\} = \{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 - |z_2|^2 < 0\}.$$

Let $z = \frac{z_1}{z_2}$. Note that $|z_1|^2 - |z_2|^2 < 0$ implies that $z_2 \neq 0$. So,

$$\begin{aligned}\Im(z) &= \frac{1}{2i}(z - \overline{z}) = \frac{1}{2i}(\frac{z_1}{z_2} - \frac{\overline{z_1}}{\overline{z_2}}) \\ &= \frac{1}{2i|z_2|^2}(z_1\overline{z_2} - z_2\overline{z_1}) = \frac{1}{|z_2|^2}(iz_2\overline{z_1} - iz_1\overline{z_2}) \\ &= \frac{1}{|z_2|^2}(\overline{z_1}, \overline{z_2}) \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.\end{aligned}$$

Since $\frac{1}{|z_2|^2} > 0$, we have

$$\Im(z) > 0 \text{ if and only if } \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \text{ defines a positive definite Hermitian inner product.}$$

That is, we can give the upper half plane a Hermitian metric $\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$. In this description, an interesting thing is the following: Since $\mathbf{SU}(1,1)$ is the group of isometries of $D_{\mathbb{C}}$ with the inner product $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Let $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$, which is called the Cayley transformation. It is easy to see that

$$\overline{C}^T \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

So, we have $C \cdot \mathbf{SU}(1,1)$ is the group of isometries of the upper half plane with the inner product $\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$. By the similar proof of Exercise 1.2.2.1, We also have that

$$C^{-1} \cdot \mathbf{SL}(2,\mathbb{R}) \cdot C = \mathbf{SU}(1,1).$$

Note that in general, $\mathbf{SU}(p,q) \ncong \mathbf{SL}(p+q,\mathbb{R})$. We will give three famous metrics:

Remark 1.2.6.2.
$$\frac{2|dz|}{1-|z|^2} \text{ on } D_{\mathbb{C}}$$
$$\frac{|dz|}{\Im(z)} \text{ on } \mathcal{U} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$$
$$\frac{2|dz|}{1+|z|^2} \text{ on } \mathbb{CP}^1.$$

Note that $\frac{2|dz|}{1+|z|^2}$ on $\mathbb{CP}^1$ is called the Fubini-study metric. Since it is invariant under inversions, it is well-defined, i.e., it is even defined at ∞: Let $w = z^{-1}$. So, we have

$$dw = -z^{-2}dz \Longrightarrow |dw| = |z|^{-2}|dz|.$$

So, we have

$$\frac{2|dw|}{1+|w|^2} = \frac{2|z|^{-2}|dz|}{1+|z|^{-2}} = \frac{2|dz|}{1+|z|^2}.$$

Hence, we can assign values at 0 or ∞ using each coordinate chart.

1.2.7 Decomposition of $\mathbf{SL}(2,\mathbb{R})$

Note that $\mathbf{SL}(2,\mathbb{R})$ is a prototype of a noncompact real semisimple Lie group.

Definition 1.2.7.1 (Class function). *Let G be a group and F an arbitrary field. A class function f is a function from G to F which is constant on the conjugacy classes of G. That is,*

$$G \xrightarrow{f} F \text{ such that } f(g^{-1}xg) = f(x) \text{ for all } g, x \in G.$$

Example 1.2.7.1. *Let G be a matrix group. For $A \in G$, define $G \xrightarrow{f} \mathbb{R}$ by $f(A) = \mathbf{tr}(A)$. f is a class function.*

Example 1.2.7.2. *Let $\mathbf{SL}(2,\mathbb{R}) \xrightarrow{f} \mathbb{R}$ by $f(A) = \mathbf{tr}(A)$. Then $f^{-1}(0) = \mathfrak{sl}(2,\mathbb{R}) \cap \mathbf{SL}(2,\mathbb{R})$.*

Exercise 1.2.7.1. *Let D^2 be an open unit disk in $\mathbb{R}^2$. Define a class function f of a solid torus $D^2 \times S^1$ as a Lie group.*

Proof. Let (r, α, θ) be a coordinate system of $D^2 \times S^1$ where $0 \le r < 1, 0 \le \alpha < 2\pi$, and $0 \le \theta < 2\pi$. Define $D^2 \times S^1 \xrightarrow{\varphi} \mathbf{SL}(2,\mathbb{R})$ by

$$\varphi(r,\alpha,\theta) = \begin{pmatrix} e^{\frac{1}{2}\log(\frac{1-r^2}{r^2+1-2r\cos\alpha})} & (\frac{2r\sin\alpha}{r^2+1-2r\cos\alpha})e^{\frac{1}{2}\log(\frac{1-r^2}{r^2+1-2r\cos\alpha})} \\ 0 & e^{-\frac{1}{2}\log(\frac{1-r^2}{r^2+1-2r\cos\alpha})} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

It is easily seen to be a diffeomorphism and a Lie group homomorphism. So, let

$$f(r,\alpha,\theta) = (\cos\theta)\big(e^{\frac{1}{2}\log(\frac{1-r^2}{r^2+1-2r\cos\alpha})} + e^{-\frac{1}{2}\log(\frac{1-r^2}{r^2+1-2r\cos\alpha})}\big) + (\sin\theta)(\frac{2r\sin\alpha}{r^2+1-2r\cos\alpha})e^{\frac{1}{2}\log(\frac{1-r^2}{r^2+1-2r\cos\alpha})}.$$

f is a class function on a solid torus $D^2 \times S^1$. □

If we look at the proof of Exercise 1.2.7.1, we actually decompose $\mathbf{SL}(2,\mathbb{R})$ or a solid torus $D^2 \times S^1$ into a product of an element of $\mathbf{SO}(2)$ and an element of upper triangular matrix. It is no coincidences. In general, we have

Theorem 1.2.7.1 (Iwasawa Decomposition). *Let G be a Lie group. We can decompose*

$$G = K \cdot A \cdot N \text{ where } K \text{ is a compact group, } A \text{ is an abelian group, and } N \text{ is a nilpotent group.}$$

Proof. We will prove this later. □

Example 1.2.7.3. *Suppose that $A \in \mathbf{SL}(2,\mathbb{R})$. we have*

$$A = K \cdot \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \text{ where } K \in \mathbf{SO}(2) \text{ and } r, t \in \mathbb{R}.$$

However, in the case of $\mathbf{GL}(n,\mathbb{R})$, without using the Iwasawa decomposition, we still have a nice decomposition by the virtue of Gram-Schmidt orthogonalization.

Proof. Let $A = (A_1, \ldots, A_n) \in \mathbf{GL}(n,\mathbb{R})$ where A_i is an $n \times 1$-column matrix for $i = 1, \ldots, n$. The Gram-Schmidt orthonormal process say that we can have $A' = (A'_1, \ldots, A'_n)$ where A'_i is orthonormal to each other:

$$\begin{aligned}
A'_1 =& A_1 \\
A'_2 =& A_2 - \frac{\langle A_2, A'_1 \rangle}{\langle A'_1, A'_1 \rangle} A'_1 \\
A'_3 =& A_3 - \frac{\langle A_3, A'_2 \rangle}{\langle A'_2, A'_2 \rangle} A'_2 - \frac{\langle A_3, A'_1 \rangle}{\langle A'_1, A'_1 \rangle} A'_1 \\
&\vdots \\
A'_n =& A_n - \frac{\langle A_n, A'_{n-1} \rangle}{\langle A'_{n-1}, A'_{n-1} \rangle} A'_{n-1} - \cdots - \frac{\langle A_n, A'_1 \rangle}{\langle A'_1, A'_1 \rangle} A'_1.
\end{aligned}$$

$$A = (A'_1, \ldots, A'_n) \cdot \begin{pmatrix}
1 & \frac{\langle A_2, A'_1 \rangle}{\langle A'_1, A'_1 \rangle} & \frac{\langle A_3, A'_1 \rangle}{\langle A'_1, A'_1 \rangle} & \frac{\langle A_4, A'_1 \rangle}{\langle A'_1, A'_1 \rangle} & \cdots & \frac{\langle A_n, A'_1 \rangle}{\langle A'_1, A'_1 \rangle} \\
0 & 1 & \frac{\langle A_3, A'_2 \rangle}{\langle A'_2, A'_2 \rangle} & \frac{\langle A_4, A'_2 \rangle}{\langle A'_2, A'_2 \rangle} & \cdots & \frac{\langle A_n, A'_2 \rangle}{\langle A'_2, A'_2 \rangle} \\
0 & 0 & 1 & \frac{\langle A_4, A'_3 \rangle}{\langle A'_3, A'_3 \rangle} & \cdots & \frac{\langle A_n, A'_3 \rangle}{\langle A'_3, A'_3 \rangle} \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 1 & \frac{\langle A_n, A'_{n-1} \rangle}{\langle A'_{n-1}, A'_{n-1} \rangle} \\
0 & 0 & \cdots & 0 & 0 & 1
\end{pmatrix}$$

Since A'_i is orthonormal to each other, $(A'_1, \ldots, A'_n) = A' \in \mathbf{O}(n)$, which is a compact group. So, we conclude that for $A = (A_1, \ldots, A_n) \in \mathbf{GL}(n,\mathbb{R})$ we have a decomposition

$$A = A' \cdot U \text{ where } U \text{ is an upper triangular and } A' \in \mathbf{O}(n).$$

□

Consider an action of $\mathbf{SO}(2)$ on $\mathbb{RP}^1 = \partial\mathcal{U} = \mathbb{R}\cup\{\infty\}$ where $\mathcal{U}$ is the upper half plane:

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}.$$

By identifying

$$\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} \longleftrightarrow \cot\theta,$$

we have a transitive action of $\mathbf{SO}(2)$ on $\mathbb{RP}^1$. So, $\mathbf{SL}(2,\mathbb{R})$ acts transitively on $\mathbb{RP}^1$. Let $\mathcal{B}$ be the isotopy subgroup of $\begin{pmatrix}1\\0\end{pmatrix}$ in $\mathbf{SL}(2,\mathbb{R})$, which called a Borel subgroup. Since we are using a homogeneous coordinate in $\mathbb{RP}^1$, it is obvious that $\mathcal{B}$ consists of upper triangular matrixes. In this description, combining the Gram-Schmidt process, we have

Example 1.2.7.4. *Let $A\in\mathbf{SL}(2,\mathbb{R})$. Then $A = A^{orth}\cdot A^{upp}$. That is,*

$$\mathbf{SL}(2,\mathbb{R}) = \mathbf{SO}(2)\times\mathcal{B}.$$

That is, for $A\in\mathbf{SL}(2,\mathbb{R})$, we have

$$A = A^{orth}\cdot\begin{pmatrix}e^t & re^t\\ 0 & e^{-t}\end{pmatrix} \text{ where } A^{orth}\in\mathbf{SO}(2) \text{ and } r,t\in\mathbb{R}.$$

Note that $\mathcal{B} = \{\begin{pmatrix}e^t & re^t\\ 0 & e^{-t}\end{pmatrix} \mid r,t\in\mathbb{R}\}$ is called a Borel subgroup. Moreover, from the above, it is easy to see that the action of $\mathbf{SL}(2,\mathbb{R})$ on $\mathbb{RP}^1$ is completely and uniquely up to $\pm\mathbf{1}_2$ determined by orthogonal matrixes in $\mathbf{SO}(2)$.
By identifying $\mathbb{RP}^1$ with the set of oriented lines in $\mathbb{R}^2$, we have an action of $\mathbf{SO}(2)$ on $\{$oriented lines in $\mathbb{R}^2\}$. Again, we have

$$\mathbf{SL}(2,\mathbb{R})/\mathbf{SO}(2)\cong\mathcal{B}\cong\mathbb{R}^2.$$

Note that $\mathbf{PSO}(2) = \mathbf{SO}(2)/\{\pm\mathbf{1}_2\}$ acts **simply transitively** on $\mathbb{RP}^1$ and $\mathbf{PSL}(2,\mathbb{R})$ acts transitively on $\mathbb{RP}^1$. Also, it is easily seen that $\mathbf{SO}(2)$ acts transitively a double covering space $(\mathbb{R}^2/\{0\})/\mathbb{R}_+$ of $\mathbb{RP}^1$.

Now, we are going to discuss a polar decomposition (Cartan decomposition) of $\mathbf{SL}(2,\mathbb{R})$. First, we state

Theorem 1.2.7.2. *Every connected Lie group G has a maximal compact Lie subgroup K such that $G\cong K\times\mathbb{R}^d$. Moreover, K is unique up to conjugation.*

Proof. We are going to prove this later. □

Example 1.2.7.5. *If G is $\mathbf{SL}(2,\mathbb{R})$, then $K=\mathbf{SO}(2)$.*

Example 1.2.7.6. *If G is $\mathbf{SL}(2,\mathbb{C})$, then $K=\mathbf{SU}(2)$. So, $G/K\cong\mathbb{R}^3$, which is a real hyperbolic 3- space.*

Also, we state

Theorem 1.2.7.3. *Let G be a Lie group and K be a maximal compact Lie subgroup of G. G is homotopy equivalent to K.*

Proof. We are going to prove this later. □

Example 1.2.7.7. $\mathbf{SL}(2,\mathbb{R})$ *is homotopy equivalent to* $\mathbf{SO}(2)$.

Proof. Since $\mathcal{B} = \{\begin{pmatrix} e^t & re^t \\ 0 & e^{-t} \end{pmatrix} \mid r, t \in \mathbb{R}\}$ is diffeomorphic to $\mathbb{R}^2$, $\mathcal{B}$ is contractible. Since

$$\mathbf{SL}(2,\mathbb{R}) = \mathbf{SO}(2) \times \mathcal{B},$$

we have $\mathbf{SL}(2,\mathbb{R})$ is homotopy equivalent to $\mathbf{SO}(2)$. □

Notation 1.2.7.1. *Let*

$$\begin{aligned} \mathcal{P}_n &= \{n \times n \text{ symmetric positive definite real matrices.}\} \\ \mathcal{S}_n &= \{A \in \mathbf{M}_n(\mathbb{R}) \mid A = A^T\}. \end{aligned}$$

Note that positive definiteness of $A \in \mathbf{M}_n(\mathbb{R})$ means $v^T A v > 0$ for $v \in \mathbb{R}^n$. Clearly, $\mathcal{P}_n \subseteq \mathcal{S}_n$. It is easily seen that $\dim \mathcal{S}_n = \frac{n(n+1)}{2}$. Note that $\mathcal{P}_n$ is a convex open set of $\mathcal{S}_n$ in the sense that if $A, B \in \mathcal{P}_n$, then $(1-t)A + tB \in \mathcal{P}_n$ for $0 \le t \le 1$. From linear algebra, we have

Theorem 1.2.7.4 (Spectral theorem). *If $X \in \mathcal{P}_n$, then there exists $B \in \mathbf{O}(n)$ such that BXB^{-1} is a diagonal matrix with positive diagonal entries $\lambda_i > 0$ for $i = 1, \ldots, n$, i.e.,*

$$BXB^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & 0 \\ 0 & \cdots & 0 & 0 & \lambda_n \end{pmatrix}$$

Define an action of $\mathbf{GL}(n,\mathbb{R})$ on $\mathcal{P}_n$ by

$$X \mapsto A^T X A \text{ by } X \in \mathcal{P}_n \text{ and } A \in \mathbf{GL}(n,\mathbb{R}).$$

Example 1.2.7.8.

$$\mathcal{P}_n = \mathbf{GL}(n,\mathbb{R})/\mathbf{O}(n).$$

Proof. We shall show that this action is well-defined transitive with the isotropy subgroup $\mathbf{O}(n)$. First, we show that it is well-defined: Note that

$$\begin{aligned} \mathcal{P}_n &= \{n \times n \text{ symmetric positive definite real matrices.}\} \\ &= \{X \in \mathbf{M}_n(\mathbb{R}) \mid X^T = X \text{ and } v^T X v > 0 \text{ for all } v \in \mathbb{R}^n\}. \end{aligned}$$

Let $X \in \mathcal{P}_n$ and $A \in \mathbf{GL}(n,\mathbb{R})$. The action is well-defined by the facts that

$$(A^T X A)^T = A^T X^T A = A^T X A \text{ and } v^T (A^T X A) v = (Av)^T X (Av) > 0.$$

The spectral theorem says that this action is transitive: Suppose that $X, Y \in \mathcal{P}_n$. Since $B^T = B^{-1}$ for $B \in \mathbf{O}(n) \subset \mathbf{GL}(n, \mathbb{R})$, without loss of generality, by the spectral theorem we can assume X is a diagonal matrix with positive diagonal entries $\lambda_i > 0$ for $i = 1, \ldots, n$ and Y is a diagonal matrix with positive diagonal entries $\eta_i > 0$ for $i = 1, \ldots, n$. Let

$$A = \begin{pmatrix} \sqrt{\frac{\eta_1}{\lambda_1}} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{\frac{\eta_2}{\lambda_2}} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sqrt{\frac{\eta_{n-1}}{\lambda_{n-1}}} & 0 \\ 0 & \cdots & 0 & 0 & \sqrt{\frac{\eta_n}{\lambda_n}} \end{pmatrix}$$

So, $A^T X A = Y$ for $A \in \mathbf{GL}(n, \mathbb{R})$. So, $\mathcal{P}_n$ is a homogeneous space of $\mathbf{GL}(n, \mathbb{R})$. Since $\mathbf{1}_n \in \mathcal{P}_n$ and $\mathbf{O}(n)$ is defined by $A^T A = A^T \mathbf{1}_n A = \mathbf{1}_n$, we conclude that $\mathbf{O}(n)$ is the isotropy subgroup. Hence,

$$\mathcal{P}_n = \mathbf{GL}(n, \mathbb{R}) / \mathbf{O}(n).$$

□

We have decomposed $\mathbf{GL}(n, \mathbb{R})$ into the product of $\mathcal{P}_n$ and $\mathbf{O}(n)$, which is a maximal compact subgroup. In general, we have

Theorem 1.2.7.5 (Cartan decomposition). *Let G be a Lie group. We can decompose*

$G = K \cdot \mathcal{P}$ where K is a compact group and $\mathcal{P}$ is a positive definite symmetric group.

Example 1.2.7.9.

$$\mathbf{SL}(2, \mathbb{R}) = \mathbf{SO}(2) \times (\mathcal{P}_2 \cap \mathbf{SL}(2, \mathbb{R})).$$

Note that If $A \in \mathcal{P}_2 \cap \mathbf{SL}(2, \mathbb{R})$, then we can have $A = B \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} B^{-1}$ where $B \in \mathbf{SO}(2)$

From the proof of Example 1.2.7.8, we can also prove

Exercise 1.2.7.2. *For all $X \in \mathcal{P}_n$, there exists a unique $Y \in \mathcal{P}_n$ such that $Y^2 = X$.*

Proof. If $X \in \mathcal{P}_n$, then there exists $B \in \mathbf{O}(n)$ such that BXB^T is a diagonal matrix with positive diagonal entries $\lambda_i > 0$ for $i = 1, \ldots, n$, i.e.,

$$BXB^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & 0 \\ 0 & \cdots & 0 & 0 & \lambda_n \end{pmatrix}$$

Let

$$A = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sqrt{\lambda_{n-1}} & 0 \\ 0 & \cdots & 0 & 0 & \sqrt{\lambda_n} \end{pmatrix}$$

So, we have $B^{-1}A^2B = X$. Since $B^T = B^{-1}$ for $B \in \mathbf{O}(n) \subset \mathbf{GL}(n, \mathbb{R})$ and

$$B^{-1}A^2B = (B^{-1}AB) \cdot (B^{-1}AB) = (B^{-1}AB)^2 = (B^TAB)^2,$$

we have $Y = B^TAB \in \mathcal{P}_n$. $\square$

The next exercise shall show that $\mathcal{S}_n$ is a Lie algebra of $\mathcal{P}_n$.

Exercise 1.2.7.3. *Show that* $\exp : \mathcal{S}_n \to \mathcal{P}_n$ *is injective.*

Proof. First, we show that $\exp : \mathcal{S}_n \to \mathcal{P}_n$ is well-defined. Let $A = (a_{ij})_{n\times n} \in \mathcal{S}_n$. In the case of $\mathcal{S}_n$, the spectral theorem says that if $A \in \mathcal{S}_n$, then there exists $B \in \mathbf{O}(n)$ such that BAB^{-1} is a diagonal matrix with nonzero diagonal entries $\lambda_i \neq 0$ for $i = 1, \ldots, n$, i.e.,

$$BAB^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & 0 \\ 0 & \cdots & 0 & 0 & \lambda_n \end{pmatrix}$$

So,

$$\begin{aligned} \exp A &= B^{-1}B\exp(A)B^{-1}B \\ &= B^{-1}\exp(BAB^{-1})B \\ &= B^{-1}\begin{pmatrix} e^{\lambda_1} & 0 & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & e^{\lambda_{n-1}} & 0 \\ 0 & \cdots & 0 & 0 & e^{\lambda_n} \end{pmatrix} B. \end{aligned}$$

Obviously, it is symmetric and since $e^x > 0$, we conclude that it is positive definite. Moreover, it is obvious that it is injective by the fact that e^x is an injective function. $\square$

Exercise 1.2.7.4. *Let G be a Lie group and H be a* ***connected topological space****. Suppose that $p : H \to G$ is a covering space. Choose an element $h_0 \in p^{-1}(e) \subset H$ to serve as the identity element in H where e is the identity in G,*
Show that there exists a unique Lie group structure on H such that p is a homomorphism of Lie group.

Proof. Note that the covering map p makes H a connected differentiable manifold (See Section 3.1.7.). Also, note that H is path-connected by the fact that H is a connected manifold. Hence, G is also path-connected. Let e be the identity of G and

$$\widetilde{H} = \{[\gamma] \mid \gamma; [0,1] \to G \text{ where } \gamma \text{ is continuous and } \gamma(0) = e.\}$$

Here, $[\gamma_1] = [\gamma_2]$ if $\gamma_1 \backsim \gamma_2$ rel $\{0, 1\}$. It is well-known that $\widetilde{H}$ is a universal covering space of G:

$$\widetilde{H} \xrightarrow{\pi} G \text{ where } \pi([\gamma]) = \gamma(1).$$

So, we also deduce that $\widetilde{H}$ is a differentiable manifold. Using the group structure of G, we can define a group structure on $\widetilde{H}$ in the following manners:

$$[\gamma_1] \cdot [\gamma_2] = [\gamma_1 \cdot \gamma_2].$$

Note that the identity element of $\widetilde{H}$ is $[\gamma_e]$, i.e., $\gamma_e([0,1]) = e$. Since $\widetilde{H}$ is also a universal covering space of H and H is a covering space of G, we have the following commutative diagram:

$$\begin{array}{ccc} \widetilde{H} & \xrightarrow{\pi} & G \\ {\scriptstyle f}\downarrow & \nearrow {\scriptstyle p} & \\ H & & \end{array}$$

Note that f is not unique. However, since we can choose f such that $f([\gamma_e]) = h_0$ and $\widetilde{H}/\Gamma \cong H$ where Γ is the group of deck transformations of f, we can give a unique Lie group structure on H such that p is a homomorphism of Lie group and h_0 is the identity of H. □

Remark 1.2.7.1. *Let $\widetilde{\mathbf{SL}(2,\mathbb{R})}$ is a Lie universal covering space of $\mathbf{SL}(2,\mathbb{R})$. This is a nonlinear Lie group, i.e., a group which is not represented by a group of matrixes.*
However, since $\mathbf{SL}(2,\mathbb{R}) \cong \mathbf{SO}(2) \times \mathbb{R}^2$ and $\widetilde{\mathbf{SO}(2)} \cong \mathbb{R}$ from $\mathbf{SO}(2) \cong S^1$, we conclude that

$$\widetilde{\mathbf{SL}(2,\mathbb{R})} \cong \widetilde{\mathbf{SO}(2)} \times \mathbb{R}^2 \cong \mathbb{R}^3.$$

We will prove the following later: Every homomorphism $\widetilde{\mathbf{SL}(2,\mathbb{R})} \to \mathbf{GL}(n,\mathbb{R})$ factors through $\mathbf{SL}(2,\mathbb{R})$, i.e.,

$$\begin{array}{ccc} \widetilde{\mathbf{SL}(2,\mathbb{R})} & \longrightarrow & \mathbf{GL}(n,\mathbb{R}) \\ \downarrow & \nearrow & \\ \mathbf{SL}(2,\mathbb{R}) & & \end{array}$$

Now, we give a definition of $\mathbf{Spin}(n)$:

Definition 1.2.7.2.

$$\mathbf{Spin}(n) \overset{def}{=} \widetilde{\mathbf{SO}(n)}.$$

Note that $\mathbf{Spin}(n) \hookrightarrow \mathbf{GL}(2^n,\mathbb{R})$ is a $2^n \times 2^n$ matrix.

1.2.8 Extension

Before we give various definitions, we want to give

Remark 1.2.8.1. *Lie classified Lie groups up to local isomorphisms by Lie algebras. Lie showed*

every solvable Lie algebra $\longleftrightarrow$ an algebra of upper triangle matrixes.

Killing and Cartan classified Lie algebras by means of simple extensions using structure theories of algebras. In this consideration, we have

a Lie algebra $\mathfrak{g}$ is semisimple $\longleftrightarrow$ $\mathfrak{g}$ has no nonzero solvable ideal.

This shows that a semisimple Lie algebra is a direct sum of simple Lie algebras. We shall see that solvable Lie algebra are iterated extensions of abelian Lie algebras.

Definition 1.2.8.1 (Extension). *Let A, B, and C be groups. B is an extension of A by C if the following sequence is exact:*

$$A \hookrightarrow B \twoheadrightarrow C.$$

That is, $A \xrightarrow{i} B$ is a monomorphism, $B \xrightarrow{j} C$ is an epimorphism, and $Im(i) = Ker(j)$.

Note that A is a normal subgroup of B and $B/A \cong C$.

Example 1.2.8.1. *Let $\mathcal{B}$ be a Borel subgroup and $\mathbb{R}_+$ be a multiplicative group of positive reals. Since we can identify*

$$\mathcal{B} \cong \{\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mid a > 0, b \in \mathbb{R}\},$$

by giving a map $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mapsto a$, we have an extension

$$\{\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid a > 0, b \in \mathbb{R}\} \hookrightarrow \mathcal{B} \twoheadrightarrow \mathbb{R}_+.$$

Definition 1.2.8.2 (Central extension). *Let A, B, and C be groups and $\mathcal{Z}(B)$ be the center of B. B is a central extension of A by C if $A \hookrightarrow B \twoheadrightarrow C$ is exact and $A \subset \mathcal{Z}(B)$.*

Example 1.2.8.2. *Let G be a group. Since $\mathcal{Z}(G) \triangleleft G$, we have a central extension*

$$\mathcal{Z}(G) \hookrightarrow G \twoheadrightarrow G/\mathcal{Z}(G).$$

Definition 1.2.8.3 (Upper central series). *For a given group G, define the following subgroups inductively: $\mathcal{Z}^0(G) = \mathbf{1}$ and $\mathcal{Z}^1(G) = \mathcal{Z}(G)$ and define $\mathcal{Z}^{i+1}(G)$ to be the subgroup of G containing $\mathcal{Z}^i(G)$ such that*

$$\mathcal{Z}^{i+1}(G)/\mathcal{Z}^i(G) = \mathcal{Z}(G/\mathcal{Z}^i(G)).$$

The following chain of subgroups is called the upper central series of G:

$$\mathbf{1} = \mathcal{Z}^0(G) \subset \mathcal{Z}(G) = \mathcal{Z}^1(G) \subset \mathcal{Z}^2(G) \subset \cdots \subset \mathcal{Z}^n(G) \subset \cdots.$$

Definition 1.2.8.4 (Nilpotent group). *A group G is nilpotent if the upper central series stops eventually. That is, there exists $n \in \mathbb{Z}$ such that $\mathcal{Z}^n(G) = G$.*

Example 1.2.8.3. *An abelian group G is nilpotent because $\mathcal{Z}^1(G) = G$. So, we can think an abelian groups as generalized nilpotent groups.*

Definition 1.2.8.5 (commutator). *Let G, H be groups. Suppose that $x, y \in G$.*

$$[x, y] \overset{def}{=} xyx^{-1}y^{-1}.$$

Also, the commutator of two groups is defined as

$$[G, H] \overset{def}{=} \langle [g, h] \mid g \in G, h \in H \rangle = \text{ a group generated by } ghg^{-1}h^{-1} \text{ for } g \in G, h \in H.$$

Inductively, we define an n-fold commutator of G by

$$\mathcal{C}_0(G) = G.$$
$$\mathcal{C}_1(G) = [G, G].$$
$$\mathcal{C}_{n+1}(G) = [\mathcal{C}_n(G), G].$$

Clearly, $G = \mathcal{C}_0(G) \supseteq \mathcal{C}_1(G)$. Suppose that $\mathcal{C}_{k-1}(G) \supseteq \mathcal{C}_k(G)$. From

$$\mathcal{C}_k(G) = [\mathcal{C}_{k-1}(G), G] \supseteq [\mathcal{C}_k(G), G] = \mathcal{C}_{k+1}(G),$$

we have $\mathcal{C}_k(G) \supseteq \mathcal{C}_{k+1}(G)$. Hence, we have:

Definition 1.2.8.6 (Lower central series). *For a given group G, we have the following chain of subgroups, which is called the lower central series of G:*

$$G = \mathcal{C}_0(G) \supseteq \mathcal{C}_1(G) \supseteq \mathcal{C}_2(G) \supseteq \cdots \supseteq \mathcal{C}_n(G) \subset \cdots .$$

Exercise 1.2.8.1. *G is nilpotent if and only if there exists N such that N-fold commutator of G is $\mathbf{1}$, i.e., $\mathcal{C}_N(G) = \mathbf{1}$.*

Proof. Note that $\mathbf{1} = [\mathcal{Z}(G/\mathcal{Z}^{k-1}(G)), G/\mathcal{Z}^{k-1}(G)] = [\mathcal{Z}^k(G)/\mathcal{Z}^{k-1}(G), G/\mathcal{Z}^{k-1}(G)]$ implies that

$$\mathcal{C}(\mathcal{Z}^k(G)) = [\mathcal{Z}^k(G), G] \subseteq \mathcal{Z}^{k-1}(G).$$

Notice that inductively, we have for all i, k,

$$\mathcal{C}_i(\mathcal{Z}^k(G)) \subseteq \mathcal{Z}^{k-i}(G).$$

Suppose that G is nilpotent, i.e., $\mathcal{Z}^N(G) = G$. Let $i, k = N$. We have

$$\mathcal{C}_N(\mathcal{Z}^N(G)) = \mathcal{C}_N(G) \subseteq \mathcal{Z}^0(G) = \mathbf{1}.$$

Note that $\mathbf{1} = [\mathcal{C}_{k-1}(G)/[\mathcal{C}_{k-1}(G), G], G/[\mathcal{C}_{k-1}(G), G]] = [\mathcal{C}_{k-1}/\mathcal{C}_k(G), G/\mathcal{C}_k(G)]$ implies that

$$\mathcal{C}_{k-1}(G)/\mathcal{C}_k(G) \subseteq \mathcal{Z}\big(G/\mathcal{C}_k(G)\big).$$

Suppose that $\mathcal{C}_N(G) = \mathbf{1}$. Clearly, $\mathbf{1} = \mathcal{C}_{N+1}(G) \subseteq \mathcal{Z}^1(G)$. Suppose that $\mathcal{C}_{N+1-k}(G) \subseteq \mathcal{Z}^k(G)$. So,

$$G/\mathcal{C}_{N+1-k}(G) \twoheadrightarrow G/\mathcal{Z}^k(G).$$

Since

$$\mathcal{C}_{N-k}(G)/\mathcal{C}_{N-k+1}(G) \subseteq \mathcal{Z}\big(G/\mathcal{C}_{N-k+1}(G)\big),$$

by the induction, we have

$$\big(\mathcal{C}_{N-k}(G) \cdot \mathcal{Z}^k(G)\big)/\mathcal{Z}^k(G) \subseteq \mathcal{Z}\big(G/\mathcal{Z}^k(G)\big) = \mathcal{Z}^{k+1}(G)/\mathcal{Z}^k(G).$$

That is,

$$\mathcal{C}_{N-k}(G) \subseteq \mathcal{C}_{N-k}(G) \cdot \mathcal{Z}^k(G) \subseteq \mathcal{Z}^{k+1}(G).$$

So, we have $G = \mathcal{C}_0(G) \subseteq \mathcal{Z}^{N+1}(G) = G$. □

By the virtue of Exercise 1.2.8.1 we conclude that G is nilpotent if and only if the lower central series of G stops eventually.

A strictly upper triangular $n \times n$ matrix is an upper triangular $n \times n$ matrix with zeros in the all diagonal entries. Let

$$\mathfrak{N}_n = \{ \text{ strictly upper triangular } n \times n \text{ matrixes. } \}$$

Note that $(\mathfrak{N}_n)^n = 0$, i.e., the multiple of any n elements is a zero matrix. Since $\mathfrak{N}_n$ does not have an identity and invertible elements, we adjoin those elements formally, i.e.,

$$\mathbf{1}_n \oplus \mathfrak{N}_n.$$

For $A, A' \in \mathfrak{N}_n$,

$$(\mathbf{1}_n + A)(\mathbf{1}_n + A') = \mathbf{1}_n + (A + A') + AA', \text{ which shows that it is closed under multiplication.}$$

Note that for $A \in \mathbf{M}_n$, we have

$$\mathbf{1}_n = (\mathbf{1}_n + A)(\mathbf{1}_n + (-1)A + A^2 + (-1)^3 A^3 + \cdots + (-1)^n A^n + \cdots).$$

If $A \in \mathfrak{N}_n$, then $A^n = 0$. So, we have

$$\mathbf{1}_n = (\mathbf{1}_n + A)(\mathbf{1}_n + (-1)A + A^2 + (-1)^3 A^3 + \cdots + (-1)^m A^m) \text{ for } m \leq n.$$

Combining closedness under multiplication, we deduce that $\mathbf{1}_n \oplus \mathfrak{N}_n$ is an algebra with invertible elements:

$$(\mathbf{1}_n + A)^{-1} = \frac{\mathbf{1}_n}{(\mathbf{1}_n + A)} = \mathbf{1}_n + (-1)A + A^2 + (-1)^3 A^3 + \cdots + (-1)^m A^m \text{ for } m \leq n.$$

Example 1.2.8.4.

$$\mathbf{1}_2 \oplus \mathfrak{N}_2 = \{\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \mid u \in \mathbb{R}\} \cong \mathbb{R}.$$

Example 1.2.8.5.

$$\textsf{\textit{Heisenberg group}}\ \mathbf{H}_3 = \mathbf{1}_3 \oplus \mathfrak{N}_3 = \{\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mid x, y, z \in \mathbb{R}\} \cong \mathbb{R}^3.$$

Consider

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x' & z' \\ 0 & 1 & y' \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x + x' & z + z' + xy' \\ 0 & 1 & y + y' \\ 0 & 0 & 1 \end{pmatrix}.$$

This implies that $\mathbf{H}_3 \xrightarrow{\varphi} \mathbb{R}^2$ is a homomorphism where

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} \text{ and } ker\varphi = \{\begin{pmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mid z \in \mathbb{R}\} \cong \mathbb{R}.$$

Also, by the above it is easy to see that

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -x & -z + xy \\ 0 & 1 & -y \\ 0 & 0 & 1 \end{pmatrix} \text{ and } [\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & x' & z' \\ 0 & 1 & y' \\ 0 & 0 & 1 \end{pmatrix}] = \begin{pmatrix} 1 & 0 & xy' - x'y \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since if $xy' - x'y = 0$ for all $x', y' \in \mathbb{R}$, then $x = y = 0$, we have $ker\varphi = \mathcal{Z}(\mathbf{H}_3)$. So, we deduce that $\mathbf{H}_3$ is a central extension of $\mathbb{R} \cong ker\varphi$ by $\mathbb{R}^2$, i.e., $ker\varphi \subseteq \mathcal{Z}(\mathbf{H}_3)$ and

$$\mathbb{R} \cong ker\varphi \hookrightarrow \mathbf{H}_3 \overset{\varphi}{\twoheadrightarrow} \mathbb{R}^2.$$

If C is a subgroup of $\mathcal{Z}(\mathbf{H}_3)$, then $C \triangleleft \mathbf{H}_3$. So, we define

$$\mathbf{H}_3^{\mathbf{red}} = \mathbf{H}_3 / \{ \begin{pmatrix} 1 & 0 & n \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mid n \in \mathbb{Z} \}.$$

Since $\mathbb{R}/\mathbb{Z} \cong S^1$, it is easy to see that $\mathbf{H}_3^{\mathbf{red}} \cong S^1 \times \mathbb{R}^2$.

Exercise 1.2.8.2. *Show that $\mathbf{H}_3^{\mathbf{red}}$ does not have a faithful matrix representation.*

Proof. □

Chapter 2

Lie Groups as Smooth Manifolds

2.1 Smooth Manifold Theory

2.1.1 Smooth manifolds and Tangent spaces

Let M be a second countable Hausdorff topological space.

Definition 2.1.1.1 (manifold). *We say M is **locally Euclidean** if every point $p \in M$ has a neighborhood $p \in U$ such that there exists a homeomorphism $U \xrightarrow{\varphi} U' \subseteq \mathbb{R}^n$, which is called a chart, where U' is an open set in $\mathbb{R}^n$.*

*We say M is a manifold if it is a **locally Euclidean** second countable Hausdorff topological space.*

Note that we call a collection of charts as an atlas and we always take a maximal atlas.

Definition 2.1.1.2 (smooth manifold). *We say M is a smooth manifold if it is a manifold with a smooth atlas, i.e., for charts $\varphi_U : U \to U' \subseteq \mathbb{R}^n$ and $\varphi_V : V \to V' \subseteq \mathbb{R}^n$, the transition map $g_{UV} \overset{def}{=} \varphi_U \circ \varphi_V^{-1} : \varphi_V(U \cap V) \to \varphi_U(U \cap V)$ is C^∞.*

In this description, we can define

Definition 2.1.1.3 (smooth map). *Let M^m be an m-dimensional smooth manifold and N^n an n-dimensional smooth manifold. We $M^m \xrightarrow{f} N^n$ is smooth if for all patches (U, φ_U) in M and (V, ψ_V) in N, we have*

$$\varphi_U(U \cap f^{-1}(V)) \xrightarrow{\psi_V \circ f \circ \varphi_U^{-1}} \psi_V(f(U) \cap V) \text{ is smooth.}$$

Definition 2.1.1.4 (Lie groups). *A Lie group G is a differentiable manifold with smooth group operations*

Note that if G is a topological manifold with continuous group operations, then it is, in fact, a Lie group in the above sense, which was suggested by David Hilbert as his fifth problem and proved by A. Gleason, D. Montgomery, and L. Zippin.

Sometimes, a Lie group is defined by a real analytic manifold (C^ω-manifold), i.e., transition maps are given by convergent Taylor series with real analytic group operations. It is well- known that $C^\omega \subsetneq C^\infty$, for example, $e^{-\frac{1}{x^2}}$. Without using the fifth problem, in fact, it can be shown that smooth group operations in a Lie group are real analytic group operations. So, no generality is lost by assuming a more restrictive definition.

Let

$$C^\infty(M) = \{f \mid f : M \to \mathbb{R} \text{ is smooth.}\}$$

We can make $C^\infty(M)$ an $\mathbb{R}$-algebra by defining operations as follows:

$$(f+g)(p) = f(p) + g(p) \text{ and } (f \cdot g)(p) = f(p) \cdot g(p).$$

Every point $p \in M$ defines an $\mathbb{R}$-algebra homomorphism between $C^\infty(M)$ and $\mathbb{R}$:

$$C^\infty(M) \xrightarrow{\epsilon_p} \mathbb{R} \text{ by } \epsilon_p(f) = f(p).$$

Since $\mathbb{R}$ is a simple algebra, $ker(\epsilon_p)$ is a maximal ideal of $C^\infty(M)$. The next theorem shall make local information into global information:

Theorem 2.1.1.1 (Partition of Unity). *Let M be a smooth manifold and $\mathcal{U} = \{U_\alpha\}$ be an open cover of M, i.e., $M = \cup U_\alpha$, then there exists a smooth partition of unity $\{\psi_\alpha\}$ subordinate to the open cover $\mathcal{U}$. That is,*

1. For each α, $\psi_\alpha \geq 0$ and $1 \equiv \sum_\alpha \psi_\alpha(p)$ for $p \in M$

2. (Smoothness) $\psi_\alpha \in C^\infty(M)$

3. (Local finiteness) Each point $p \in M$ has an open set U such that

$$supp(\psi_\alpha) \cap U \neq \emptyset \text{ for only finite number of } \alpha$$

4. (Subordination) For each α,

$$supp(\psi_\alpha) = \{u \in M \mid \psi_\alpha(u) \neq 0\} \subseteq U_\alpha.$$

Let $f \in C^\infty(U)$ where U is a small open neighborhood of p in M and we define an equivalent class $[f]_p$, which is called the germ of f at p, in the following ways:

$$[f_1]_p = [f_2]_p \text{ if there exists an open set } W \subseteq U \text{ such that } p \in W \text{ and } f_1 \equiv f_2 \text{ on } W.$$

Let $\gamma : (-\epsilon, \epsilon) \to U$ be a path in U such that $\gamma(0) = p$. So, we have a path $f \circ \gamma : (-\epsilon, \epsilon) \to \mathbb{R}$. We define a derivative operator D_γ of f with respect to γ by $D_\gamma([f]_p) = (f \circ \gamma)'(0)$. It is easy to see that D_γ defines an $\mathbb{R}$-linear homomorphism from the set of equivalent classes at p of $C^\infty(U)$ to $\mathbb{R}$:

$$\begin{aligned} D_\gamma(c) &= 0 \text{ for } c \in \mathbb{R} \\ D_\gamma(c_1[f_1]_p + c_2[f_2]_p) &= c_1 D_\gamma([f_1]_p) + c_2 D_\gamma([f_2]_p) \\ D_\gamma([f_1]_p \cdot [f_2]_p) &= D_\gamma(f_1)\epsilon_p([f_2]_p) + \epsilon_p(f_1) D_\gamma([f_2]_p). \end{aligned}$$

Remark 2.1.1.1. *Foe simplicity, we will write f as $[f]_p$. As far as local notions are concerned such as derivative operators and tangent spaces, etc., readers should think that we only deal with the germs, not functions. So, there will be no confusion. However, when we deal with global notions such as tangent bundles and derivations, we will give appropriate remarks*

We define an equivalent class $[\gamma]_p$ of smooth paths through p in the followings:

$$[\gamma_1]_p = [\gamma_2]_p \Longleftrightarrow \gamma_1 \backsim \gamma_2 \text{ if } \gamma_1(0) = \gamma_2(0) = p \text{ and } \gamma_1'(0) = \gamma_2'(0).$$

Let $Der(C^\infty(M), \epsilon_p)$ be the set of all derivative operators induced from equivalent classes of smooth paths such that $\gamma(0) = p$.

Definition 2.1.1.5 (Tangent Spaces). *We define a tangent space $T_p(M)$ at p to be $Der(C^\infty(M), \epsilon_p)$.*

The next lemma shall show that we can make $T_p(M)$ a vector space.

Lemma 2.1.1.1. *For $\gamma_1, \gamma_2 \in T_p(M)$ and $c_1, c_2 \in \mathbb{R}$, there exists a unique $[\gamma]_p$ such that for all $f \in C^\infty(M)$,*

$$(f \circ \gamma)'(0) = c_1 D_{\gamma_1}(f) + c_2 D_{\gamma_2}(f).$$

Proof. Since we are dealing with an infinitesimal path, without loss of generality we can assume $U = \mathbb{R}^n$ and $p = \mathbf{0}$. Moreover, since we are dealing with a smooth function f, for sufficiently small t near $\mathbf{0}$, we always that

$$f(c_1\gamma_1(t) + c_2\gamma_2(t)) = f(c_1\gamma_1(t)) + f(c_2\gamma_2(t)).$$

Let $\gamma(t) = c_1\gamma_1(t) + c_2\gamma_2(t)$. Uniqueness follows from the definition of equivalent classes. $\square$

Defining $(c_1 D_{\gamma_1} + c_2 D_{\gamma_2})(f) = c_1 D_{\gamma_1}(f)) + c_2 D_{\gamma_2}(f)$, by Lemma 2.1.1.1, we have

$$c_1 D_{\gamma_1} + c_2 D_{\gamma_2} \in T_p(M).$$

Let φ be a chart of an open neighborhood U of $p \in M$ such that $\varphi(p) = \mathbf{0}$. So, for $m \in M$, we have coordinates

$$\varphi(m) = (x_1(m), \dots, x_n(m)).$$

Let $\varphi \circ \gamma_i(t) = (0, \dots, 0, t, 0, \dots, 0)$, i.e., i-th coordinate is t. For all $f \in C^\infty(U)$, we have

$$D_{\gamma_i}(f) = \frac{d(f \circ \gamma_i)(t)}{dt}|_{t=0} = \frac{\partial f(m)}{\partial x_i(m)}|_{m=p}.$$

So, we have $D_{\gamma_i} = \frac{\partial}{\partial x_i} = \partial_i \in T_p(M)$.

Exercise 2.1.1.1. *$\{\partial_1, \dots, \partial_n\}$ is a basis of $T_p(M) = Der(C^\infty(M), \epsilon_p)$.*

Proof. Suppose that $\sum_{i=1}^n c_i \partial_i = 0$. Take $f = x_i$. Since $\sum_{i=1}^n c_i \partial_i f = c_i = 0$, we conclude that $\{\partial_1, \dots, \partial_n\}$ are linearly independent.

Now, we are going to show that it is a spanning set, which is not so trivial. Also, we remark that the following proof works only in a C^∞-category. First of all, we want to remind you that

$$\mathfrak{m}_p \overset{def}{=} ker(\epsilon_p) = \{[f]_p \in C^\infty(U)_p \mid \epsilon_p(f) = f(p) = 0\} \text{ is a maximal ideal of } C^\infty(U).$$

Again, $C^\infty(U)_p$ means a local ring, which is the set of germs. From now on, we omit the local ring notions for simplicity. However, reader should think them as equivalent classes. So, as far as local properties are concerned, hoping reader's generosity we will write

$$\mathfrak{m}_p \overset{def}{=} ker(\epsilon_p) = \{f \in C^\infty(U) \mid \epsilon_p(f) = f(p) = 0\} \text{ is a maximal ideal of } C^\infty(U).$$

From the property of a derivative operator

$$D_\gamma(c_1 f_1 + c_2 f_2) = c_1 D_\gamma(f_1) + c_2 D_\gamma(f_2),$$

it easy to see that a derivative operator D_γ defines a linear functional on $C^\infty(U)$. So, it will define a linear functional on $\mathfrak{m}_p$. Let

$$\mathfrak{m}_p^2 = \{F \in C^\infty(U) \mid F = \sum_{i=1}^{k} f_i \cdot g_i \text{ where } f_i, g_i \in \mathfrak{m}_p\}.$$

By the following property of a derivative operator

$$D_\gamma(f_1 \cdot f_2) = D_\gamma(f_1)\epsilon_p(f_2) + \epsilon_p(f_1) D_\gamma(f_2),$$

it is easy to see that D_γ defines a linear functional on $\mathfrak{m}_p/\mathfrak{m}_p^2$. Note that the maximality of $\mathfrak{m}_p$ in $C^\infty(U)$ shows that $C^\infty(U)/\mathfrak{m}_p$ is a field. Since $\mathfrak{m}_p/\mathfrak{m}_p^2$ is a $C^\infty(U)/\mathfrak{m}_p$-module, it is a vector space. Now, we shall show that $Der(C^\infty(M), \epsilon_p)$ is isomorphic to the dual vector space of $\mathfrak{m}_p/\mathfrak{m}_p^2$, i.e., $(\mathfrak{m}_p/\mathfrak{m}_p^2)^*$. Since we already showed that there is a well-defined identification from $Der(C^\infty(M), \epsilon_p)$ to $(\mathfrak{m}_p/\mathfrak{m}_p^2)^*$, it suffices to show that it is one-to-one and onto.
Suppose that for all $f \in \mathfrak{m}_p/\mathfrak{m}_p^2$, we have $D_\gamma(f) = 0$. So, $D_\gamma(x_i) = 0$ for $i = 1, \ldots, n$. That is,

$$\frac{d(x_i \circ \gamma)(t)}{dt}|_{t=0} = 0 \text{ for } i = 1, \ldots, n, \text{ which implies that } \gamma'(0) = 0.$$

So, $\gamma \backsim \gamma_1$ where $\gamma_1(t) = p$, i.e., constant function. Hence, $D_\gamma = 0$.
In order to prove that it is onto, we need a preparation theorem from calculus:

Theorem 2.1.1.2. *Let $f \in C^\infty(U)$ where U is a small open neighborhood of p such that $\varphi(U)$ is convex in $\mathbb{R}^n$ for a chart $\varphi(u) = (x_1(u), \ldots, x_n(u))$. For each $u, v \in U$, we have*

$$\begin{aligned} f(u) = f(p) &+ \sum_{i=1}^{n} \frac{\partial f(u)}{\partial x_i(u)}|_{u=p}(x_i(u) - x_i(p)) \\ &+ \sum_{i,j=1}^{n} (x_i(u) - x_i(p))(x_j(u) - x_j(p)) \int_0^1 (1-t) \frac{\partial^2 f(v)}{\partial x_i(v) \partial x_j(v)}|_{\varphi(v)=\varphi(p)+t(\varphi(u)-\varphi(p))} dt. \end{aligned}$$

The above theorem shows that $\mathfrak{m}_p/\mathfrak{m}_p^2$ is a n-dimensional vector space with a basis $x_i(u) - x_i(p)$ for $i = 1, \ldots, n$ in the C^∞-category. In general, $\mathfrak{m}_p/\mathfrak{m}_p^2$ is an infinite dimensional vector space.
Suppose that $L \in (\mathfrak{m}_p/\mathfrak{m}_p^2)^*$. Note that L is completely determined by $L(x_i(u) - x_i(p))$ for $i = 1, \ldots, n$. Let

$$\gamma(t) = \varphi^{-1}\big(L(x_1(u) - x_1(p))t, \ldots, L(x_n(u) - x_n(p))t\big), \text{ which is a path in } U.$$

It is easy to see that $D_\gamma(x_i) = \frac{d(x_i \circ \gamma(t))}{dt}|_{t=0} = L(x_i(u) - x_i(p))$ for $i = 1, \ldots, n$.

So, we have for $f \in \mathfrak{m}_p/\mathfrak{m}_p^2$,

$$\begin{aligned}
D_\gamma(f) &= D_\gamma\big(f(p) + \sum_{i=1}^{n} \frac{\partial f(u)}{\partial x_i(u)}|_{u=p}(x_i(u) - x_i(p))\big) = D_\gamma\big(\sum_{i=1}^{n} \frac{\partial f(u)}{\partial x_i(u)}|_{u=p}(x_i(u) - x_i(p))\big) \\
&= \sum_{i=1}^{n} \frac{\partial f(u)}{\partial x_i(u)}|_{u=p} D_\gamma\big(x_i(u) - x_i(p)\big) = \sum_{i=1}^{n} \frac{\partial f(u)}{\partial x_i(u)}|_{u=p} D_\gamma\big(x_i(u)\big) \\
&= \sum_{i=1}^{n} \frac{\partial f(u)}{\partial x_i(u)}|_{u=p} \frac{d(x_i \circ \gamma(t))}{dt}|_{t=0} = \sum_{i=1}^{n} \frac{\partial f(u)}{\partial x_i(u)}|_{u=p} L(x_i(u) - x_i(p)) \\
&= L\big(\sum_{i=1}^{n} \frac{\partial f(u)}{\partial x_i(u)}|_{u=p}(x_i(u) - x_i(p))\big) = L(f).
\end{aligned}$$

So, we show that $Der(C^\infty(M), \epsilon_p) = (\mathfrak{m}_p/\mathfrak{m}_p^2)^*$, which is an n-dimensional vector space. Since we know that $\{\partial_1, \ldots, \partial_n\}$ are linearly independent in $T_p(M) = Der(C^\infty(M), \epsilon_p)$, it is a basis. ☐

It is worth remarking what we showed in the proof of Exercise 2.1.1.1.

Remark 2.1.1.2. *We showed that $T_p(M)$ is an n-dimensional vector space. So, $T_p(M) \cong \mathbb{R}^n$. Since $T_p(M)$ is a local notion, we have $T_p(M) \cong T_q(M)$ fro $p, q \in M$.*
We also showed that $T_p(M) = Der(C^\infty(M), \epsilon_p) = (\mathfrak{m}_p/\mathfrak{m}_p^2)^*$. So, $(T_P(M))^* = \mathfrak{m}_p/\mathfrak{m}_p^2$. That is, $Der(C^\infty(M), \epsilon_p) \times \mathfrak{m}_p/\mathfrak{m}_p^2 \to \mathbb{R}$ is a perfect paring.

Let $M^m \xrightarrow{\psi} N^n$ be a smooth map. We define $(d\psi)_p$ in the following way: Let $D \in Der(C^\infty(M), \epsilon_p)$ and $f \in C^\infty(U)$ where $U \subseteq N$ is a small open neighborhood of $\psi(p)$. Define

$$((d\psi)_p(D))(f) \overset{def}{=} D(f \circ \psi) \text{ and } \psi^*(f) \overset{def}{=} f \circ \psi.$$

A straightforward checking will show that $(d\psi)_p(D) \in Der(C^\infty(N), \epsilon_{\psi(p)})$ and $(d\psi)_p$ induces a homomorphism from $T_p(M)$ to $T_{\psi(p)}(N)$. Moreover, chain rules will show that if

$$p \in M^m \xrightarrow{\psi} N^n \xrightarrow{\varphi} Q^q,$$

then we have $d(\varphi \circ \psi)_p = (d\varphi)_{\psi(p)} \circ (d\psi)_p$ where

$$T_p(M^m) \xrightarrow{(d\psi)_p} T_{\psi(p)}(N^n) \xrightarrow{(d\varphi)_{\psi(p)}} T_{(\varphi\circ\psi)(p)}(Q^q) \text{ and } T_p(M^m) \xrightarrow{d(\varphi\circ\psi)_p} T_{(\varphi\circ\psi)(p)}(Q^q).$$

Also, it is easy to see that if $M^m \xrightarrow{\psi} N^n$ be a diffeomorphism, then $(d\psi)_p$ is an isomorphism between $T_p(M)$ to $T_{\psi(p)}(N)$.

Definition 2.1.1.6 (Tangent Bundles). *Define the tangent bundle $T(U)$ over an open set $U \subseteq M$ to be a disjoint union of tangent spaces. That is,*

$$T(U) \overset{def}{=} \coprod_{p\in U} T_p(M).$$

Note that since $T_p(M) \cong \mathbb{R}^n$, we give the topology of $T(U)$ from the topology of $U \times \mathbb{R}^n$. So, it has a natural smooth structure. Also, note that locally we have $T(U) \cong \mathbb{R}^{2n}$.

2.1.2 Vector fields as Derivations

Definition 2.1.2.1 (Vector fields). *Let* $T(M) \xrightarrow{\pi} M$ *be* $\pi(p \times T_p(M)) = p$.
A section $M \xrightarrow{s} T(M)$ *is a smooth map such that* $\pi \circ s(p) = p$. *That is,* $\pi \circ s = id_M$. *A vector field* X *is a section of* $T(M)$.

In general, for any vector bundle $E \xrightarrow{\pi} M$, we say that s is a section if $\pi \circ s = id_M$. So, what the above definition means is that a vector field is a section of **tangent bundle.**

Definition 2.1.2.2 (Derivations). *A derivation* D *of an* $\mathbb{R}$*-algebra* A *is a linear map such that* $D(fg) = D(f)g + fD(g)$ *for all* $f, g \in A$. *That is, the set* $Der(A)$ *of all derivations of* A *is given by*

$$Der(A) = \{D \in \mathbf{End}(A) \mid D(fg) = D(f)g + fD(g) \text{ for all } f, g \in A\}.$$

It is worth remarking that $Der(A)$ is a left A-module: Let $D \in Der(A)$ and $f, g, h \in A$.

$$fD(gh) = (fD(g))h + g(fD(h)) = (fD)(gh).$$

Exercise 2.1.2.1. *Show that we can make* $Der(A)$ *a Lie algebra.*

Proof. By defining for all $a, d \in \mathbb{R}$, all $D_1, D_2 \in Der(A)$, and all $f \in A$,

$$(aD_1 + bD_2)(f) = aD_1(f) + bD_2(f),$$

we conclude that $Der(A)$ is a vector space. Also, defining a multiplicative structure $[,]$ by

$$[D_1, D_2](f) = D_1(D_2(f)) - D_2(D_1(f)),$$

a straightforward checking will show that

1. $[D_1, D_2] \in Der(A)$
2. $[,]$ is $\mathbb{R}$-bilinear
3. $[D_1, D_2] = -[D_2, D_2]$
4. $[D_1, [D_2, D_3]] + [D_2, [D_3, D_1]] + [D_3, [D_1, D_2]] = 0.$

So, $Der(A)$ is a Lie algebra. ☐

Since $C^\infty(U)$ is an $\mathbb{R}$-algebra for any open $U \subset M$, we have $Der(C^\infty(M))$. Let $\mathbf{Vect}(U) = \mathfrak{X}(U)$ be the set of all vector fields over for an open $U \subset M$. Now, we shall show that $\mathfrak{X}(U) = Der(C^\infty(U))$. So, $\mathfrak{X}(U)$ is a Lie algebra.

Theorem 2.1.2.1. $\mathfrak{X}(U) \subseteq Der(C^\infty(U))$ *for any open* $U \subset M$.

Proof. Note that $X_p \in T_p(M) = Der(C^\infty(M), \epsilon_p)$ and in Exercise 2.1.1.1, we show that $\{\partial_1, \ldots, \partial_n\}$ is a basis of $T_p(M) = Der(C^\infty(M), \epsilon_p)$. However, it was a kind of misleading notation. We should have written $\{\partial_{1,p}, \ldots, \partial_{n,p}\}$ as a basis of $T_p(M) = Der(C^\infty(M), \epsilon_p)$. Now, let

$$\partial_i(p) = \partial_{i,p}. \text{ That is, } \pi \circ \partial_i = id_M.$$

We want to emphasize that ∂_i need not be a vector field, since it need not be smooth. However, what is true is the following: If $X \in \mathfrak{X}(U)$ for an open $U \subset M$, there exist $\partial_i \in \mathfrak{X}(U)$ such that

$$X_u = \sum_{i=1}^{k} f_i(u)\partial_{i,u} \text{ where } f_i \in C^\infty(U), u \in U \text{ and } k \leq n.$$

Let $\psi, \varphi \in C^\infty(U)$. It is obvious that $X_u(a\psi + b\varphi) = aX_u(\psi) + bX_u(\varphi)$, which shows that $X_u \in$ **End**$(C^\infty(U))$. Also, we have

$$\begin{aligned} X_u(\psi \cdot \varphi) = \sum_{i=1}^{k} f_i(u)\partial_{i,u}(\psi \cdot \varphi) &= \sum_{i=1}^{k} f_i(u)\epsilon_u(\psi)\partial_{i,u}(\varphi) + \sum_{i=1}^{k} f_i(u)\epsilon_u(\varphi)\partial_{i,u}(\psi) \\ &= \psi(u)\sum_{i=1}^{k} f_i(u)\partial_{i,u}(\varphi) + \varphi(u)\sum_{i=1}^{k} f_i(u)\partial_{i,u}(\psi) = \psi(u)X_u(\varphi) + X_u(\psi)\varphi(u). \end{aligned}$$

That is, $X(\psi \cdot \varphi) = \psi X(\varphi) + X(\psi)\varphi$, which implies that $\mathfrak{X}(U) \subseteq Der(C^\infty(U))$. □

In order to show the reversed inclusion, we need some preliminary. First of all, **we want to clarify a vector field notation**

Notation 2.1.2.1. *We will write $\{\partial_1, \ldots, \partial_n\}$ as linearly independent vector fields. That is, $\partial_i \in \mathfrak{X}(U)$ for an open $U \subset M$ implies that ∂_i is smooth over U. However, there is another notation, which expresses a vector field. Let U be an open set in M and $\mathcal{V} = \{V_\alpha\}$ be an atlas of U. So, for each open set V_α, we have a chart $(x_{\alpha,1}(v), \ldots, x_{\alpha,n}(v))$. So, for each ζ such that $\zeta_p \in T_p(U)$ and $\pi \circ \zeta = id_U$, we can have*

$$\zeta_{(x_{\alpha,1}(v),\ldots,x_{\alpha,n}(v))} = \sum_{i=1}^{n} f_i(x_{\alpha,1}(v), \ldots, x_{\alpha,n}(v))\frac{\partial}{\partial x_{\alpha,i}}.$$

Note that $\frac{\partial}{\partial x_{\alpha,i}} \in \mathfrak{X}(V_\alpha)$. In this description, we say $\zeta \in \mathfrak{X}(U)$ if $f_i(x_{\alpha,1}(v), \ldots, x_{\alpha,n}(v)) \in C^\infty(U)$ for $i = 1, \ldots, n$ and all α. For simplicity, we will write $\frac{\partial}{\partial x_i}$ as $\frac{\partial}{\partial x_{\alpha,i}}$ in the understanding that x_i is a local coordinate of each V_α.

Let $D \in Der(C^\infty(U))$ and $f \in C^\infty(U)$. By definition, $D(f) \in C^\infty(U)$. So,

$$D(f)(p) \stackrel{def}{=} D_p(f) = \epsilon_p(Df) \in \mathbb{R} \text{ for } p \in U.$$

In this consideration, we can expect that D_p might give a derivative operator. However, strictly speaking, a derivative operator was defined on a local ring $C^\infty(U)_p$ and a derivation was defined on a genuine function space. So, first of all, we will show that it is not a real obstacle. Note that that's is one of reasons why we kept maintaining function notations instead of more logical equivalent class notations.

Lemma 2.1.2.1. *If $[f]_p = [g]_p$, then $D_p(f) = D_p(g)$.*

Proof. Suppose that $f|_V \equiv 0$ for $V \subseteq U$. Choose a bump function $\varphi \in C^\infty(U)$ such that

$$\varphi(x) = 0 \text{ on a compact } K \subset V \text{ and } \varphi|_{U/V} \equiv 1.$$

So, $\varphi f = f$ on U. By the definition of D, we have

$$D(f) = D(\varphi f) = D(\varphi)f + \varphi D(f).$$

So, we have $D(f)|_K \equiv 0$. by letting $K = \{p\}$ for $p \in V$, we have $D(f)|_V \equiv 0$.

If $[f]_p = [g]_p$, there exists an open set $W \subseteq U$ such that $p \in W$ and $f = g$ on W. So,

$$(D(f) - D(g))|_W = D(f - g)|_W \equiv 0.$$

That is, $D_p(f) = D_p(g)$ for all $p \in W$. □

From Lemma 2.1.2.1 and the Leibnitz rule of D as the definition of D ,it is easy to see that $D_p(f.g) = 0$ for all $[f]_p, [g]_p \in \mathfrak{m}_p$. So, we have for each $p \in U$

$$D_p \in (\mathfrak{m}_\mathfrak{p}/\mathfrak{m}_p^2)^* = T_p(M) \stackrel{def}{=} Der(C^\infty(U), \epsilon_p).$$

Hence, we conclude that each $D \in Der(C^\infty(U))$ satisfy $\pi \circ D = id_U$ for $T(U) \xrightarrow{\pi} U$ for any open $U \subseteq M$. Now, we show that D_p varies smoothly as p varies. Keeping the content of Notation 2.1.2.1 in your mind, we can write at least.

$$D_p = \sum_{i=1}^{n} f_i(x_1(p), \ldots, x_n(p)) \frac{\partial}{\partial x_i}.$$

It suffices to show that each $f_i(x_1(v), \ldots, x_n(p)) \in C^\infty(U)$. Note that by definition $D : C^\infty(U) \to C^\infty(U)$. So, we have

$$D_p(x_i) = f_i(p) \in C^\infty(U).$$

Therefore, we have

Theorem 2.1.2.2. *$\mathfrak{X}(U) \supseteq Der(C^\infty(U))$ for any open $U \subset M$.*

2.1.3 Lie bracket of Vector fields

Exercise 2.1.3.1. *Compute $[\zeta, \xi]$ for vector fields $\zeta, \xi \in \mathfrak{X}(M)$.*

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

$$\zeta = \sum_{i=1}^{n} \zeta_i(x_1, \ldots, x_n) \frac{\partial}{\partial x_i} \text{ and } \xi = \sum_{i=1}^{n} \xi_i(x_1, \ldots, x_n) \frac{\partial}{\partial x_i}.$$

$$\begin{aligned}[\zeta, \xi]f &= \sum_{i=1}^{n} \zeta_i(x_1, \ldots, x_n) \frac{\partial}{\partial x_i}(\sum_{j=1}^{n} \xi_j(x_1, \ldots, x_n) \frac{\partial f}{\partial x_j}) - \sum_{i=1}^{n} \xi_i(x_1, \ldots, x_n) \frac{\partial}{\partial x_i}(\sum_{j=1}^{n} \zeta_j(x_1, \ldots, x_n) \frac{\partial f}{\partial x_j}) \\ &= \sum_{i,j=1}^{n} \big(\zeta_i(x_1, \ldots, x_n) \frac{\partial \xi_j}{\partial x_i} - \xi_i(x_1, \ldots, x_n) \frac{\partial \zeta_j}{\partial x_i}\big) \frac{\partial f}{\partial x_j}\end{aligned}$$

So, we have

$$[\zeta, \xi] = \sum_{i,j=1}^{n} \big(\zeta_i(x_1, \ldots, x_n) \frac{\partial \xi_j}{\partial x_i} - \xi_i(x_1, \ldots, x_n) \frac{\partial \zeta_j}{\partial x_i}\big) \frac{\partial}{\partial x_j}.$$

□

Since we know $Der(C^\infty(M))$ is a left $C^\infty(M)$ module, we have $\mathfrak{X}(M)$ is a left $C^\infty(M)$ module. Also, for $f, g \in C^\infty(M)$, it is easy to see that

$$[f\zeta, g\xi] = f \cdot \zeta(g)\xi - g \cdot \xi(f)\zeta + f \cdot g[\zeta, \xi].$$

Example 2.1.3.1. *If $\partial_i, \partial_j \in \mathfrak{X}(U)$, then*

$$[\partial_i, \partial_j] = 0.$$

Example 2.1.3.2. *If $\partial_k \in \mathfrak{X}(U)$, then*

$$[\partial_k, \sum_{j=1}^n f_j \frac{\partial}{\partial x_j}] = \sum_{j=1}^n \frac{\partial f_j}{\partial x_k} \frac{\partial}{\partial x_j}.$$

Definition 2.1.3.1 (Parallel vector fields and Linear vector fields). *Let $\xi \in \mathfrak{X}(U)$, i.e.,*

$$\xi = \sum_{i=1}^n \xi_i(x_1, \ldots, x_n) \frac{\partial}{\partial x_i}.$$

We say ξ is a ***linear*** *vector field if each ξ_i is linear. We say ξ is a* ***parallel or constant*** *vector field if each ξ_i is constant.*

Now, we are going to show that $\mathbf{M}_n(\mathbb{R}) = \mathfrak{gl}(n, \mathbb{R})$ generates linear vector fields ξ on $\mathbb{R}^n$, i.e., $\xi \in \mathfrak{X}(\mathbb{R}^n)$: Let $A = (a_{ij})_{n\times n} \in \mathbf{M}_n(\mathbb{R})$ and $x_1, \ldots, x_n$ be Euclidean coordinates of $\mathbb{R}^n$. Define

$$\xi_A = \sum_{i,j=1}^n a_{ij} x_j \frac{\partial}{\partial x_i}.$$

Since the manifold which we are dealing with is $\mathbb{R}^n$ itself, we have $\xi_A \in \mathfrak{X}(\mathbb{R}^n)$. That is, $\mathbf{M}_n(\mathbb{R}) = \mathfrak{gl}(n, \mathbb{R})$ **generates all the linear vector fields of $\mathbb{R}^n$.** The following exercise will exhibit that $\mathbf{M}_n(\mathbb{R}) = \mathfrak{gl}(n, \mathbb{R})$ acts on the set of linear vector fields on $\mathbb{R}^n$ with respect to Lie bracket actions.

Exercise 2.1.3.2. *Show that $[\xi_A, \xi_B] = \xi_{AB-BA}$ for $A, B \in \mathbf{M}_n(\mathbb{R})$.*

Proof. Let

$$\xi_A = \sum_{i,j=1}^n a_{ij} x_j \frac{\partial}{\partial x_i} \text{ and } \xi_B = \sum_{i,j=1}^n b_{ij} x_j \frac{\partial}{\partial x_i}.$$

We have

$$\begin{aligned}
[\xi_A, \xi_B] &= \sum_{i,k=1}^n \Big((\sum_{j=1}^n a_{ij} x_j) \frac{\partial(\sum_{j=1}^n b_{kj} x_j)}{\partial x_i} - (\sum_{j=1}^n b_{ij} x_j) \frac{\partial(\sum_{j=1}^n a_{kj} x_j)}{\partial x_i} \Big) \frac{\partial}{\partial x_k} \\
&= \sum_{i,k=1}^n \Big((\sum_{j=1}^n a_{ij} x_j) b_{ki} - (\sum_{j=1}^n b_{ij} x_j) a_{ki} \Big) \frac{\partial}{\partial x_k} = \sum_{i,k=1}^n \Big(\sum_{j=1}^n (a_{ij} b_{ki} - b_{ij} a_{ki}) x_j \Big) \frac{\partial}{\partial x_k} \\
&= \sum_{j,k=1}^n \Big(\sum_{i=1}^n (a_{ij} b_{ki} - b_{ij} a_{ki}) x_j \Big) \frac{\partial}{\partial x_k} = \xi_{AB-BA}.
\end{aligned}$$

□

Example 2.1.3.3. $\mathbf{M}_n(\mathbb{R}) = \mathfrak{gl}(n, \mathbb{R})$ *acts on the set of parallel vector fields on* $\mathbb{R}^n$*.*

Proof. First of all, note that the action of ξ_A is Lie brackets. Since every parallel vector field of $\mathbb{R}^n$ is generated by $\frac{\partial}{\partial x_i}$ for $i = 1, \ldots, n$, it suffices to consider just one case:

$$[\xi_A, \frac{\partial}{\partial x_i}] = [\sum_{i,j=1}^{n} a_{ij} x_j \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_k}] = \sum_{i=1}^{n} a_{ik} \frac{\partial}{\partial x_i}, \text{ which is a parallel vector field.}$$

□

Example 2.1.3.4 (Compare to Example 2.1.3.1). *Let our manifold be* $\mathbb{R}^n$*. We have*

$$[-x_i \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_i}] = \frac{\partial}{\partial x_i} \text{ and } [\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0.$$

Example 2.1.3.5. *Show that* $\mathbf{affin}(\mathbb{R}) = \{(ax + b)\frac{\partial}{\partial x} \mid a, b \in \mathbb{R}\}$ *is a 2-dimensional Lie algebra of vector fields on* $\mathbb{R}$*.*

Proof. We have to check Lie bracket operations of $\frac{\partial}{\partial x}$ and $x\frac{\partial}{\partial x}$:

$$[x\frac{\partial}{\partial x}, \frac{\partial}{\partial x}] = -\frac{\partial}{\partial x} \text{ and } [\frac{\partial}{\partial x}, \frac{\partial}{\partial x}] = [x\frac{\partial}{\partial x}, x\frac{\partial}{\partial x}] = 0.$$

So, it is a Lie algebra generated by $\frac{\partial}{\partial x}$ and $x\frac{\partial}{\partial x}$. □

Definition 2.1.3.2 (Flows or One parameter subgroups). *We say* $\Phi(t, x) = \Phi_t(x)$ *is a global flow (or one parameter subgroup) on* M *if* $\Phi : \mathbb{R} \times M \to M$ *is smooth and it satisfies* $\Phi_0 = id_M$ *and* $\Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2}$*, i.e.,* $\Phi(t_1 + t_2, x) = \Phi(t_1, \Phi(t_2, x))$ *for all* $t_1, t_2 \in \mathbb{R}$*.*

This definition has immediate consequences: If Φ_t is a flow, then Φ_t has an inverse by definition, namely, Φ_{-t}. So, the smoothness gives that $\Phi_t \in \mathbf{Diff}(M)$, the set of diffeomorphisms of M. Moreover, the condition $\Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2}$ gives us that $\mathbb{R} \xrightarrow{\Phi} \mathbf{Diff}(M)$ is a homomorphism. The image of this homomorphism is a subgroup of $\mathbf{Diff}(M)$, which is parameterized by one dimensional group $\mathbb{R}$. That is why Φ is often called a **one parameter subgroup**.

Now, we want to relate a flow to a vector field of M. Let $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$. We know that γ gives a tangent vector as a derivative operator, i.e., $D_\gamma \in T_p(M) = Der(C^\infty(M), \epsilon_p)$. Since $\Phi(t, x) \in m$, fixing x and varying t it is easy to see that $\Phi(t, x)$ gives a path on M. That is, Φ gives a tangent vector $D_{\Phi(t,p)} \in T_p(M)$. So, for $f \in C^\infty(M)$, by the definition of a derivative operator, we have

$$D_{\Phi(t,p)}(f) = \frac{d}{dt}(f \circ \Phi(t, p))|_{t=0}.$$

Now, p varies in the above equation. The right hand side is smooth by the smoothness of Φ. So, we conclude that $D_{\Phi(t,p)}$ is smooth with respect to p. That is, $D_{\Phi(t,p)} \in \mathfrak{X}(M)$, which shows that every global flow defines a vector field.

Example 2.1.3.6. *Let* $t, x \in \mathbb{R}$ *and* $\delta(t, x) = xe^t$*. Obviously,* $\delta(t, x)$ *is a flow in* $\mathbb{R}$*. The corresponding vector field is* $x\frac{\partial}{\partial x}$*.*

Proof. If $\xi \in \mathfrak{X}(\mathbb{R})$, then $\xi = f(x)\frac{\partial}{\partial x}$ where $f \in C^\infty(\mathbb{R})$. So, $f(p) = \xi_p(x)$. Let $g(x) = x$. We have

$$D_{\delta(t,x)}(g) = \frac{d}{dt}(g \circ \delta(t,x))|_{t=0} = \frac{d}{dt}(xe^t)|_{t=0} = x.$$

That is, the corresponding vector field is $x\frac{\partial}{\partial x}$. □

Example 2.1.3.7. *Let $s, x \in \mathbb{R}$ and $\tau(s,x) = x + s$. Obviously, $\tau(s,x)$ is a flow in $\mathbb{R}$. The corresponding vector field is $\frac{\partial}{\partial x}$.*

Proof. Let $g(x) = x$. We have

$$D_{\tau(s,x)}(g) = \frac{d}{ds}(g \circ \tau(s,x))|_{s=0} = \frac{d}{ds}(x+s)|_{s=0} = 1.$$

That is, the corresponding vector field is $\frac{\partial}{\partial x}$. □

Note that in general, for $A \in \mathfrak{gl}(n,\mathbb{R}) = \mathbf{M}_n(\mathbb{R})$, the corresponding linear flow is

$$\Phi(t,A) = \exp(tA).$$

Definition 2.1.3.3 (Abelian Lie algebras). *We say a Lie algebra $\mathfrak{g}$ is abelian if $[X,Y] = 0$ for all $X, Y \in \mathfrak{g}$.*

In the proof of Example 2.1.3.5, we show that $[-x\frac{\partial}{\partial x}, \frac{\partial}{\partial x}] = \frac{\partial}{\partial x}$. So, $\{(ax+b)\frac{\partial}{\partial x} \mid a, b \in \mathbb{R}\}$ is not an abelian Lie algebra. From Example 2.1.3.6 and 2.1.3.7, we know that $\delta(t,x) = xe^t$ induces the corresponding vector field is $x\frac{\partial}{\partial x}$ and $\tau(s,x) = x+s$ induces the corresponding vector field is $\frac{\partial}{\partial x}$. The reason why $[-x\frac{\partial}{\partial x}, \frac{\partial}{\partial x}] \neq 0$ is that the corresponding flows do not commute each other:

$$\delta_t \circ \tau_s \circ \delta_{-t} = \tau_{se^t}.$$

Exercise 2.1.3.3. *Show that $\{(ax^2+bx+c)\frac{\partial}{\partial x} \mid a, b, c \in \mathbb{R}\}$ forms a Lie subalgebra of $\mathfrak{X}(\mathbb{R})$ and it is isomorphic to $\mathfrak{sl}_2(\mathbb{R})$.*

Proof. We have to check Lie bracket operations of $\frac{\partial}{\partial x}$, $x\frac{\partial}{\partial x}$ and $x^2\frac{\partial}{\partial x}$: It is easy to see that

$$[x\frac{\partial}{\partial x}, \frac{\partial}{\partial x}] = -\frac{\partial}{\partial x} \text{ and } [\frac{\partial}{\partial x}, \frac{\partial}{\partial x}] = [x\frac{\partial}{\partial x}, x\frac{\partial}{\partial x}] = [x^2\frac{\partial}{\partial x}, x^2\frac{\partial}{\partial x}] = 0.$$

$$[x^2\frac{\partial}{\partial x}, x\frac{\partial}{\partial x}] = -x^2\frac{\partial}{\partial x} \text{ and } [x^2\frac{\partial}{\partial x}, \frac{\partial}{\partial x}] = -2x\frac{\partial}{\partial x}.$$

So, it is a Lie algebra generated by $\frac{\partial}{\partial x}$, $x\frac{\partial}{\partial x}$ and $x^2\frac{\partial}{\partial x}$. We want to remind you that

$$\mathfrak{sl}_2(\mathbb{R}) = \{\text{traceless elements of } \mathfrak{M}_2(\mathbb{R})\} = \{\begin{pmatrix} a & b \\ c & -a \end{pmatrix} \mid a, b, c, d \in \mathbb{R}\}.$$

Let

$$A_1 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \text{ and } A_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

It is easy to see that $[A_i, A_i] = 0$ for $i = 1, 2, 3$ and

$$[A_2, A_1] = -A_1, [A_3, A_2] = -A_3 \text{ and } [A_3, A_1] = -2A_2.$$

Define a map

$$A_1 \mapsto \frac{\partial}{\partial x}, A_2 \mapsto x\frac{\partial}{\partial x} \text{ and } A_3 \mapsto x^2\frac{\partial}{\partial x}, \text{ which is a Lie algebra isomorphism.}$$

□

Exercise 2.1.3.4. *Show that* $\{(a_n x^n + \cdots + a_1 x + a_0)\frac{\partial}{\partial x} \mid a_i \in \mathbb{R}$ *for* $i = 0, \ldots, n\}$ *does not form a Lie subalgebra of* $\mathfrak{X}(\mathbb{R})$ *if* $n > 2$.

Proof. It is easy to see that

$$[x^n\frac{\partial}{\partial x}, x^2\frac{\partial}{\partial x}] = (2-n)x^{n+1}\frac{\partial}{\partial x}.$$

So, if $n > 2$, then $(2-n)x^{n+1}\frac{\partial}{\partial x} \notin \{(a_n x^n + \cdots + a_1 x + a_0)\frac{\partial}{\partial x} \mid a_i \in \mathbb{R} \text{ for } i = 0, \ldots, n\}$. Hence, it is not a Lie algebra. □

2.1.4 Existence and Uniqueness of Integral Curves

We want to remind you that the definition of a global flow, i.e., Definition 2.1.3.2. In the notation of Definition 2.1.3.2, we say $\Phi(t, x)$ is a **local flow** if the domain is just $(-\delta, \epsilon) \times U$ where open $U \subseteq M$, not necessarily the whole manifold M. An integral curve γ on M is weaker concept of a local flow. That is, the domain of γ needs to be $(-\delta, \epsilon) \times \{p\}$ for $p \in M$, which is more or less equivalent to say that we don't need the smoothness of p variable. We give rigorous definitions:

Definition 2.1.4.1 (Integral curves). *Let* $X \in \mathfrak{X}(M)$. *We say* $\gamma : (-\delta, \epsilon) \to M$ *is an integral curve of* X *if it is smooth and for all* $f \in C^\infty(M)$ *and* $s \in (-\delta, \epsilon)$,

$$\frac{d}{dt}(f \circ \gamma(t))|_{t=s} = X_{\gamma(s)} f \overset{def}{=} \epsilon_{\gamma(s)}(Xf).$$

Definition 2.1.4.2 (Local Flows). *We say* $\Phi(t, x) = \Phi_t(x)$ *is a local flow on* M *if there exists an open* $U \subseteq M$ *such that* $\Phi : (-\delta, \epsilon) \times U \to M$ *is smooth and it satisfies that* Φ_0 *is an inclusion* $U \hookrightarrow M$ *and* $\Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2}$, *i.e.,* $\Phi(t_1 + t_2, x) = \Phi(t_1, \Phi(t_2, x))$ *whenever both sides of this equation are defined.*

Theorem 2.1.4.1. *Integral curves of a given vector field* X *always exist*

Proof. Let $X \in \mathfrak{X}(M)$ and M is n-dimensional manifold. So, we can write

$$X_p = \sum_{i=1}^{n} f_i(p)\frac{\partial}{\partial x_i} \text{ where } f_i \in C^\infty(M) \text{ and } p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}.$$

Take a chart $(x_1, \dots, x_n)$ around $p \in U \subseteq M$, i.e., $U \cong \mathbb{R}^n$, such that $(x_1(p), \dots, x_n(p)) = (0, \dots, 0)$. Now, we are in an Euclidean space. Let $\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))$ such that $\gamma(0) = (0, \dots, 0)$. Look at the following equation:

$$\gamma'(t) = \begin{pmatrix} \frac{d\gamma_i(t)}{dt} \\ \vdots \\ \frac{\gamma_n(t)}{dt} \end{pmatrix} = \begin{pmatrix} f_1(\gamma_1(t), \dots, \gamma_n(t)) \\ \vdots \\ f_n(\gamma_1(t), \dots, \gamma_n(t)) \end{pmatrix}.$$

The fundamental existence and uniqueness theorem of a system of first order differential equations tells us that we can have a unique solution subject to the initial condition $\gamma(0) = (0, \dots, 0)$. So, integral curves exist in each chart. □

Remark 2.1.4.1. *Note that each point $p \in M$ has a **maximal** interval $(a(p), b(p))$ in which $\gamma_p(t)$ is smooth. Also, what the uniqueness statement really tells us is that if $\gamma_p(t)$ with $\gamma_p(0) = p$ is an integral curve of a given vector field $X \in \mathfrak{X}(M)$ with domain $(a(p), b(p))$, then for all $s \in (a(p), b(p))$, we have*

$$\gamma_p'(s) = X_{\gamma_p(s)}.$$

That is, if we let $\gamma_{\gamma_p(s)}(t)$ be the integral curve of $X \in \mathfrak{X}(M)$ at $\gamma_p(s)$, then as long as $s, t, t + s \in (a(p), b(p))$, we have

$$\gamma_{\gamma_p(s)}(t) = \gamma_p(t + s).$$

Now we construct a local flow $\Phi(t, x)$ from integral curves of $X \in \mathfrak{X}(M)$. We know that there exists a unique integral curve $\gamma_p(t)$ at each point $p \in M$. Let

$$\Phi(t, p) = \gamma_p(t).$$

By Remark 2.1.4.1, we have for $t_1, t_2, t_1 + t_2 \in (a(p), b(p))$,

$$\Phi(t_1 + t_2, p) = \gamma_p(t_1 + t_2) = \gamma_{\gamma_p(t_2)}(t_1) = \Phi(t_1, \gamma_p(t_2)) = \Phi(t_1, \Phi(t_2, p)).$$

Moreover, by construction, $\Phi(0, p)$ is an **inclusion** from U to M.

Now, we will show smoothness: We know that $\Phi(t, p)$ is smooth at the variable t on the domain of $\gamma_p(t)$, i.e., $(a(p), b(p))$. However, what we do not know is whether or not $\Phi(t, p)$ is smooth at the variable p. To prove this we need a theorem from ordinary differential equations.

The fundamental existence and uniqueness theorem of a system of first order differential equations with varying the initial conditions $\Psi(0, b_1, \dots, b_n) = \gamma_{(b_1, \dots, b_n)}(0)$ tells us that there exists an open $(-\delta, \epsilon) \times U$ where $(b_1, \dots, b_n) \in U \subseteq \mathbb{R}^n$ such that we can have a unique solution $\Psi(t, x_1, \dots, x_n)$ subject to the initial condition $\Psi(0, b_1, \dots, b_n) = \gamma_{(b_1, \dots, b_n)}(0)$.

The uniqueness tells us that $\Phi(t, p) = \gamma_p(t)$ must be the solution. That is, we can find an open neighborhood $U_p \subseteq M$ of p and an open interval $(-\delta_p, \epsilon_p)$ such that on $(-\delta_p, \epsilon_p) \times U_p$, $\Phi(t, p) = \gamma_p(t)$ is **smooth**. Hence, we proved

Theorem 2.1.4.2. *For a given vector field $X \in \mathfrak{X}(M)$ and $m \in M$, there exist an open neighborhood U_m of m, $(-\delta_m, \delta_m) \subseteq \mathbb{R}$, and a local flow $\Phi_{U_m}(t, p)$ on $(-\delta_m, \delta_m) \times U_m$, which gives unique integral curves $\gamma_p(t) = \Phi(t, p)$ at each point $p \in U_m \subseteq M$.*

Remark 2.1.4.2. *Actually, we can have more. Let $(a(p), b(p))$ be the maximal interval in which $\gamma_p(t)$ is smooth.*

Let

$$W = \bigcup_{p\in M} (a(p), b(p)) \times \{p\}$$

$$\mathcal{D}_s \overset{def}{=} \{p \in M \mid (s,p) \in W\} = \{p \in M \mid s \in (a(p), b(p))\}$$

First thing we have to check is the following:

Theorem 2.1.4.3. *W is open in $\mathbb{R} \times M$ and $\mathcal{D}_s$ is open in W.*

Proof. Suppose that $(s,p) \in W$ for $s \in \mathbb{R}$ and $p \in M$. Without loss of generality, we can assume $s > 0$. Also, we can assume that for small $\epsilon > 0$, $(s+\epsilon, p) \in W$. The proof of the smoothness of a local flow shows that there exists an open set $(-\delta, \delta) \times U_p \subseteq W$ where U_p is an open neighborhood of p and $(-\delta,\delta) \subseteq (a(m), b(m))$ for all $m \in U_p$. Let $\Phi(t,m)$ be the local flow on $(-\delta,\delta) \times U_p$. Take some large enough $k \in \mathbb{N}$ such that $\frac{s+\epsilon}{k} \in (0,\delta)$ and for all $m \in U_p$,

$$\gamma_m(s+\epsilon) \overset{def}{=} \Phi(s+\epsilon, m) = \overbrace{\Phi_{\frac{s+\epsilon}{k}} \circ \cdots \circ \Phi_{\frac{s+\epsilon}{k}}}^{k}(m).$$

So, we have $(s,p) \in (0, s+\epsilon] \times U_p \subseteq W$. Hence, $(s,p) \in (0, s+\epsilon) \times U_p \subseteq W$. Thus, W is an open set.

Suppose $m \in \mathcal{D}_s$. So, $(s,m) \in W$. Since W is open, there exists an open set

$$\{t \mid |t-s| < \epsilon\} \times U \subseteq W \text{ where } m \in U \text{ is open in } M.$$

So, $\{s\} \times U \subseteq W$. So, $m \in U \subseteq \mathcal{D}_s$. ∎

Theorem 2.1.4.4. *For a given vector field $X \in \mathfrak{X}(M)$ and $m \in M$, there exist an open set $U \subseteq M$ and a local flow $\Phi(t,p)$ on $\bigcup_{p\in U}(a(p), b(p)) \times \{p\}$, which gives unique integral curves $\gamma_p(t) = \Phi(t,p)$ at each point $p \in U_m \subseteq M$.*

Proof. Note that the same proof of Theorem 2.1.4.3 shall show that

$$\bigcup_{p\in U} (a(p), b(p)) \times \{p\} \text{ is open.}$$

Suppose that $(s,p) \in W$ for $s \in \mathbb{R}$ and $p \in M$. Without loss of generality, we can assume $s > 0$. By Theorem 2.1.4.2, we can find the domains of local flows at each points $m \in \gamma_p([0,s])$. Let $(-\delta_m, \delta_m) \times U_m$ be the domain of $\Phi_{U_m}(t,x)$. Since $\{U_m\}$ is an open cover of $\gamma_p([0,s])$ and $\gamma_p([0,s])$ is compact, we can find a finite number of U_{m_i} for $i = 1, \ldots, k$. Let $a = \min_{1\leq i\leq k}\{\delta_{m_i}\}$. By the uniqueness of a local flow, $\Phi_{U_{m_i}}(t,x) = \Phi_{U_{m_j}}(t,x)$ on $(-a,a) \times (U_{m_i} \cap U_{m_j})$. Hence, we have a local flow

$$\Phi : (-a,a) \times \bigcup_{i=1}^{k} U_{m_j} \to M.$$

Take some large enough $k \in \mathbb{N}$ such that $\frac{s}{k} \in (0,a)$ and for all $m \in U_p$. Consider

$$\gamma_m(s) \overset{def}{=} \Phi(s,m) = \overbrace{\Phi_{\frac{s}{k}} \circ \cdots \circ \Phi_{\frac{s}{k}}}^{k}(m).$$

It is easy to see that for $i = 1, \ldots, k$,

$$\overbrace{\Phi_{\frac{s}{k}} \circ \cdots \circ \Phi_{\frac{s}{k}}}^{i}(m) \subseteq \gamma_p([0,s]) \subseteq \bigcup_{i=1}^{k} U_{m_j}.$$

That is, $\Phi(t,p)$ is smooth at s as well as p. Since $s \in (a(p), b(p))$ is arbitrary, we prove the theorem. $\square$

The proof of Theorem 2.1.4.4 has interesting consequences: Fix s and let A be the domain of Φ_s. If $p \in A$, then Φ_s is smooth at p. So, $s \in (a(p), b(p))$. Hence, $A \subseteq \mathcal{D}_s$. The proof of Theorem 2.1.4.4 tells us that $\mathcal{D}_s \subseteq A$. Hence, the domain of Φ_s is $\mathcal{D}_s$. From this, we have

Theorem 2.1.4.5. *$\Phi_s : \mathcal{D}_s \to \mathcal{D}_{-s}$ is a diffeomorphism with inverse Φ_{-s}. That is, every flow is a local diffeomorphism.*

Proof. Let $m \in \mathcal{D}_s$. By definition, $\Phi_s \circ \Phi_{-s}(m) = \Phi_0(m) = m$ and $\Phi_{-s} \circ \Phi_s(m) = \Phi_0(m) = m$. The proof of Theorem 2.1.4.4 shows that $\Phi_s : \mathcal{D}_s \to \mathcal{D}_{-s}$ is smooth. $\square$

Also, it is easy to see that

Theorem 2.1.4.6. *$\bigcup_{s>0} \mathcal{D}_s = M$ and* **Domain**$(\Phi_{s_1} \circ \Phi_{s_2}) \subseteq \mathcal{D}_{s_1+s_2}$.

Proof. Obviously, $\bigcup_{s>0} \mathcal{D}_s \subseteq M$. For $m \in M$, since there exists $0 \in (a(m), b(m))$ in which $\gamma_m(t)$ is smooth, there exists $s > 0$ such that $m \in \mathcal{D}_s$. So, $\bigcup_{s>0} \mathcal{D}_s = M$. If $p \in$ **Domain**$(\Phi_{s_1} \circ \Phi_{s_2})$, then $\Phi_{s_1} \circ \Phi_{s_2}(p) = \Phi_{s_1+s_2}(p)$. So, $p \in \mathcal{D}_{s_1+s_2}$. $\square$

Definition 2.1.4.3 (Complete vector fields). *We say $X \in \mathfrak{X}(M)$ is a complete vector field if it is generated by a global flow.*

Note that not every vector field is complete.

Example 2.1.4.1. *$\frac{\partial}{\partial x} \in \mathfrak{X}(\mathbb{R}/\{0\})$ is not complete.*

Proof. Suppose that $\Phi(t,x)$ is a global flow of $\frac{\partial}{\partial x}$. So, we have

$$\Phi(t,x) = \gamma_x(t).$$

Now, we solve the following differential equation:

$$\frac{d\gamma_p(t)}{dt} = 1 \text{ and } \gamma_p(0) = p.$$

It is easy to see that $\gamma_p(t) = t + p$. So,

$$\Phi(t,x) = t + x.$$

However, when $x = -t$, $t = 0$ implies $x = 0$. That is, $\Phi(t,x) = t + x$ is not defined if $t = -x$ on $\mathbb{R}/\{0\}$. Note that $\frac{\partial}{\partial x} \in \mathfrak{X}(\mathbb{R})$ is complete by the flow $\Phi(t,x) = t + x$. $\square$

Example 2.1.4.2. *$x^2 \frac{\partial}{\partial x} \in \mathfrak{X}(\mathbb{R})$ is not complete.*

Proof. Suppose that $\Phi(t,x)$ is a global flow of $x^2\frac{\partial}{\partial x}$. So, we have

$$\Phi(t,x) = \gamma_x(t).$$

Now, we solve the following differential equation:

$$\frac{d\gamma_p(t)}{dt} = \gamma_p^2(t) \text{ and } \gamma_p(0) = p \Longrightarrow \int_0^t dt = \int_0^t \frac{d\gamma_p(t)}{\gamma_p^2(t)}.$$

It is easy to see that $\gamma_p(t) = \frac{p}{1-tp}$. So,

$$\Phi(t,x) = \frac{x}{1-tx}.$$

However, when $t = \frac{1}{x}$, $\Phi(t,x) = \infty$. So, $\Phi(t,x)$ is not defined if $t = \frac{1}{x}$ on $\mathbb{R}$. □

Let (U_1, x) be a chart around the origin of $\mathbb{R} \cup \{\infty\}$ and (U_2, y) be a chart around ∞ of $\mathbb{R} \cup \{\infty\}$. Of course, as usual, the transition function is given by $x \mapsto \frac{1}{x} = y$. From this, we can smoothly extend $X = x^2\frac{\partial}{\partial x}$ to ∞. That is, we have a smooth vector field $X \in \mathfrak{X}(\mathbb{R} \cup \{\infty\})$, which is given by $x^2\frac{\partial}{\partial x}$ and $-\frac{\partial}{\partial y}$ on each chart. Now, this is complete. The global flow

$$\Xi : \mathbb{R} \times (\mathbb{R} \cup \{\infty\}) \to \mathbb{R} \cup \{\infty\}$$

is given by $\Phi_1(t,x) = \frac{x}{1-tx}$ or $\Psi_1(t,x) = \frac{1}{\Phi_1(t,x)}$ on $\mathbb{R} \times U_1$ and $\Phi_2(t,y) = \frac{1}{y-t}$ or $\Psi_2(t,x) = \frac{1}{\Phi_2(t,y)}$ on $\mathbb{R} \times U_2$. Since they agree on $U_1 \cap U_2$, the uniqueness shows that they are can be formed into a global flow Ξ. Actually, what gives the existence of Ξ is the compactness of $\mathbb{R} \cup \{\infty\}$. In general, we have

Exercise 2.1.4.1. *If M is compact, every vector field $X \in \mathfrak{X}(M)$ is complete. That is, each vector field gives a global flow $\xi_t \in \mathbf{Diff}(M)$ if M is compact.*

Proof. Let M be compact and $X \in \mathfrak{X}(M)$. By Theorem 2.1.4.2, we can find local flows, which generates X. Let $(-\delta_p, \epsilon_p) \times U_p$ be the domain of $\Phi(t,m)_p$. Since $\{U_p\}$ is an open cover and M is compact, we can find a finite number of flows $\Phi(t,m)_{p_i}$ for $i = 1, \ldots, k$. Let $a = \min_{1 \le i \le k}\{\delta_{p_i}, \epsilon_{p_i}\}$. By the uniqueness of a local flow, $\Phi(t,m)_{p_i} = \Phi(t,m)_{p_j}$ on $(-a,a) \times (U_{p_i} \cap U_{p_j})$. Hence, we have a local flow

$$\Phi : (-a,a) \times M \to M \text{ with } \Phi(0,m) = id_M.$$

Let $r \in \mathbb{R}$. There exist unique $n \in \mathbb{Z}$ and $|b| < \frac{a}{2}$ such that $r = n\frac{a}{2} + b$. Define

$$\Phi(r,m) = \begin{cases} \overbrace{\Phi_{\frac{a}{2}} \circ \cdots \circ \Phi_{\frac{a}{2}}}^{n} \circ \Phi_b & \text{if } r \ge 0 \\ \underbrace{\Phi_{-\frac{a}{2}} \circ \cdots \circ \Phi_{-\frac{a}{2}}}_{n} \circ \Phi_b & \text{if } r < 0. \end{cases}$$

The chain rule gives that $\Phi : \mathbb{R} \times M \to M$ is smooth. Also, by construction, $\Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2}$. Hence, it is a global flow which generates $X \in \mathfrak{X}(M)$. □

Theorem 2.1.4.7 (Normal forms for nonsingular vector fields). *Let M be an n-dimensional smooth manifold and $p \in M$. Suppose that $X \in \mathfrak{X}(M)$ and $X_p \neq 0$. Then there exists a smooth coordinate chart (U, φ) at p where*

$$X|_U = \varphi * (\frac{\partial}{\partial y})|_U.$$

Proof. We know that there exists a local flow ξ_t around a small neighborhood $V \subseteq M$ of p, which generates X on V. That is, $\xi : (-\epsilon, \epsilon) \times V \to M$ where

$$X_{(x_1,\ldots,x_n)} = \frac{d}{dt}\xi_t(x_1, \ldots, x_n)|_{t=0}.$$

Since $X_p \neq 0$, without loss of generality, we can shrink V smaller so that we can choose a coordinate system on $(V, \varphi = (x_1, \ldots, x_n))$ such that $p = (0, \ldots, 0)$ and

$$T_p(V) \text{ has a basis } \{X_p, \frac{\partial}{\partial x_2}|_p, \ldots, \frac{\partial}{\partial x_n}|_p\}, \text{ i.e., } X_p = \frac{\partial}{\partial x_1}|_p.$$

Let

$$W = \{(x_1, \ldots, x_n) \in V \mid x_1 = 0\}.$$

Note that W is nothing but a smooth hypersurface in V. So, $\dim(W) = n-1$. Define on $(-\epsilon, \epsilon)\times W$,

$$F(y, x_2, \ldots, x_n) = \xi_y(0, x_2, \ldots, x_n).$$

Note that $F(0, 0, \ldots, x_i, 0, \ldots, 0) = \xi_0(0, \ldots, x_i, 0, \ldots, 0) = (0, \ldots, x_i, 0, \ldots, 0)$ for $i = 2, \ldots, n$. Clearly, F is smooth on $(-\epsilon, \epsilon) \times W$. So, we have

$$dF_p : T_p((-\epsilon, \epsilon) \times W) \to T_p(M) = T_p(V).$$

Since $T_P((-\epsilon, \epsilon) \times W) = T_p((-\epsilon, \epsilon)) \times T_p(W)$, $T_p((-\epsilon, \epsilon) \times W)$ has a basis $\{\frac{\partial}{\partial y}|_p, \frac{\partial}{\partial x_2}|_p, \ldots, \frac{\partial}{\partial x_n}|_p\}$. It is easy to see that

$$dF_p(\frac{\partial}{\partial x_i}|_p) = \sum_{j=1}^{n} \frac{\partial F_j}{\partial x_i}|_p \frac{\partial}{\partial x_j}|_p = \frac{\partial}{\partial x_i}|_p \text{ for } i = 2, \ldots, n.$$

Moreover, by the definition of a local flow ξ_t, we have

$$dF_p(\frac{\partial}{\partial y}|_p) = \sum_{j=1}^{n} \frac{\partial F_j}{\partial y}|_p \frac{\partial}{\partial x_j}|_p = X_p.$$

We note that

$$\begin{aligned} X_{\xi_s(x_1,\ldots,x_n)} &= \frac{d}{dt}\xi_t(\xi_s(x_1, \ldots, x_n))|_{t=0} \\ &= \frac{d}{dt}\xi_{t+s}(x_1, \ldots, x_n)|_{t=0} = \frac{d}{dt}\xi_t(x_1, \ldots, x_n)|_{t=s} \end{aligned}$$

From this and $F(y, x_2, \ldots, x_n) = \xi_y(0, x_2, \ldots, x_n)$, we have

$$\begin{aligned} \frac{\partial F}{\partial y}|_{(a,a_2,\ldots,a_n)} &= \frac{\partial \xi_y(0, x_2, \ldots, x_n)}{\partial y}|_{(a,a_2,\ldots,a_n)} \\ &= \frac{d\xi_y(0, a_2, \ldots, a_n)}{dy}|_{y=a} = \frac{d}{dt}\xi_t(\xi_a(0, a_2, \ldots, a_n)|_{t=0} \\ &= X_{\xi_a(0,a_2,\ldots,a_n)} = X_{F(a,a_2,\ldots,a_n)}. \end{aligned}$$

That is,

$$dF_{(y,x_2,\ldots,x_n)}(\frac{\partial}{\partial y}|_{(y,x_2,\ldots,x_n)}) = X_{F(y,x_2,\ldots,x_n)}.$$

Since the differential of F is nonsingular at $\{0\} \times \overbrace{(0,\ldots,0)}^{n-1}$, then we can use $\{y, x_1, \ldots, x_n\}$ as coordinates from the inverse function theorem, if necessary, by shrinking $(-\epsilon,\epsilon) \times W$. Since $F : (-\epsilon,\epsilon) \times W \to M$, there exists an open set $(-\epsilon,\epsilon) \times W \overset{F}{\cong} U \subseteq M$ such that

$$\{y \circ F^{-1}, x_2 \circ F^{-1}, \ldots, x_n \circ F^{-1}\} \text{ is a smooth coordinate chart on } U \text{ and } X|_U = dF(\frac{\partial}{\partial y})|_U.$$

□

2.1.5 Left actions of Diff(M) on $C^\infty(M)$.

Let $\varphi \in \mathbf{Diff}(M)$ and $f \in C^\infty(M)$. Define

$$(\varphi \cdot f) \overset{def}{=} f \circ \varphi^{-1}.$$

It is easy to see to see this is an action, i.e.,

$$\psi \cdot (\varphi \cdot f) = (\psi \circ \varphi) \cdot f.$$

Example 2.1.5.1. *Let $X \in \mathfrak{X}(M)$ be complete and ξ_t be the corresponding global flow. We already have shown that $\xi_t \in \mathbf{Diff}(M)$ and the inverse of ξ_t is ξ_{-t}. So, we have*

$$\begin{aligned} Xf &= \frac{d}{dt}|_{t=0} f(\xi_t(p)) = \frac{d}{dt}|_{t=0}((\xi_t)^{-1} \cdot f) \\ &= \frac{d}{dt}|_{t=0}(\xi_{-t} \cdot f) = -\frac{d}{dt}|_{t=0}(\xi_t \cdot f). \end{aligned}$$

Now, we are about to define a left action of $\mathbf{Diff}(M)$ on $\mathfrak{X}(M) \overset{def}{=} \mathbf{Vect}(M)$. Note that this is a prototype of general theory that a Lie group (in this case, $\mathbf{Diff}(M)$) acts on the Lie algebra (in this case, $\mathfrak{X}(M)$). We want to remind that for $\varphi \in \mathbf{Diff}(M)$ we have

$$\begin{array}{ccc} TM & \xrightarrow{d\varphi} & TM \\ \downarrow & & \downarrow \\ M & \xrightarrow{\varphi} & M. \end{array}$$

So, for $X \in \mathfrak{X}(M)$,

$$d\varphi(X) \in \mathfrak{X}(M) \text{ and } d\varphi_{\varphi^{-1}(p)}(X_{\varphi^{-1}(p)}) \overset{def}{=} (d\varphi(X))_p \in T_p(M).$$

Note that

$$X_p(f) \overset{def}{=} (Xf)(p).$$

Note that we define a vector field to be a derivation. So, by definition, we have

$$(d\varphi_{\varphi^{-1}(p)}(X_{\varphi^{-1}(p)}))(f) \overset{def}{=} X_{\varphi^{-1}(p)}(f \circ \varphi) \overset{def}{=} (X(f \circ \varphi))(\varphi^{-1}(p)).$$

Of course, if you use the Jacobian of $J\varphi$, then you will get the same answer.

Remark 2.1.5.1. *Obviously, the above* ***does not*** *imply*

$$d\varphi_{\varphi^{-1}(p)}(X_{\varphi^{-1}(p)})(f) = X_{\varphi^{-1}(p)}(f \circ \varphi) = X_p(f), \text{ i.e., } (d\varphi(X))_p = X_p.$$

However, in some situation there exists $\varphi \in \mathbf{Diff}(M)$*, which makes*

$$d\varphi_{\varphi^{-1}(p)}(X_{\varphi^{-1}(p)})(f) = (d\varphi(X))_p = X_p \text{ for all } p \in M.$$

Later, those will be our one of main objects to study.

Let $\varphi \in \mathbf{Diff}(M)$ and $X \in \mathfrak{X}(M)$. We define

$$\varphi \cdot X \overset{def}{=} d\varphi(X).$$

That is,

$$(\varphi \cdot X)_p = d\varphi_{\varphi^{-1}(p)}(X_{\varphi^{-1}(p)}) \overset{def}{=} d\varphi \circ X \circ \varphi^{-1}.$$

It is an action, since

$$\psi \cdot (\varphi \cdot X) = d\psi(d\varphi(X)) = d(\psi \circ \varphi)(X) = (\psi \circ \varphi) \cdot X.$$

This action has naturality: For all $f \in C^\infty(M)$,

$$\begin{aligned}((\varphi \cdot X)(\varphi \cdot f))(p) &= (\varphi \cdot X)_p(\varphi \cdot f) = (d\varphi(X))_p(\varphi \cdot f) \\ &= (d\varphi)_{\varphi^{-1}(p)}(X_{\varphi^{-1}(p)})(\varphi \cdot f) = X_{\varphi^{-1}(p)}((\varphi \cdot f) \circ \varphi) \\ &= X_{\varphi^{-1}(p)}(f) = (Xf)(\varphi^{-1}(p)) \\ &= (\varphi \cdot (Xf))(p)\end{aligned}.$$

That is, $(\varphi \cdot X)(\varphi \cdot f) = \varphi \cdot (Xf)$. Observe that if $\varphi \in \mathbf{Diff}(M)$, then it is easy to see that $\varphi : Der(C^\infty(M)) \times C^\infty(M) \to C^\infty(M)$ defines a bilinear map where $Der(C^\infty(M)) = \mathfrak{X}(M)$, i.e.,

$$\varphi(Y, f) = (\varphi \cdot Y)(\varphi \cdot f) = \varphi \cdot (Yf).$$

Remark 2.1.5.2. *The action of* $\mathbf{Diff}(M)$ *on* $\mathfrak{X}(M)$ *does not extend the group of smooth maps between* M *to* M*, i.e.,* $C^\infty(M, M)$*, since we need invertibility. However, in the case of differential forms, a pullback is well-defined. Now, we will show this point of view.*

We want to remind you that a covector is an element of $T_p^*(M) = \mathbf{Hom}(T_P(M), \mathbb{R})$. A smooth section of $T^*(M) = (T(M))^*$ is called a covector field or 1-form. Every $f \in C^\infty(M)$ defines a 1-form: Let $f : M \to \mathbb{R}$ be smooth. So, we have

$$T_p(M) \xrightarrow{df_p} T_{f(p)}(\mathbb{R}) \cong \mathbb{R}.$$

Hence, $df_p \in \mathbf{Hom}(T_P(M), \mathbb{R}) = T_p^*(M)$. Another way of saying this situation is the following: d defines a homomorphism between $C^\infty(M)$ and the space of 1-forms. That is,

$$C^\infty(M) \xrightarrow{d} \Omega^1(M) = \{ \text{ sections of } T^*(M)\} \text{ i.e., } f \mapsto df.$$

Let $\varphi : M \to N$ be an arbitrary smooth map. We have $d\varphi : T(M) \to T(N)$. Since $T^*(M) = \mathbf{Hom}(T(M), \mathbb{R})$, $T^*(N) = \mathbf{Hom}(T(N), \mathbb{R})$, and $d\varphi$ is nothing but the Jacobian $J\varphi$ of φ, using the linearity of $d\varphi$, it is easy to see that

$$(d\varphi)^T : T^*(N) \to T^*(M).$$

Let $\omega \in \Omega^1(N)$ and $p \in M$. Define

$$(\varphi^*\omega)_p \overset{def}{=} (d\varphi)_p^T(\omega_{\varphi(p)}) \overset{def}{=} (d\varphi)^T \circ \omega \circ \varphi.$$

Note that for finite dimensional vector spaces V, W and the duals V^*, W^*, any linear map f gives the following commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{\cong} & V^* \\ f\big\downarrow & & \big\downarrow f^T \\ W & \xrightarrow{\cong} & W^*. \end{array}$$

So, it is easy to see that

$$(\varphi^*\omega)_p(X_p) = (d\varphi)_p^T(\omega_{\varphi(p)})(X_p) = \omega_{\varphi(p)}(d\varphi_p(X_p)) = \omega_{\varphi(p)}((d\varphi(X))_{\varphi(p)}).$$

Since the right-hand term is a smooth function as p varies, $\varphi^*\omega$ is smooth as p varies. Hence, we have

$$\varphi^* : \Omega^1(N) \to \Omega^1(M).$$

Example 2.1.5.2. *Let $M = \mathbb{R}^2$ and $X = \frac{\partial}{\partial x}, Y = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$. Clearly, $X, Y \in \mathfrak{X}(\mathbb{R}^2)$. It is easy to see that the corresponding global flow of X is $\xi_t(x, y) = (x + t, y)$ and the corresponding global flow of Y is $\eta_t(x, y) = (xe^t, ye^t)$. We have*

$$\frac{d}{dt}|_{t=0}(\xi_t \cdot Y) = -[X, Y].$$

Proof. For all $(a, b) \in \mathbb{R}^2$,

$$d\xi_t(a, b) = J\xi_t = \begin{pmatrix} \frac{\partial x+t}{\partial x} & \frac{\partial x+t}{\partial y} \\ \frac{\partial y}{\partial x} & \frac{\partial y}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that

$$d\xi_t : T_{(a,b)}(\mathbb{R}^2) \to T_{\xi_t(a,b)}(\mathbb{R}^2) = T_{(a+t,y)}(\mathbb{R}^2).$$

Also, we have

$$Y_{(a,b)} = a\frac{\partial}{\partial y} + b\frac{\partial}{\partial y} \text{ and } Y_{\xi_{-t}(a,b)} = Y_{(a-t,b)} = (a - t)\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}.$$

By definition

$$\begin{aligned} \xi_t \cdot Y &= d\xi_t \circ Y \circ \xi_{-t} = (d\xi_t)_{\xi_{-t}}(Y_{\xi_{-t}}) \\ &= \mathbf{I}_2 \cdot \begin{pmatrix} a - t \\ b \end{pmatrix} = (a - t, b) \end{aligned}.$$

That is, $\xi_t \cdot Y = (a-t)\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} \in T_{(a+t,y)}(\mathbb{R}^2)$. So,

$$\frac{d}{dt}|_{t=0}(\xi_t \cdot Y) = \frac{d}{dt}|_{t=0}(x-t)\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} = -\frac{\partial}{\partial x}.$$

Since it is easy to see that $-[x,y] = -\frac{\partial}{\partial x}$, we have

$$\frac{d}{dt}|_{t=0}(\xi_t \cdot Y) = -[X,Y].$$

□

Again, this is no coincidences. This will be the content of the next section.

2.1.6 Lie Derivatives

Let M be a smooth manifold and $X, Y \in \mathfrak{X}(M)$. Suppose that $\xi_t \in \mathbf{Diff}(M)$ be a global flow of X. Note that $\xi_t \cdot Y \in \mathfrak{X}(M)$ is a path of vector fields. Define the Lie derivative of Y by X to be

$$\mathcal{L}_X(Y) \overset{def}{=} \frac{d}{dt}|_{t=0}(\xi_{-t} \cdot Y).$$

Theorem 2.1.6.1.

$$\mathcal{L}_X(Y) = [X,Y].$$

Proof. Actually, X does not need to be complete even though we will assume that it is complete. That is, it is sufficient that ξ_t is a local flow for our proof. Notice that all the actions which we have defined so far in this section still make sense if we apply them to appropriate domains.

Let ξ_t be the global flow of $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$. Note that $\xi_0 \in \mathbf{Diff}(M)$ is the identity map and

$$\begin{aligned}(\mathcal{L}_X f)(p) &= -\big(\frac{d}{dt}|_{t=0}(\xi_t \cdot f)\big)(p) = -\frac{d}{dt}|_{t=0}(f \circ \xi_{-t}(p)) \\ &= \frac{d}{dt}|_{t=0}(f \circ \xi_t(p)) = X_p f.\end{aligned}$$

Note that from the above $Xf = -\big(\frac{d}{dt}|_{t=0}(\xi_t \cdot f)\big)$ and by $\mathbb{R}$-bilinearity of actions, we have $t(Y(f)) = Y(tf)$ for $t \in \mathbb{R}$. So, using bilinearity of the actions, we have

$$\begin{aligned}X(Yf) &= -\frac{d}{dt}|_{t=0}(\xi_t \cdot (Yf)) = -\frac{d}{dt}|_{t=0}(\xi_t \cdot Y)(\xi_t \cdot f) = -\lim_{t\to 0}\frac{(\xi_t \cdot Y)(\xi_t \cdot f) - Yf}{t} \\ &= -\lim_{t\to 0}\frac{(\xi_t \cdot Y)(\xi_t \cdot f) - (\xi_t \cdot Y)(f) + (\xi_t \cdot Y)(f) - Yf}{t} \\ &= -\lim_{t\to 0}\frac{1}{t}(\xi_t \cdot Y)((\xi_t \cdot f) - f) + \lim_{t\to 0}\frac{(\xi_t \cdot Y - Y)(f)}{t} \\ &= -\lim_{t\to 0}(\xi_t \cdot Y)(\frac{(\xi_t \cdot f) - f}{t}) + \lim_{t\to 0}\frac{(\xi_t \cdot Y - Y)(f)}{t} \\ &= -\big((\xi_t \cdot Y)|_{t=0}\big)\frac{d}{dt}|_{t=0}(\xi_t \cdot f) - \frac{d}{dt}|_{t=0}(\xi_t \cdot Y)\big((\xi_t \cdot f)|_{t=0}\big) \\ &= Y(Xf) + \mathcal{L}_X Y(f).\end{aligned}$$

Hence,

$$\mathcal{L}_X(Y)(f) = X(Yf) - Y(Xf) = [X,Y](f)$$

$\square$

Corollary 2.1.6.1.1.

$$\frac{d}{dt}|_{t=0}(\xi_{-t}\cdot Y) = [X,Y].$$

What does it mean $[X,Y]=0$ for $X,Y\in\mathfrak{X}(M)$? One obvious answer would be $X(Y(f)) = Y(X(f))$ for all $f\in C^\infty(M)$. That is, second order differential operators , i.e., XY and YX are the same. Corollary 2.1.6.1.1 has another interesting consequence: It says that if $[X,Y]=0$, then

$$\frac{d}{dt}|_{t=0}(\xi_{-t}\cdot Y) = 0.$$

In particular, since $[X,X]=0$, $\frac{d}{dt}|_{t=0}(\xi_{-t}\cdot X)=0$. We note that

$$\begin{aligned}
\frac{d}{dt}\xi_{-t}\cdot Y|_{t=s} &= \frac{d}{dt}\xi_{-t-s}\cdot Y|_{t=0} = \frac{d}{dt}(\xi_{-t}\circ\xi_{-s})\cdot Y|_{t=0}\\
&= \frac{d}{dt}(\xi_{-s}\circ\xi_{-t})\cdot Y|_{t=0} = \frac{d}{dt}(\xi_{-s})\cdot(\xi_{-t}\cdot Y)|_{t=0}\\
&= \frac{d}{dt}(\xi_{-s})|_{t=0}\cdot(\xi_{-t}\cdot Y)|_{t=0} + (\xi_{-s}|_{t=0})\cdot(\frac{d}{dt}\xi_{-t}\cdot Y|_{t=0})\\
&= (\xi_{-s})\cdot(\frac{d}{dt}\xi_{-t}\cdot Y|_{t=0}) = 0.
\end{aligned}$$

One important conclusion of the above is the following: If $[X,Y]=0$, then since $\frac{d}{dt}\xi_{-t}\cdot Y|_{t=s}=0$ for all $s\in\mathbb{R}$, we have

$$Y = \xi_0\cdot Y = \xi_s\cdot Y.$$

Now, we investigate the flow of $\xi_t\cdot Y$.

Exercise 2.1.6.1. *Let $\varphi\in\mathbf{Diff}(M)$ and $X\in\mathfrak{X}(M)$ be complete. Show that $\varphi\cdot X$ is also complete and the global flow is given by $\varphi\circ\eta_t\circ\varphi^{-1}$ where η_t is the global flow of X.*

Proof. Let $p\in M$ and $\varphi^{-1}(q)=p$. We have

$$(\varphi\cdot X)_q = d\varphi_{\varphi^{-1}(q)}(X_{\varphi^{-1}(q)}) = \frac{d}{dt}(\varphi\circ\eta_t(\varphi^{-1}(q))|_{t=0}.$$

So, let $\psi_t(q)=\varphi\circ\eta_t\circ\varphi^{-1}(q)$. Since $\psi_t\in\mathbf{Diff}(M)$ and ψ_t generates $\varphi\cdot X$ by the above, the uniqueness of a global flow says that ψ_t is the global flow of $\varphi\cdot X$. $\square$

Let η_s be the global flow of Y and ξ_t be a global flow of X. If $[X,Y]=0$, then $Y=\xi_t\cdot Y$. So, by Exercise 2.1.6.1, we have

$$\eta_s = \xi_t\circ\eta_s\circ\xi_{-t} \Longleftrightarrow \eta_s\circ\xi_t = \xi_t\circ\eta_s.$$

Hence, we have

Theorem 2.1.6.2. *Vector fields commute each other if and only if their global flows commute each other.*

2.1.7 Remarks on vector bundles

We do not need this section later. So, you can skip this if you wish.
An n-dimensional real smooth vector bundle over smooth manifold M^m is a smooth map $E^{n+m} \xrightarrow{\pi} M^m$ satisfying the local triviality and each fiber $\pi^{-1}(p)$ is an n-dimensional real vector space. Note that a structure group is $\mathbf{GL}(n, \mathbb{R})$. Let

$$\Gamma(E) = \{C^\infty - \text{sections of } \pi, \text{ i.e., } \pi \circ s = id_M\}.$$

Example 2.1.7.1. $\Gamma(T(M)) = \mathfrak{X}(M)$ *and* $\Gamma(T^*(M)) = \Omega^1(M)$.

One of a nice classification theorem of vector bundles is the following:

Theorem 2.1.7.1. *There is one-to-one correspondence between smooth vector bundles of finite rank over a smooth manifold M and finitely generated projective $C^\infty(M)$-modules.*

Sketch of proof. It is a well-know theorem in Algebraic Geometry that there is one-to-one correspondence between smooth vector bundles of finite rank n over a smooth manifold M and finitely generated locally free $C^\infty(M)$-modules. Here, finitely generated locally free $C^\infty(M)$-modules $\mathcal{A}$ means that there exists an open neighborhood $U_p \subseteq M$ such that $\mathcal{A}(U_p) \cong (C^\infty(U_p))^n$ for each $p \in M$. Note that we should have said locally free sheaves over a structure sheaf. A sheaf has more structure than a module. However, since we can construct a unique global object from compatible local data in the case of C^∞, $\mathcal{A}$ are already sheave. Since locally free $C^\infty(U_p)$-modules is locally projective $C^\infty(U_p)$-modules and $C^\infty(M)$-modules is sheave, we can construct a global splitting g from each local splitting, so it is projective $C^\infty(M)$-modules:

$$0 \to \mathcal{C}(M) \xrightarrow{h} \mathcal{B}(M) \underset{g}{\overset{f}{\rightleftarrows}} \mathcal{A}(M) \to 0.$$

If $\mathcal{A}(M)$ is a finitely generated projective $C^\infty(M)$-module, then $\mathcal{A}(U_p)$ is a finitely generated projective $C^\infty(U_p)$-module. Since $C^\infty(U_p)$ is a **local ring** as a stalk, $\mathcal{A}(U_p)$ is a finitely generated free $C^\infty(U_p)$-module. □

From this we have:
$\Gamma(E)$ is a finitely generated free $C^\infty(M)$-module if and only if there exist sections $s_1, \ldots, s_n \in \Gamma(E)$ such that every section $s \in \Gamma(E)$ is a linear combination $s = \sum_{i=1}^n f_i s_i$ where $f_i \in C^\infty(M)$ if and only if there are n sections which are everywhere linearly independent if and only if E is a trivial vector bundle. Note that not every vector bundle is trivial.

Example 2.1.7.2. *Since every section of $T(S^2)$ must vanish somewhere, $\mathfrak{X}(S^2)$ is not a free $C^\infty(S^2)$-module.*

Note that there exists a normal bundle ν of S^2 such that $T(S^2) \oplus \nu$ is a trivial bundle. In a compact manifold, this phenomenon is always true. Before proving this, we give a definition of the Whitney sum: If $E_1^{n_1}$ and $E_2^{n_2}$ are vector bundles over M, we can make Whitney sum $E_1^{n_1} \oplus E_2^{n_2}$ over M of $E_1^{n_1}$ and $E_2^{n_2}$, which is a vector bundle over M whose fiber $(E_1^{n_1} \oplus E_2^{n_2})_p$ is $E_{1p}^{n_1} \oplus E_{2p}^{n_2}$ and structure group is given by

$$\begin{pmatrix} g^1_{\alpha\beta} & 0 \\ 0 & g^2_{\alpha\beta} \end{pmatrix} \text{ where } g^1_{\alpha\beta}, g^2_{\alpha\beta} \text{ are in the structure groups of } E_1^{n_1} \text{ and } E_2^{n_2}, \text{ respectively.}$$

By the virtue of Theorem 2.1.7.1, we can give two proofs of the following:

Exercise 2.1.7.1. *Let M^m be a compact Hausdorff smooth manifold and ξ be a vector bundle over M^m. Prove that ξ sits inside a trivial bundle over $\mathbb{R}^n$ for some $n \geq m$.*

Proof 1. Regarding M^m as the zero section of $E(\xi)$, one can embed some neighborhood $U \subset E(\xi)$ of M^m in some $\mathbb{R}^n$ by the Whitney embedding theorem. Using the local triviality of vector bundle, we have a trivial bundle $\tau_U \oplus \nu_U$ over U where the tangent bundle τ_U and normal bundle ν_U. Since ξ projects non trivially to the normal space to $T_x(M)$ in $T_x(U)$ using the differential of the embedding, this induces an isomorphism of ξ with the normal bundle λ to M in U. But $\lambda \oplus \nu_U|_M \simeq \nu_M$ and so $\xi \oplus \nu_U|_M \oplus \tau_M \simeq \lambda \oplus \nu_U|_M \oplus \tau_M \simeq \nu_M \oplus \tau_M$ which is a trivial bundle. □

Proof 2. It is a well-know theorem in Algebra that a projective module is a direct summand of a free module. By Theorem 2.1.7.1 and the Whitney sum, there exists a vector bundle E' such that

$$\Gamma(E) \oplus \Gamma(E') \text{ is free.}$$

□

2.1.8 Frobenius's Theorem

Definition 2.1.8.1 (k-plane distributions or k-plane fields). *An k-plane field E over M is a subbundle of $T(M)$, the tangent bundle of M. That is*

$$\Gamma(E) \subseteq \Gamma(T(M)) = \mathfrak{X}(M).$$

Definition 2.1.8.2 (Integral submanifold). *Let E be a k-plane field over M. We say a manifold of S is an integral submanifold of E if $S \xrightarrow{f} M$ is an injective immersion and for all $s \in S$,*

$$d_{f(s)}(T_s(S)) = E_{f(s)}.$$

Note that we do not need f to be an embedding. That is, $f(S)$ need not be homeomorphic to S.

Definition 2.1.8.3 (Integrability of E). *Let E be a k-plane field over M. We say E is integrable if each $p \in M$, there exists an integral submanifold of E through p.*

Suppose that E is an integrable k-plane field. We can find an integral submanifold at each point $p \in M$. From these submanifold, in an obvious way, we can make the maximal connected integral submanifolds to E.

Definition 2.1.8.4 (Foliation and Leaf). *The decomposition of M into the maximal connected submanifolds to E is called a foliation $\mathcal{F}$ of M. Each maximal connected integral submanifold to E is called a leaf of the foliation $\mathcal{F}$. Plainly speaking, a foliation is the family of leaves.*

Example 2.1.8.1. *Let G be a torus $\mathbb{R}^2/\mathbb{Z}^2$, which is a Lie group. Define $f : \mathbb{R} \to G$ by*

$$t \mapsto (t, \alpha t) \text{ mod } \mathbb{Z}^2.$$

It is easy to see that if α is irrational, then f is injective immersion and $\mathbf{Im}(f)$ *is dense in G. However, since G is compact and $\mathbb{R}$ is not, we conclude that* $\mathbf{Im}(f)$ *is not homeomorphic to $\mathbb{R}$ and f is not proper. Note that* $\mathbf{Im}(f)$ *is a submanifold of G. Usually, we call* $\mathbf{Im}(f)$ *is the skew line on the torus. By taking different irrational number α, we can foliate G by* $\mathbf{Im}(f_\alpha)$*, which are leaves. That is, for $(x, y) \in G$, a leaf is given by $(x, y) + \mathbb{R}(\alpha, 1)$ mod $\mathbb{Z}^2$, i.e., $f_\alpha(\mathbb{R})$.*

This is a general phenomenon: In general, if $f : H \to G$ is an injective Lie group homomorphism, then the cosets of H foliates G. Note that G/H need not be a Lie group.

Example 2.1.8.2. *Every line field (1-plane field) is integrable.*

Proof. Let E be a line field of M. Locally, there exists a vector field $v_p \in \mathfrak{X}(U_p)$ such that v_p spans E_p where $U_p \subseteq M$ is an open neighborhood of p. Since locally there always exists an integral curve generating v_p, which is obviously a submanifold, we conclude that E is integrable. $\square$

Example 2.1.8.3. *We know that every line field E is integrable. Since $\Gamma(E)$ is generated by a single vector field $X \in \mathfrak{X}(M)$, we have*

$$\Gamma(E) = C^\infty(M) \cdot X.$$

Since $[f_1X, f_2X] = (f_1(Xf_2) - f_2(Xf_1))X$ and $f_1(Xf_2) - f_2(Xf_1) \in C^\infty(M)$ for all $f_1, f_2 \in C^\infty(M)$, we conclude that $\Gamma(E)$ is a Lie subalgebra.

This is the general theorem, which we shall give. First we need the followings:

Definition 2.1.8.5. *Let $M \xrightarrow{\varphi} N$ be a smooth map and $X \in \mathfrak{X}(M), Y \in \mathfrak{X}(N)$. We say X and Y are φ-related if for all $p \in M$,*

$$d\varphi(X_p) = Y_{\varphi(p)}.$$

Example 2.1.8.4. *Let $\varphi \in \mathbf{Diff}(M)$ and $X \in \mathfrak{X}(M)$. Then X and $\varphi \cdot X$ are φ-related.*

Exercise 2.1.8.1. *If X^i is φ-related to Y^i for $i = 1, 2$, then $[X^1, X^2]$ is φ-related to $[Y^1, Y^2]$.*

Proof. Let $f \in C^\infty(N)$ and $p \in M$. Note that

$$(X^i(f \circ \varphi))(p) = Y^i_{\varphi(p)} f = (Y^i f)(\varphi(p)) = (Y^i f) \circ \varphi(p).$$

So, we have

$$\begin{aligned} d\varphi([X^1, X^2]_p)(f) &= [X^1, X^2]_p(f \circ \varphi) = X^1_p(X^2(f \circ \varphi)) - X^2_p(X^1(f \circ \varphi)) \\ &= X^1_p((Y^2 f) \circ \varphi) - X^2_p((Y^1 f) \circ \varphi) = Y^1_{\varphi(p)}(Y^2 f) - Y^2_{\varphi(p)}(Y^1 f) \\ &= [Y^1, Y^2]_{\varphi(p)}(f). \end{aligned}$$

$\square$

Theorem 2.1.8.1. *If E is integrable, then $\Gamma(E)$ is a Lie subalgebra of $\mathfrak{X}(M)$.*

Proof. Suppose that $X^1, X^2 \in \Gamma(E)$. We want to show $[X^1, X^2] \in \Gamma(E)$, equivalently, $[X^1, X^2]_p \in E_p$ for all $p \in M$. Let S be an integral submanifold through p, i.e.,

$$S \overset{\varphi}{\hookrightarrow} M.$$

If $X^1, X^2 \in \Gamma(E)$, then there exists $Y^1, Y^2 \in \mathfrak{X}(S)$ such that $d\varphi(Y^i) = X^i \circ \varphi$ for $i = 1, 2$. That is, Y^i is φ-related to X^i. Since $[Y^1, Y^2] \in \mathfrak{X}(S)$ by the fact that $\mathfrak{X}(S)$ is a Lie algebra and $[Y^1, Y^2]$ is φ-related to $[X^1, X^2]$, we have

$$[X^1, X^2]_\varphi = d\varphi([Y^1, Y^2]) \in \Gamma(E).$$

$\square$

The Frobenius's Theorem, which is due to Clebsh, Deahna, and of course Frobenius, is the converse of Theorem 2.1.8.1.

Theorem 2.1.8.2 (Frobenius's Integrability Theorem). *A k-plane field E is integrable if and only if $\Gamma(E)$ is a Lie subalgebra.*

Remark 2.1.8.1. *Note that Frobenius's Theorem is a prototype of the following theorem, which we shall prove later: Let G be a Lie group and $\mathfrak{g}$ be the Lie algebra. Suppose that $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. Then there exists a Lie group H whose Lie algebra is $\mathfrak{h}$ and an injective immersive homomorphism $f : H \to G$.*

Before proving this, we give some definitions and an example of a nonintegrable k-plane field. Let G be a Lie group. We define homomorphisms from G to $\mathbf{Diff}(G)$ by the following ways:

$$G \xrightarrow{l} \mathbf{Diff}(G) \text{ by } g \mapsto l_g \text{ where } l_g(h) = gh \text{ for } h \in G$$
$$G \xrightarrow{r} \mathbf{Diff}(G) \text{ by } g \mapsto r_{g^{-1}} \text{ where } r_{g^{-1}}(h) = hg^{-1} \text{ for } h \in G.$$

We call l_g a left multiplication by g and r_g a right multiplication by g. Note that to make r a homomorphism, we send g to $r_{g^{-1}}$.

Definition 2.1.8.6 (Left-invariant vector fields). *Let $X \in \mathfrak{X}(G)$. We say X is a left-invariant vector field (resp. right-invariant vector field) if $l_g \cdot X = X$ (resp. $r_g \cdot X = X$) for all g.*

Note that

$$l_g \cdot X \overset{def}{=} dl_g \circ X \circ l_{g^{-1}}. \text{ i.e., } (l_g \cdot X)_h = (dl_g)_{g^{-1}h}(X_{g^{-1}h}).$$

We often write $dl_g(X_{g^{-1}h})$ as $(dl_g)_{g^{-1}h}(X_{g^{-1}h})$. In this description, we have $X \in \mathfrak{X}(G)$ is a left-invariant vector field if $dl_g(X_{g^{-1}h}) = X_h$ for all $h \in G$. Using the smooth group operation of G, it is easy to see that if $dl_h(X_e) = X_h$ for all $h \in G$ then $X \in \mathfrak{X}(G)$ is a left-invariant vector field:

$$dl_g(X_{g^{-1}h}) = dl_g(dl_{g^{-1}h}(X_e)) = dl_g \circ dl_{g^{-1}h}(X_e) = dl_h(X_e) = X_h$$

Hence, a left-invariant vector field is completely determined by evaluation at e, the identity of G. So, an obvious question would be " Can an element in $T_e(G)$ give a left-invariant vector field?"

Exercise 2.1.8.2. *$\mathfrak{LX}(G)$, the set of left-invariant vector fields, is isomorphic to $T_e(G)$ as vector spaces. Hence, $\dim \mathfrak{LX}(G) = \dim T_eG$ as vector spaces.*

Proof. Define

$$\epsilon : \mathfrak{LX}(G) \to T_eG \text{ by } \epsilon(X) = X_e.$$

Note that the above discussion already showed that it is a monomorphism. We are going to show it is onto. If $v \in T_e(G)$, let

$$X_h = dl_h(v).$$

The subtle part is the smoothness of this construction. We have to show that

$$dl_h(v)(f) \text{ is smooth for all } f \in C^\infty(G).$$

That is, regarding v as a first order differential operator, can $(dl_h(v)(f))(g) = v(f \circ l_h(g)) = v(f(hg))$ be smooth as g and h vary? Since G has a smooth group operation, without loss of generality, we can assume $g = e$. Since $f(h)$ is smooth and v is a first order differential operator, $v(f(h))$ is smooth as h varies. Another thing to check is that $dl_h(v)$ is left-invariant. A straitforward check would show that it is so. From this, we deduce that the set of left-invariant vector fields $\mathfrak{LX}(G)$ is isomorphic to $T_e(G)$ as vector spaces. □

Interesting consequences of Exercise 2.1.8.2 are the following: Since $\dim \mathfrak{LX}(G) = \dim T_eG$ as vector spaces, it shows that the tangent bundle $T(G)$ has linearly independent vector fields as many as $\dim G$, That is, every Lie group is parallelizable. For example, we know that $\mathbb{H}_1$, the set of unit length quaternions is Lie group isomorphic to S^3. Hence, S^3 is parallelizable. Moreover, by purely dimensional reason, we have $\mathfrak{LX}(G) \cong \mathfrak{X}(G)$ as $\mathbb{R}$-vector spaces. So, we have

Theorem 2.1.8.3. *Let G be a Lie group with $\dim G = n$. The set of smooth vector fields $\mathfrak{X}(G)$ has n linearly independent smooth vector fields, i.e., a basis, which consists of left-invariant vector fields.*

Note that $\mathfrak{LX}(G)$ and $\mathfrak{RX}(G)$ are not $C^\infty(G)$-submodules unlike $\mathfrak{X}(G)$, i.e., fX need not be a left-invariant vector field for $f \in C^\infty(G)$ and $X \in \mathfrak{LX}(G)$. So, $\mathfrak{LX}(G)$ is not isomorphic to $\mathfrak{X}(G)$ as $C^\infty(G)$-modules.

Exercise 2.1.8.3. *Let $\mathfrak{RX}(G)$ be the set of right-invariant vector fields. Show that $\mathfrak{LX}(G)$ and $\mathfrak{RX}(G)$ are subalgebras of $\mathfrak{X}(G)$ with respect to commutators and they are isomorphic.*

Proof. Note that a similar proof of Exercise 2.1.8.2 shall show that $\mathfrak{RX}(G)$ is isomorphic to $T_e(G)$. Hence, by the virtue of Exercise 2.1.8.2, it suffices to show that $[X, Y] \in \mathfrak{LX}(G)$ for $X, Y \in \mathfrak{LX}(G)$. The same proof can apply the other case. A moment of thought would give you that X is l_g-related to X and Y is l_g-related to Y for any $g \in G$. By Exercise 2.1.8.1, we conclude that $[X, Y]$ are l_g-related to $[X, Y]$ for any $g \in G$. That is,

$$l_g \cdot [X, Y] = [X, Y] \text{ for any } g \in G.$$

Hence, $[X, Y]$ is a left-invariant vector field. □

By the virtue of Exercise 2.1.8.3, it is obvious that there will be no differences between $\mathfrak{LX}(G)$ and $\mathfrak{RX}(G)$ as far as algebras are concerned. So, from now on we exclusively are working on $\mathfrak{LX}(G)$.

Exercise 2.1.8.4. *Show that every left-invariant vector field X on a Lie group G is complete.*

Proof. Suppose that $X \in \mathfrak{LX}(G)$ and $\Phi : (-\epsilon, \epsilon) \times U \to G$ be a corresponding local flow where U is an open sets in G. Letting $\Phi_t^g(h) = l_g(\Phi_t(l_{g^{-1}}(h))$ for $h \in l_g(U)$, we have a local flow

$$\Phi_t^g : (-\epsilon, \epsilon) \times U \to G.$$

The upshot is that the fact that X is left-invariant makes Φ_t^g also a local flow on $l_g(U)$ of X: For $h \in l_g(U)$,

$$\frac{d}{dt}|_{t=0} l_g(\Phi_t(l_{g^{-1}}(h)) = dl_g(X_{g^{-1}h}) = X_h.$$

Hence, in this way, we can have a local flow

$$\Phi : (-\epsilon, \epsilon) \times G \to G.$$

Now we want to remind you of the proof of Exercise 2.1.4.1. When M was compact, the compactness gave the above kind of a local flow. Once we had that kind of a local flow, without using the compactness assumption, we could extend a local flow from $(-\epsilon, \epsilon)$ to $\mathbb{R}$. Indeed, it is a general theorem. If you have a local flow such that $\Phi : (-\epsilon, \epsilon) \times M \to M$ for a smooth manifold M, you can extend a global flow $\Phi : \mathbb{R} \times M \to M$ by the construction in the proof of Exercise 2.1.4.1. So, in our case, Φ can be extended to a global flow. Hence, X is complete. □

Exercise 2.1.8.5. *Show that X is a left-invariant vector field on G if and only if one parameter subgroup of X is a right multiplication.*

Proof. Suppose that X is a left invariant vector field on G. Exercise 2.1.8.4 tells you that there exists a one parameter subgroup Φ_t. Since $l_g \cdot X = X$, by Exercise 2.1.6.1, we have

$$l_g \circ \Phi_t = \Phi_t \circ l_g.$$

So we have for all $g, h \in G$,

$$g\Phi_t(h) = l_g \circ \Phi_t(h) = \Phi_t \circ l_g(h) = \Phi_t(gh).$$

Letting $h = e$, we have for all $g \in G$

$$\Phi_t(g) = g\Phi_t(e) = r_{\Phi_t(e)}(g).$$

That is, Φ_t is a just right multiplication by $\Phi_t(e)$. Now, suppose that a one parameter subgroup Φ_t, which generates X, is a right multiplication, i.e., $\Phi_t(g) = ga_t$ where $a_t \in G$. Note that Exercise 2.1.6.1 also tells that the one parameter subgroup of $l_g \cdot X$ is $l_g \circ \Phi_t \circ l_{g^{-1}}$. Since

$$l_g \circ \Phi_t \circ l_{g^{-1}}(h) = g\Phi_t(g^{-1}h) = gg^{-1}ha_t = \Phi_t(h),$$

we conclude that $l_g \cdot X = X$. □

Now we go back to the main theme of this section. We give an example of a nonintegrable k-plane field.

Exercise 2.1.8.6. *Find an $\mathbb{R}^2$-plane field on $\mathbb{R}^3$ which is not integrable.*

Proof. Consider a Heisenberg group

$$\mathbf{H}_3 = \{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mid x, y, z \in \mathbb{R} \} \cong \mathbb{R}^3.$$

Note that

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\xi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x-\xi & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\theta \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x & -\theta x + z \\ 0 & 1 & y-\theta \\ 0 & 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x & z-\delta \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

Obviously,

$$\Xi_\xi(x,y,z) = \begin{pmatrix} 1 & x-\xi & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \text{ is a global flow of } \mathbf{H}_3 \text{ where } \xi \in \mathbb{R},$$
$$\Theta_\theta(x,y,z) = \begin{pmatrix} 1 & x & -\theta x + z \\ 0 & 1 & y-\theta \\ 0 & 0 & 1 \end{pmatrix} \text{ is a global flow of } \mathbf{H}_3 \text{ where } \theta \in \mathbb{R}, \text{ and}$$
$$\Delta_\delta(x,y,z) = \begin{pmatrix} 1 & x & z-\delta \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \text{ is a global flow of } \mathbf{H}_3 \text{ where } \delta \in \mathbb{R}.$$

From Exercise 2.1.8.5, we know that Ξ_ξ, Θ_θ, and Δ_δ generate left-invariant vector fields, since Ξ_ξ, Θ_θ, and Δ_δ are right multiplications. By identifying

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} = (x,y,z),$$

it easy to see that

$$X = \frac{\partial}{\partial x} \text{ is the left-invariant vector field of } \Xi_\xi = (x-\xi, y, z),$$
$$Y = \frac{\partial}{\partial y} + x\frac{\partial}{\partial z} \text{ is the left-invariant vector field of } \Theta_\theta = (x, y-\theta, z-\theta x), \text{ and}$$
$$Z = \frac{\partial}{\partial z} \text{ is the left-invariant vector field of } \Delta_\delta = (x,y,z-\delta).$$

Moreover, X, Y, and Z generate $\mathfrak{X}(\mathbf{H}_3)$. Note that this is a nice example of Theorem 2.1.8.3. Let E be spanned by X and Y, i.e.,

$$E = \text{Span}\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}\}.$$

Obviously, $\frac{\partial}{\partial z} \notin \Gamma(E)$. However, we have

$$[\frac{\partial}{\partial x}, \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}] = \frac{\partial}{\partial z}.$$

So, the Frobenius theorem says that it is not integrable.

Now, we investigate E further: We will give a geometrical explanation of nonintegrability of E. One of geometrical meanings of Frobenius theorem is that E must look like a flat 2-plane in order to be integrable. That is, the curvature form of E as a vector bundle over $\mathbb{R}^3$ must be zero. Even though you do not know how to get the curvature form of a vector bundle, at least intuitively it would be clear that E does not look flat from the factor $x\frac{\partial}{\partial z}$ in the formula

$$E = \text{Span}\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}\}.$$

In general, a k-plane distribution is integrable if and only if the curvature form is zero. It is somewhat obvious, since the curvature form of a tangent bundle is always zero. Now we show that even though E is not integrable, which is not so nice, E still has some nice property. Let $\gamma : \mathbb{R} \to \mathbf{H}_3$ be a curve. We say $\gamma(t)$ is E-horizontal if

$$\gamma'(t) \in E_{\gamma(t)}.$$

Obviously, for fixed $(a_1, a_2, a_3) \in \mathbb{R}^3$, the curves $\Theta_\theta(a_1, a_2, a_3)$ and $\Xi_\xi(a_1, a_2, a_3)$ are E-horizontal. We claim that for a given $p, q \in \mathbb{R}^3$, there exists an E-horizontal curve from p to q. Let

$$\omega = dz - xdy.$$

It is easy to see that $\omega(X) = \omega(Y) = 0$ and $\omega(Z) \neq 0$. So, ω is everywhere nonzero. Since

$$E_p = \mathbf{ker}(\omega_p : T_p(\mathbf{H}_3) \to \mathbb{R}),$$

we have $\omega(\gamma'(t)) = 0$ if and only if $\gamma(t)$ is E-horizontal. Now, we construct $\gamma(t)$ such that $\omega(\gamma'(t)) = 0$ from p to q. It suffices to construct piecewise smooth $\gamma(t)$, since there exists a local integral E-horizontal curve. Without loss of generality, we can assume $p = (0, 0, 0)$. Let $q = (q_1, q_2, q_3)$. Since $\omega_{(x(t),a,b)}(x'(t), 0, 0) = 0$, any line curve parallel to x-axis is horizontal. Moreover, $\omega_{(0,y(t),b)}(0, y'(t), 0) = 0$,, any line parallel to y-axis in yz-plane is horizontal. Hence, we have piecewise smooth E-horizontal curve $\gamma(t)$ from $(0, 0, 0)$ to (q_1, q_2, r). From this, it suffices to show that there exists an E-horizontal curve $\gamma(t)$ from $(0, 0, 0)$ to $(0, 0, a)$ for a given $a \in \mathbb{R}$. By the formula $\omega(\gamma'(t)) = 0$, we have to show that there exists $\gamma(t) = (x(t), y(t), z(t)) : [0, s_1] \to \mathbb{R}^3$ from $(0, 0, 0)$ to $(0, 0, a)$ such that

$$\frac{dz(t)}{dt} = x(t)\frac{dy(t)}{dt}.$$

Since

$$z(s) = \int_0^s x(t)dy(t),$$

it suffices to show that there exists a closed curve $\delta(t) = (x(t), y(t))$ in xy-plane, which gives $z(s_1) = a$. Since δ is a boundary of A, i.e., $\delta = \partial A$, by stoke's theorem

$$z(s_1) = \int_0^{s_1} x(t)dy(t) = \int_\delta xdy = \int_A dxdy = \text{ area of } A.$$

That is, we can find $\delta(t)$. Note that $-dxdy$ is the curvature form of E and $\omega = dz - xdy$ is the connection of E. ☐

We note that sometimes people call an integrable distribution as an involutive or completely integrable and E in Exercise 2.1.8.6 as an integrable distribution. Also, it is easy to see that 0-dimensional foliation of M is nothing but a differentiable structure of M. Now we give a proof of the Frobenius theorem. Note that the only direction we need to be proved is that if $\Gamma(E)$ is a Lie subalgebra where E is an k-plane distribution on M, then E is integrable.

Proof of The Frobenius Theorem. We prove this by induction on rank of E. When E is 0-plane distribution, it is vacuously true. Suppose that the theorem holds for rank $r-1$ and E is an r-plane distribution on M with a Lie subalgebra $\Gamma(E)$. Since we want to show there exists a submanifold N_p at each point p of M such that $T_p(N_p) = E_p$, obviously it suffices to prove this in a local chart.

Moreover, by the local triviality of a vector bundle, we can assume that $\Gamma(E)|_U$ is generated by r linearly independent vector fields, $X_1, \ldots, X_r$. Of course, X_i is everywhere nonzero for being an element of a basis for $i = 1, \ldots, r$. So, by Theorem 2.1.4.7, there exists a coordinate system $(x^1, \ldots, x^n)$ on $U \subseteq M$ with $(x^1(p), \ldots, x^n(p)) = (0, \ldots, 0)$ such that

$$X_r = \frac{\partial}{\partial x^1}.$$

So, there exists a function $f \in C^\infty(U)$ such that $X_r(f) = 1$ and $f(p) = 0$. Note that $f = x^1$ and $X_r(f) = df(X_r) = 1$. The main point is that by construction $f : U \to \mathbb{R}$ is a submersion. So, by the implicit function theorem, $f^{-1}(0)$ is an $n-1$ dimensional submanifold of U through p with $T(f^{-1}(0)) = \mathbf{ker}(df)$. Let $f^{-1}(0) = M'$. Note that we do not know whether or not $X_i \in TM'$. So, we let

$$Y_i = X_i - X_i(f)X_r \text{ for } i < r \text{ and } Y_r = X_r.$$

Clearly, E is again spanned by $Y_1, \ldots, Y_r$ and $E' = \mathbf{span}\{Y_1, \ldots, Y_{r-1}\} \subseteq \mathbf{ker}(df) = TM'$. Note that $Y_r = X_r \notin TM'$, since $X_r(f) = 1$. Obviously, we have E' is an $(r-1)$-pane field on M'. We claim that E' is integrable. By the induction hypothesis, it suffices to show that $\Gamma(E')|_{M'}$ is a Lie subalgebra of $\mathfrak{X}(M')$. Let $Y^1, Y^2 \in \Gamma(E')|_{M'}$. So,

$$Y^1 = \sum_{i=1}^{r-1} c_i^1(x)Y_i \text{ and } Y^2 = \sum_{i=1}^{r-1} c_i^2(x)Y_i.$$

So, we have

$$[Y^1, Y^2] = \sum_{i,j=1}^{r-1} c_i^1(x)c_j^2(x)[Y_i, Y_j] + \sum_{i,j=1}^{r-1} c_i^1(x)(Y_i c_j^2(x))Y_j - \sum_{i,j=1}^{r-1} c_j^2(x)(Y_j c_i^1(x))Y_i.$$

Clearly,

$$\sum_{i,j=1}^{r-1} c_i^1(x)(Y_i c_j^2(x))Y_j - \sum_{i,j=1}^{r-1} c_j^2(x)(Y_j c_i^1(x))Y_i \in \Gamma(E')|_{M'}.$$

So, it suffices to show that

$$\sum_{i,j=1}^{r-1} c_i^1(x)c_j^2(x)[Y_i, Y_j] \in \Gamma(E')|_{M'}.$$

That is, it suffices to show that for $i, j < r$, we can write

$$[Y_i, Y_j] = \sum_{k=1}^{r-1} b_{ij}^k(x)Y_k.$$

Since $\Gamma(E)|_U$ is a Lie subalgebra by the assumption, we can write

$$[Y_i, Y_j] = \sum_{k=1}^{r} b_{ij}^k(x)Y_k.$$

Since $Y_r(f) = 1$ and $Y_i(f) = 0$ for $i < r$, we have $[Y_i, Y_j]f = 0$ for $i, j < r$. Hence, we have for $i, j < r$

$$0 = [Y_i, Y_j]f = \sum_{k=1}^{r} b_{ij}^k(x) Y_k f = b_{ij}^r.$$

Therefore, by the induction hypothesis, E' is integrable. That is, there exists a $(r-1)-$submanifold $S' \subseteq M' \subset U$ for E' through p. Note that $E = X_r \oplus E' \subset TU$. Letting ξ_t be the flow of X_r, define

$$S = \bigcup_{|t|<\epsilon} \xi_t(S').$$

Since $E = X_r \oplus E'$ and ξ_t is a local diffeomorphism, S is diffeomorphic to a subset $W \subseteq \mathbb{R} \times \xi_0(S') \cong \mathbb{R} \times S'$. Theorem 2.1.4.3 tells that W is an open set of $\mathbb{R} \times S'$. So, it is a $r - 1$ submanifold and obviously $p \in S$ and $E = X_{r,p} + T_pS' = T_pS$. □

2.2 Graded Algebras of Smooth Manifolds

2.2.1 Exterior Algebra

Definition 2.2.1.1 (Tensor product). *Let V and W be vector spaces. A m-tensor product is a universal object in the sense that every m-multilinear map $\varphi : V \times \cdots \times V \to W$ factors through a universal m-multilinear map $V \times \cdots \times V \to V \otimes \cdots \otimes V$. That is,*

$$\begin{array}{ccc} V \times \cdots \times V & \xrightarrow{\varphi} & W \\ \otimes \big\downarrow & \nearrow \widetilde{\varphi} & \\ \underbrace{V \otimes \cdots \otimes V}_{m} & & \end{array}$$

Of course, the universal object always exists uniquely. Observe the following: Let $\varphi_1 : V_1 \to W_1$ and $\varphi_2 : V_2 \to W_2$ be linear maps between vector spaces. Since $W_1 \otimes W_2$ always exists uniquely (up to isomorphism), we have

$$V_1 \times V_2 \xrightarrow{\varphi_1 \times \varphi_2} W_1 \times W_2 \xrightarrow{\otimes} W_1 \otimes W_2.$$

That is, we have a bilinear map $\varphi_1 \otimes \varphi_2 : V_1 \times V_2 \to W_1 \otimes W_2$. From this, we deduce that

an m-multilinear linear map $\otimes$ l-multilinear linear map is an $(m + l)$-multilinear map.

Suppose that V_1, V_2 are vector spaces over a field k. Note that since k is a field, there is a canonical isomorphism $k \otimes k \cong k$. So, given linear maps $\varphi_1 : V_1 \to k$ and $\varphi_2 : V_2 \to k$, we have

$$\varphi_1 \otimes \varphi_2 : V_1 \times V_2 \to k.$$

We often write $\otimes^m V = \overbrace{V \otimes \cdots \otimes V}^{m}$ and $\bigotimes^m V \otimes \bigotimes^l V = \bigotimes^{m+l} V$. Using tensor products, we have a graded algebra of vector spaces.

Definition 2.2.1.2 (Graded algebra). *Let V be a vector space. We let*

$$\mathcal{A} = \bigoplus_{m \geq 0} \mathcal{A}_m \text{ where } \mathcal{A}_m \stackrel{def}{=} \otimes^m V.$$

By defining multiplicative structures $\mathcal{A}_m \times \mathcal{A}_l \xrightarrow{\otimes} \mathcal{A}_{m+l}$, we have an associative graded algebra structure.

Let $\mathcal{S}_n$ be a symmetric group, i.e., the group of permutations of n-element. We define an action of $\mathcal{S}_n$ on $\otimes^m V$ by the following way: For $\sigma \in \mathcal{S}_n$, letting $(-1)^\sigma$ be the sign of σ,

$$\sigma \cdot (v_1 \otimes \cdots \otimes v_n) \stackrel{def}{=} (-1)^\sigma v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}.$$

Using the above, we define that pth exterior product of V

$$\bigwedge^p V = \{v \in \otimes^p V \mid \sigma \cdot v = v \text{ for all } \sigma \in \mathcal{S}_p\} = (\otimes^p V)^{\mathcal{S}_p}.$$

Note that if $A \subseteq X$ is a subspace of X, then A is called a retract of X if there is a retraction $f : X \to A$, i.e., $f(a) = a$ for all $a \in A$.

Theorem 2.2.1.1. *Given any finite group G and a vector space A with G-action, i.e., $g \cdot v = \rho_g(v)$ where $\rho : G \to \mathbf{GL}(V)$ is a linear representation, then*

$$V^G = \{v \in V \mid g \cdot v = v \text{ for all } g \in G\} \text{ is a retract of } V.$$

Proof. we construct a retraction $\varphi : V \to V^G$ by the following way:

$$\varphi(v) = \frac{1}{|G|} \sum_{g \in G} g \cdot v.$$

It is easy to see that $h \cdot \varphi(v) = v$ for all $h \in G$, which shows that it is well defined and $\varphi(v) = v$ for all $v \in V^G$, which says that it is a retraction. ☐

Since $\mathcal{S}_p$ is finite, we deduce that in this description, we define

$$\varphi : \otimes^p V \to \wedge^p V \text{ by } \varphi(v) = \sum_{\sigma \in \mathcal{S}_p} \sigma \cdot v.$$

Note that still $\sigma \cdot \varphi(v) = \varphi(v)$ for all $\sigma \in \mathcal{S}_p$.

Example 2.2.1.1. *Let $v_1, v_2 \in V$. Then $\varphi(v_1 \otimes v_2) = v_1 \wedge v_2 = v_1 \otimes v_2 - v_2 \otimes v_1$. Hence, $v \wedge v = 0$ for $v \in V$ and $v_1 \wedge v_2 = -v_2 \wedge v_1$.*

Example 2.2.1.2. *Let $v_{i_1}, \ldots, v_{i_m} \in V$. So, we have*

$$v_{i_1} \otimes \cdots \otimes v_{i_m} \mapsto v_{i_1} \wedge \cdots \wedge v_{i_m}.$$

If $i_j = i_{j+1}$, then by Example 2.2.1.1, we have $v_{i_1} \wedge \cdots \wedge v_{i_m} = 0$. So, we can always assume $i_1 < i_2 < \cdots < i_m$ in a wedge product.

Example 2.2.1.3. *If $\{v_1, \ldots, v_n\}$ is a basis of V, it is easy to see that $\{v_{i_1} \otimes \cdots \otimes v_{i_m} \mid 1 \leq i_j \leq n\}$ is a basis of $\otimes^m V$. So, $\dim \otimes^m V = n^m$. By Exercise 2.2.1.2, we have*

$$\{v_{i_1} \wedge \cdots \wedge v_{i_m} \mid 1 \leq i_1 < \cdots < i_m \leq n\}$$

is a basis of $\wedge^m V$ and $\dim \wedge^m V = \binom{n}{m}$. Especially, $\wedge^n V$ is generated by $v_1 \wedge \cdots \wedge v_n$.

Example 2.2.1.4. *Let $\alpha \in \wedge^p V, \beta \in \wedge^q V$. Since $\wedge^m V$ is a retract of $\otimes^m V$, we can think $\alpha \in \otimes^p V$ and $\beta \in \otimes^q V$. So, we get $\alpha \otimes \beta \in \otimes^{p+q} V$. Hence,*

$$\alpha \otimes \beta \overset{\varphi}{\mapsto} \alpha \wedge \beta \in \wedge^{p+q} V.$$

So, the graded (associative) algebra of tensor products gives the graded (associative) algebra of exterior product, which is called ***an exterior algebra.*** *Also, by Exercise 2.2.1.1, it is easy to see that for $\alpha \in \wedge^p V, \beta \in \wedge^q V$, we have*

$$\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha = (-1)^{|\alpha||\beta|} \beta \wedge \alpha.$$

The relation is called a ***commutative graded algebra.***

Example 2.2.1.5. *We know that given linear maps $\varphi_1 : V_1 \to k$ and $\varphi_2 : V_2 \to k$ where V_1, V_2 are vector space over k, we have*

$$\varphi_1 \otimes \varphi_2 : V_1 \times V_2 \to k \text{ by } (v_1, v_2) \mapsto \varphi_1(v_1)\varphi_2(v_2).$$

So, we also have

$$\varphi_1 \wedge \varphi_2 : V_1 \times V_2 \to k \text{ by } (v_1, v_2) \mapsto \varphi_1(v_1)\varphi_2(v_2) - \varphi_2(v_1)\varphi_1(v_2).$$

Clearly, $\varphi \wedge \varphi = 0$ and $\varphi_1 \wedge \varphi_2 = -\varphi_2 \wedge \varphi_1$.

Exercise 2.2.1.5 shows that we can construct a pth exterior product of the dual space V^* of V. Also, by Exercise 2.2.1.4, we also have an exterior algebra of the dual V^*. In particular, we can associate an n-dimensional vector space V to $\wedge^n V^*$ as an analogy that we can associate $\mathbf{GL}(n, \mathbb{R})$ to $\mathbb{R}^* = \mathbf{GL}(1, \mathbb{R})$ by the **det** function.

Definition 2.2.1.3. *Let $\mathcal{A} = \bigoplus_{m \geq 0} \mathcal{A}_m$ and $\mathcal{B} = \bigoplus_{m \geq 0} \mathcal{B}_m$ be graded algebras. Letting $\mathcal{A}_m = \mathcal{B}_m = 0$ if $m < 0$, a linear map $F : \mathcal{A} \to \mathcal{B}$ is called a graded linear map of degree d if $F : \mathcal{A}_p \to \mathcal{B}_{p+m}$.*

Example 2.2.1.6. *Let $f : V \to W$ be a homomorphism of vector spaces. In an obvious way, it is easy to see that f induces graded homomorphisms of exterior products having degree 0,*

$$f^{(p)} : \wedge^p V \to \wedge^p W.$$

Example 2.2.1.7. *Let $M^m \xrightarrow{f} N^n$ be a smooth map between smooth manifolds. For $f(p) \in N$, we have $(df)^T : T^*_{f(p)} N \to T^*_p M$. So, $(df)^T$ induces a homomorphism of commutative graded algebras with degree 0,*

$$\begin{array}{ccc} \wedge^k T^* M & \xleftarrow{((df)^T)^{(k)}} & \wedge^k T^* N \\ \downarrow & & \downarrow \\ M & \xrightarrow{f} & N. \end{array}$$

Let $\Gamma(\wedge^k T^*M) = \Omega^k(M)$, the set of k-forms. Obviously, $\Omega^*(M)$ and $\Omega^*(N)$ are commutative graded algebras and we also have

$$f^* : \Omega^k N \to \Omega^k M \text{ by } f^*(\omega) = (df)^T \circ \omega \circ f.$$

Note that $(g \circ f)^* = f^* \circ g^*$.

2.2.2 Graded derivations

Definition 2.2.2.1 (Graded derivations). *Let $\mathcal{A}$ be a graded algebra and $D : \mathcal{A} \to \mathcal{A}$ be a graded homomorphism of degree d, i.e., $D(\mathcal{A}_m) \subseteq \mathcal{A}_{m+d}$. We say D is a graded derivation of degree d if for $\alpha, \beta \in \mathcal{A}$,*

$$D(\alpha \cdot \beta) = D(\alpha) \cdot \beta + (-1)^{|\alpha|d} \alpha \cdot D(\beta).$$

Example 2.2.2.1. *Let V be a vector space and V^* be the dual. Every $w \in V^*$ defines a graded derivation ι_w of $\wedge^* V$ of degree -1, which is called an interior multiplication by w*

Proof. Define

$$\iota_w = \begin{cases} 0 & \text{on } \wedge^0 V = k, \text{ scalar} \\ w & \text{on } \wedge^1 V = V \\ \sum_{i=1}^p (-1)^{i-1} v_1 \wedge \cdots \wedge w(v_i) \wedge \cdots \wedge v_p & \text{on } \wedge^p V. \end{cases}$$

It is easy to see that ι_w is a graded derivation of degree -1. □

The above shows that any linear functional $w \in V^*$ defines an interior multiplication, i.e., a derivation of degree -1, on exterior algebra $\wedge^* V$.

Example 2.2.2.2. *Every vector field $X \in \mathfrak{X}(M)$ defines an interior multiplication, a derivation of degree -1, ι_X on $\Omega^*(M)$.*

Proof. Note that we already know that $\Omega^*(M) = \oplus_{k \geq 0} \Omega^k(M)$ is an associative commutative graded algebra with $\Omega^0(M) = C^\infty(M)$. Let $X \in \mathfrak{X}(M)$. For $\omega \in \Omega^k(M)$, define

$$\iota_X(\omega)(Y_1, \ldots, Y_{k-1}) = \omega(X, Y_1, \ldots, Y_{k-1}) \text{ where } Y_i \in \mathfrak{X}(M).$$

From this, we have

$$\iota_X : \Omega^k(M) \to \Omega^{k-1}(M).$$

Since for $p \in M$, $\iota_{X_p} : \wedge^k T_p^* M \to \wedge^{k-1} T_p^* M$ defines a derivation of degree -1 by Exercise 2.2.2.1, ι_X is a derivation of degree -1 on $\Omega^*(M)$. □

Note that Exercise 2.1.2.1 showed that for any ungraded algebra $\mathcal{A}$, $Der(\mathcal{A})$ forms a Lie algebra under

$$[X, Y] = X \circ Y - Y \circ X \text{ where } X, Y \in Der(\mathcal{A}).$$

Exercise 2.2.2.1. *Let $\mathcal{A}$ be an associative graded commutative algebra and $Der_k(\mathcal{A})$ be the set of derivations on $\mathcal{A}$ of degree k. Given graded derivation $X \in Der_k(\mathcal{A})$ and $Y \in Der_l(\mathcal{A})$, the graded commutator*

$$[X, Y] = X \circ Y - (-1)^{kl} Y \circ X$$

is a graded derivation of degree $k + l$. Moreover, show that

$$Der(\mathcal{A}) = \bigoplus_{k \geq 0} Der_k(\mathcal{A})$$

forms a graded Lie algebra.

Proof. It is obvious that By defining for all $a, d \in \mathbb{R}$, all $X_1, X_2 \in Der_k(\mathcal{A})$, and all $f \in \mathcal{A}$,

$$(aX_1 + bX_2)(f) = aX_1(f) + bX_2(f),$$

we conclude that $Der_k(\mathcal{A})$ is a vector space. Let $X \in Der_k(\mathcal{A})$, $Y \in Der_l(\mathcal{A})$ and $Z \in Der_j(\mathcal{A})$. Obviously, the degree of $[X, Y]$ is $k + l$. Moreover,

$$\begin{aligned}
[X, Y](\alpha \cdot \beta) &= (X \circ Y - (-1)^{kl} Y \circ X)(\alpha \cdot \beta) = X \circ Y(\alpha \cdot \beta) - (-1)^{kl} Y \circ X(\alpha \cdot \beta) \\
&= X((Y\alpha) \cdot \beta + (-1)^{|\alpha|l} \alpha \cdot (Y\beta)) - (-1)^{kl} Y((X\alpha) \cdot \beta + (-1)^{|\alpha|k} \alpha \cdot (X\beta)) \\
&= ([X, Y]\alpha) \cdot \beta + (-1)^{(|\alpha|+l)k} Y\alpha \cdot X\beta + (-1)^{|\alpha|l} X\alpha \cdot Y\beta + (-1)^{|\alpha|l+|\alpha|k} \alpha \cdot XY\beta \\
&- (-1)^{kl+|\alpha|l+kl} X\alpha \cdot Y\beta - (-1)^{kl+|\alpha|k} Y\alpha \cdot X\beta - (-1)^{kl+|\alpha|k+|\alpha|l} \alpha \cdot YX\beta \\
&= ([X, Y]\alpha) \cdot \beta + (-1)^{|\alpha|(k+l)} \alpha \cdot ([X, Y]\beta) \\
&+ (-1)^{(|\alpha|+l)k} Y\alpha \cdot X\beta + (-1)^{|\alpha|l} X\alpha \cdot Y\beta - (-1)^{kl+|\alpha|l+kl} X\alpha \cdot Y\beta - (-1)^{kl+|\alpha|k} Y\alpha \cdot X\beta \\
&= ([X, Y]\alpha) \cdot \beta + (-1)^{|\alpha|(k+l)} \alpha \cdot ([X, Y]\beta).
\end{aligned}$$

This shows that $[X, Y] \in Der_{k+l}(\mathcal{A})$. Moreover, a straightforward computation shall show that

$$[X, Y] = (-1)^{kl+1}[Y, X] \text{ and } (-1)^{kj}[X, [Y, Z]] + (-1)^{lk}[Y, [Z, X]] + (-1)^{jl}[Z, [X, Y]] = 0.$$

Hence, $Der(\mathcal{A}) = \bigoplus_{k \geq 0} Der_k(\mathcal{A})$ forms a graded Lie algebra. □

Example 2.2.2.3.

$$[\iota_X, \iota_Y] = \iota_X \circ \iota_Y + \iota_Y \circ \iota_X.$$

Exercise 2.2.2.2. *Let V be an n-dimensional vector space. Show that*

$$\wedge^k V^* \cong (\wedge^k V)^*.$$

Proof. Example 2.2.1.3 says that if $\{e_1, \ldots, e_n\}$ is a basis of V, then

$$\{e_{i_1} \wedge \cdots \wedge e_{i_k} \mid 1 \leq i_1 < \cdots < i_k \leq n\}$$

is a basis of a vector space $\wedge^k V$ and $\dim \wedge^k V = \binom{n}{k}$. So, $\dim(\wedge^k V)^* = \binom{n}{k}$. By the same consideration, we have

$$\{e^*_{i_1} \wedge \cdots \wedge e^*_{i_k} \mid 1 \leq i_1 < \cdots < i_k \leq n\}$$

is a basis of a vector space $\wedge^k V^*$ and $\dim \wedge^k V^* = \binom{n}{k}$. So, to show $\wedge^k V^* \cong (\wedge^k V)^*$ it suffices to construct a monomorphism

$$\varphi : \wedge^k V^* \to (\wedge^k V)^*.$$

Letting $\binom{n}{k} = N$, define

$$\varphi(e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*)(e_{j_1} \wedge \cdots \wedge e_{j_k}) = e_{i_1}^*(e_{j_1}) \cdots e_{i_k}^*(e_{j_k}) \text{ and}$$
$$\varphi(\sum_{i=1}^N c_i e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*) = \sum_{i=1}^N c_i \varphi(e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*).$$

By construction φ is well-defined and linear. Suppose that $\varphi(\sum_{i=1}^N c_i e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*) = 0$. Since

$$0 = \varphi(\sum_{i=1}^N c_i e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*)(e_{j_1} \wedge \cdots \wedge e_{j_k}) = c_j \text{ for each } j,$$

we conclude that $\sum_{i=1}^N c_i e_{i_1}^* \wedge \cdots \wedge e_{i_k}^* = 0$, So, it is a monomorphism. □

Let $v \in V$ where V is a vector space. Define an exterior multiplication e_v by

$$e_v : \wedge^k V \to \wedge^{k+1} V \text{ by } w \mapsto v \wedge w.$$

It is easy to see that

$$(e_v)^T : (\wedge^{k+1} V)^* \to (\wedge^k V)^*.$$

Now, take $V = T_pM$ where M is a smooth manifold and $p \in M$. For $X \in \mathfrak{X}(M)$, we have

$$\begin{array}{ccc} (\wedge^{k+1} T_pM)^* & \xrightarrow{(e_{X_p})^T} & (\wedge^k T_pM)^* \\ \cong\downarrow & & \cong\downarrow \\ \wedge^{k+1} T_p^*M & \xrightarrow{\iota_{X_p}} & \wedge^k T_p^*M. \end{array}$$

It is easy to see that the above diagram commutes. So, since $e_v \circ e_v = 0$ by the skew-symmetry, we conclude that

$$\iota_X \circ \iota_X = 0 \text{ for } X \in \mathfrak{X}(M).$$

If you are suspicious about the proof of $\iota_X \circ \iota_X = 0$, see Corollary 2.2.2.1.1.

Definition 2.2.2.2 (Vector bundle morphisms). *Let E and F be a vector bundle over a smooth manifold M.*

$$\begin{array}{ccc} E & \xrightarrow{\Phi} & F \\ \downarrow & & \downarrow \\ M & \xrightarrow{id} & M \end{array}$$

We say Φ is a vector bundle morphism if Φ is smooth and for each $p \in M$, $\Phi|_{E_p} : E_p \to F_p$ is linear. Also,

$$\mathbf{HOM}(E, F) = \textit{the set of all vector bundle morphisms from } E \textit{ to } F.$$

Define $\Phi_* : \Gamma(E) \to \Gamma(F)$ by

$$(\Phi_*(\xi))(p) = \Phi(\xi(p)) \text{ for all } \xi \in \Gamma(E) \text{ and } p \in M.$$

Using the smoothness and linearity of Φ, Φ_* is well-defined and $\Phi_*(f \cdot \xi) = f \cdot \Phi_*(\xi)$ for $f \in C^\infty(M)$ and $\xi \in \Gamma(E)$. So, Φ_* is a $C^\infty(M)$-module homomorphism. That is,

$$\mathbf{HOM}(E, F) \hookrightarrow \Gamma(\mathbf{Hom}(E, F)).$$

Suppose that $\psi \in \Gamma(\mathbf{Hom}(E, F))$. By definition, $\Psi(p, m)$ is smooth at the first variable and linear at the second variable where $p \in M$ and $m \in E_p$. Also, the linearity at the second variable says that it is also smooth at the second variable. So, using local triviality of E, the composition of Ψ with a trivialization gives an element in $\mathbf{HOM}(E, F)$. Hence, we have one-to-one correspondence between $\mathbf{HOM}(E, F)$ and $\Gamma(\mathbf{Hom}(E, F))$.

Theorem 2.2.2.1. *ι_X are the only derivations of negative degree on $\Omega^*(M)$.*

Proof. Suppose that $\omega \in \Omega^*(M)$. Choose a partition of unity $\{f_\alpha\}$ subordinate to a coordinate covering $\{U_\alpha\}$. So,

$$\omega = \sum_\alpha f_\alpha \cdot \omega.$$

Note that $f_\alpha \cdot \omega = 0$ outside of a coordinate patch $(x_{1,\alpha}, \ldots, x_{n,\alpha})$ in $U_\alpha \subseteq \mathbb{R}^n$. Hence, using a local coordinate, we can write

$$f_\alpha \cdot \omega = \sum_{1 \le i_1 < \cdots < i_k \le n} a_{i_1 \cdots i_k}(x_{1,\alpha}, \ldots, x_{n,\alpha}) dx_{i_1,\alpha} \wedge \cdots \wedge dx_{i_k,\alpha}.$$

So, k-form is locally generated by C^∞-functions and 1-forms. Hence, Using a partition of unity, we conclude that $\Omega^*(M)$ is generated by $\Omega^0(M)$ and $\Omega^1(M)$:

$$\omega = \sum_\alpha \sum_{1 \le i_1 < \cdots < i_k \le n} a_{i_1 \cdots i_k}(x_{1,\alpha}, \ldots, x_{n,\alpha}) dx_{i_1,\alpha} \wedge \cdots \wedge dx_{i_k,\alpha}.$$

Remember what the above notation really means: By the local finiteness of a partition of unity, for $p \in M$, there exists an open neighborhood U of p such that only finite number of α covering U. So, we have

$$\omega|_U = \sum_{l=1}^{N} \sum_{1 \le i_1 < \cdots < i_k \le n} a_{i_1 \cdots i_k}(x_{1,\alpha_l}, \ldots, x_{n,\alpha_l}) dx_{i_1,\alpha_l} \wedge \cdots \wedge dx_{i_k,\alpha_l}.$$

It is obvious that any negative degree derivations is always zero on $\Omega^0(M)$, and they are zero on $\Omega^1(M)$ unless degree is -1. So, we conclude that -1 is the only negative degree of nonzero negative derivations and -1 degree derivations are uniquely defined by the value on $\Omega^1(M)$. Note that -1 degree derivations are elements of $\mathbf{HOM}(T^*M, \mathbb{R})$. So, since

$$\mathbf{HOM}(T^*M, \mathbb{R}) \Longleftrightarrow \Gamma(\mathbf{Hom}(T^*M, \mathbb{R})),$$

we deduce that -1 degree derivations are generated by $\mathfrak{X}(M)$. Therefore, ι_X are the only derivations of negative degree on $\Omega^*(M)$. $\square$

Corollary 2.2.2.1.1.

$$\iota_X \circ \iota_X = 0 \textit{ for } X \in \mathfrak{X}(M).$$

Proof. Now, by Exercise 2.2.2.1, we deduce that $[\iota_X, \iota_X] = 2\iota_X \circ \iota_X$ is a derivation of degree -2. So, Theorem 2.2.2.1 says that it must be zero. So,

$$\iota_X \circ \iota_X = 0 \text{ for } X \in \mathfrak{X}(M).$$

We can also give another proof: Since by skew symmetry of ω where $\omega \in \Omega^k(M)$, we have

$$-\omega(X, X, Y_1, \dots, Y_{k-2}) = \omega(X, X, Y_1, \dots, Y_{k-2}).$$

Hence,

$$0 = \omega(X, X, Y_1, \dots, Y_{k-2}) = ((\iota_X \circ \iota_X)\omega)(Y_1, \dots, Y_{k-2}).$$

□

Theorem 2.2.2.2. *There exists a unique derivation d of degree 1 on $\Omega^*(M)$ such that $d \circ d = 0$ and $d|_{\Omega^0(M)}$ is ordinary differential.*

Proof. Suppose that $M = \mathbb{R}^n$. Let $x_1, \dots, x_n$ be a coordinate system of $\mathbb{R}^n$. For $\omega \in \Omega^k(\mathbb{R}^n)$, define

$$\begin{aligned} d(\omega) &= d(\sum_{1 \le i_1 < \cdots < i_k \le n} f_{i_1 \cdots i_k}(x_1, \dots, x_n) dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \\ &\overset{def}{=} \sum_{1 \le i_1 < \cdots < i_k \le n} df_{i_1 \cdots i_k}(x_1, \dots, x_n) \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &\overset{def}{=} \sum_{1 \le i_1 < \cdots < i_k \le n} \sum_{l=1}^{n} \frac{\partial}{\partial x_l} f_{i_1 \cdots i_k}(x_1, \dots, x_n) dx_l \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \end{aligned}$$

It is easy to see that d is a derivation of degree 1 on $\Omega^*(\mathbb{R}^n)$ such that $d \circ d = 0$ and $d|_{\Omega^0(\mathbb{R}^n)}$ is ordinary differential. Now suppose that d' is another derivation of degree 1 on $\Omega^*(\mathbb{R}^n)$ such that $d' \circ d' = 0$ and $d'|_{\Omega^0(\mathbb{R}^n)}$ is ordinary differential. Note that we have $d'f = df$ for $f \in C^\infty(\mathbf{R}^n)$, since they are just ordinary differential by the assumptions. In particular,

$$dx_i = d'x_i \text{ and } d' \circ dx_i = d' \circ d'x_i = 0 \text{ for } i = 1, \dots, n.$$

So, we have

$$\begin{aligned} d'(\omega) &= d'(\sum_{1 \le i_1 < \cdots < i_k \le n} f_{i_1 \cdots i_k}(x_1, \dots, x_n) dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \\ &= \sum_{1 \le i_1 < \cdots < i_k \le n} d' f_{i_1 \cdots i_k}(x_1, \dots, x_n) \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &+ \sum_{1 \le i_1 < \cdots < i_k \le n} f_{i_1 \cdots i_k}(x_1, \dots, x_n) \cdot d'(dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \\ &= \sum_{1 \le i_1 < \cdots < i_k \le n} \sum_{l=1}^{n} \frac{\partial}{\partial x_l} f_{i_1 \cdots i_k}(x_1, \dots, x_n) d'x_l \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &= \sum_{1 \le i_1 < \cdots < i_k \le n} \sum_{l=1}^{n} \frac{\partial}{\partial x_l} f_{i_1 \cdots i_k}(x_1, \dots, x_n) dx_l \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} = d\omega \end{aligned}$$

Hence, at least in $\mathbb{R}^n$, we have a unique derivation of degree 1 satisfying the properties with respect to a given coordinate system. It will be proved that a derivation is invariant under changes of coordinates in the setting of a manifold, more generally. Now, let $\{U_\alpha\}$ be an open cover consisting of coordinate patches, i.e., $\varphi_\alpha = (x_{1,\alpha}, \dots, x_{n,\alpha})$. So, we do have a unique derivation d_{U_α} satisfying the properties on each open set U_α. Now suppose that $\omega \in \Omega^k(U_\alpha \cap U_\beta)$. So, on $U_\alpha \cap U_\alpha$, ω has two representations, namely,

$$\begin{aligned}\omega &= \sum_{1 \le i_1 < \cdots < i_k \le n} f_{i_1 \cdots i_k}(x_{1,\alpha}, \dots, x_{n,\alpha}) d_{U_\alpha} x_{i_1,\alpha} \wedge \cdots \wedge d_{U_\alpha} x_{i_k,\alpha} \\ &= \sum_{1 \le i_1 < \cdots < i_k \le n} g_{i_1 \cdots i_k}(x_{1,\beta}, \dots, x_{n,\beta}) d_{U_\beta} x_{i_1,\beta} \wedge \cdots \wedge d_{U_\beta} x_{i_k,\beta}.\end{aligned}$$

Choosing a bump function we can make $\omega_\alpha \in \Omega^k(U_\alpha)$ with $\omega_\alpha = \omega$ on $U_\alpha \cap U_\beta$ and $\omega_\alpha \equiv 0$ on $U_\alpha/(U_\alpha \cap U_\beta)$. By the same reason, we have such $\omega_\beta \in \Omega^k(U_\beta)$. Let $p \in U_\alpha \cap U_\beta$. We want to show

$$d_{U_\alpha}(\omega_\alpha)_p = d_{U_\beta}(\omega_\beta)_p.$$

If you are brave enough, you can prove this using k-form. However, fortunately, if we use the linearity and the properties of a derivation, it is sufficient to prove this for 0-forms and 1-forms. Suppose $\omega = f \in \Omega^0(U_\alpha \cap U_\beta)$. So,

$$\text{we have} \begin{cases} \omega_\alpha = f \circ \varphi_\alpha^{-1}(x_{1,\alpha}, \dots, x_{n,\alpha}) \overset{def}{=} f(x_{1,\alpha}, \dots, x_{n,\alpha}) \\ \omega_\beta = f \circ \varphi_\beta^{-1}(x_{1,\beta}, \dots, x_{n,\beta}) \overset{def}{=} f(x_{1,\beta}, \dots, x_{n,\beta}). \end{cases}$$

Note that if we think $x_{i,\alpha}$ as a function with coordinates $x_{1,\beta}, \dots, x_{n,\beta}$, then by the fact d_{U_α} is ordinary differential on functions, we have

$$d_{U_\alpha} x_{i,\alpha} = dx_{i,\alpha}(x_{1,\beta}, \dots, x_{n,\beta}) = \sum_{k=1}^n \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}} dx_{k,\beta} = \sum_{k=1}^n \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}} d_{U_\beta} x_{k,\beta}$$

Hence, we have

$$\begin{aligned} d_{U_\alpha}(\omega_\alpha)_p &= d_{U_\alpha} f(x_{1,\alpha}, \dots, x_{n,\alpha})|_p = \sum_{l=1}^n \frac{\partial}{\partial x_{l,\alpha}} f(x_{1,\alpha}, \dots, x_{n,\alpha}) d_{U_\alpha} x_{l,\alpha}|_p \\ &= \sum_{l=1}^n \sum_{k=1}^n \frac{\partial}{\partial x_{l,\alpha}} f(x_{1,\alpha}, \dots, x_{n,\alpha}) \frac{\partial x_{l,\alpha}}{\partial x_{k,\beta}} d_{U_\beta} x_{k,\beta}|_p = \sum_{k=1}^n \frac{\partial}{\partial x_{k,\beta}} f(x_{1,\beta}, \dots, x_{n,\beta}) d_{U_\beta} x_{k,\beta}|_p \\ &= d_{U_\beta} f(x_{1,\beta}, \dots, x_{n,\beta})|_p = d_{U_\beta}(\omega_\beta)_p \end{aligned}$$

Suppose $\omega \in \Omega^1(U_\alpha \cap U_\beta)$. Without loss of generality, we can assume $\omega_\alpha = f(x_{1,\alpha}, \dots, x_{n,\alpha}) d_{U_\alpha} x_{i,\alpha}$. So, we also have

$$\omega_\beta = \sum_{k=1}^n g_k(x_{1,\beta}, \dots, x_{n,\beta}) d_{U_\beta} x_{k,\beta} \text{ where } f(x_{1,\alpha}, \dots, x_{n,\alpha}) \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}} = g_k(x_{1,\beta}, \dots, x_{n,\beta})$$

So, for $p \in U_\alpha \cap U_\beta$.

$$\begin{aligned}
d_{U_\alpha}(\omega_\alpha)_p &= d_{U_\alpha}(f(x_{1,\alpha},\ldots,x_{n,\alpha})d_{U_\alpha}x_{i,\alpha})|_p = d_{U_\alpha}f(x_{1,\alpha},\ldots,x_{n,\alpha}) \wedge d_{U_\alpha}x_{i,\alpha}|_p \\
&= d_{U_\alpha}f(x_{1,\alpha},\ldots,x_{n,\alpha}) \wedge \sum_{k=1}^{n} \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}} d_{U_\beta}x_{k,\beta}|_p = d_{U_\beta}f(x_{1,\beta},\ldots,x_{n,\beta}) \wedge \sum_{k=1}^{n} \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}} d_{U_\beta}x_{k,\beta}|_p \\
&= d_{U_\beta}(f(x_{1,\beta},\ldots,x_{n,\beta}) \cdot \sum_{k=1}^{n} \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}} d_{U_\beta}x_{k,\beta})|_p - f(x_{1,\beta},\ldots,x_{n,\beta}) \cdot d_{U_\beta}(\sum_{k=1}^{n} \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}}) \wedge d_{U_\beta}x_{k,\beta}|_p \\
&= d_{U_\beta}(f(x_{1,\beta},\ldots,x_{n,\beta}) \sum_{k=1}^{n} \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}} d_{U_\beta}x_{k,\beta})|_p - f(x_{1,\beta},\ldots,x_{n,\beta}) \sum_{l=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 x_{i,\alpha}}{\partial x_{k,\beta}\partial x_{l,\beta}} d_{U_\beta}x_{l,\beta} \wedge d_{U_\beta}x_{k,\beta}|_p \\
&= d_{U_\beta}(f(x_{1,\beta},\ldots,x_{n,\beta}) \cdot \sum_{k=1}^{n} \frac{\partial x_{i,\alpha}}{\partial x_{k,\beta}} d_{U_\beta}x_{k,\beta})|_p = d_{U_\beta}(\omega_\beta)_p.
\end{aligned}$$

Note that by the skew-symmetry,

$$f(x_{1,\beta},\ldots,x_{n,\beta}) \cdot \sum_{l=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 x_{i,\alpha}}{\partial x_{k,\beta}\partial x_{l,\beta}} d_{U_\beta}x_{l,\beta} \wedge d_{U_\beta}x_{k,\beta} = 0.$$

From this, we have a well-defined notion of a derivation d_M by setting for $\omega \in \Omega^k(M)$

$$(d_M\omega)|_{U_\alpha} = d_{U_\alpha}\omega|_{U_\alpha} \text{ where } \omega|_{U_\alpha} \text{ is the restriction of } \omega \text{ on } U_\alpha.$$

It is easy to check that d_M satisfies all the required properties. We will show the uniqueness of d_M. Suppose that d'_M is another derivation with the properties. The proof shall use the uniqueness of a local derivation, since we already know that each point of M has a unique local derivation. Let $p \in V \subseteq M$ be an open set with a unique local derivation d_V and $\omega \in \Omega^k(V)$. So, we have ω^* the extention of ω by a bump function, i.e.,

$$\begin{aligned}
&\omega^* = \omega \text{ on } V \\
&\omega^* \equiv 0 \text{ on open } U \text{ where } V \subseteq \overline{V} \subseteq U.
\end{aligned}$$

The main point is that we do not know whether or not $d'\omega^*|_{M/\overline{V}} \equiv 0$. If this is true, then by the facts that d' is a derivation with the properties and the uniqueness of d_V, we have

$$(d'\omega^*)|_V = d_V\omega. \text{ Hence, } d'_M = d_M.$$

So, it suffices to show that if k-form $\omega|_W \equiv 0$ for an open set $W \subseteq M$, then $(d'\omega)|_W \equiv 0$. Suppose that k-form $\omega|_W \equiv 0$ on an open set $W \subseteq M$. Choose a bump function $\varphi \in C^\infty(M)$ such that

$$\varphi(x) = 0 \text{ on a compact } K \subset W \text{ and } \varphi|_{M/W} \equiv 1.$$

So, $\varphi\omega = \omega$ on M. By the property of a derivation, we have

$$d'\omega = d'(\varphi\omega) = (d'\varphi)\omega + \varphi d'\omega.$$

So, we have $(d'\omega)|_K \equiv 0$. By letting $K = \{p\}$ for $p \in W$, we have $(d'\omega)|_W \equiv 0$. □

Remark 2.2.2.1. *It is easy to see that $\Omega^*(M)$ as a graded algebra is generated by $\Omega^0(M)$ and $d\Omega^0(M)$. If M is compact, then $d\Omega^0(M)$ has a finite basis.*

Theorem 2.2.2.3. *Let $M^m \xrightarrow{\varphi} N^n$ be a smooth map. Then φ induces a graded algebra homomorphism with the following commuting diagram*

$$\begin{array}{ccc} \Omega^k(N) & \xrightarrow{\varphi^k} & \Omega^k(M) \\ d\Big\downarrow & & d\Big\downarrow \\ \Omega^{k+1}(N) & \xrightarrow{\varphi^{k+1}} & \Omega^{k+1}(M). \end{array}$$

That is, $d(\varphi^*\omega) = \varphi^*(d\omega)$ for $\omega \in \Omega^*(N)$.

Proof. Note that

$$(\varphi^*\omega)_p \overset{def}{=} (J\varphi)_p^T(\omega_{\varphi(p)}) \overset{def}{=} (J\varphi)^T \circ \omega \circ \varphi \text{ where } J\varphi \text{ is the Jacobian of } \varphi.$$

When $f \in \Omega^0(N)$, $\varphi^* f \overset{def}{=} f \circ \varphi$. So, we have for $\omega \in \Omega^k(N)$ and $X^1, \ldots, X^k \in \mathfrak{X}(M)$,

$$(\varphi^*\omega)_p(X_p^1, \ldots, X_p^k) = \omega_{\varphi(p)}(J\varphi(X_p^1)_{\varphi(p)}, \ldots, J\varphi(X_p^k)_{\varphi(p)}).$$

Again, the right-hand side is a smooth function as p varies. So, $\varphi^*\omega$ is smooth as p varies. Hence, $\varphi^* : \Omega^*(N) \to \Omega^*(M)$. We shall show that $(\varphi^*(\omega \wedge \xi))_p = (\varphi^*\omega)_p \wedge (\varphi^*\xi)_p$ for all $p \in M$, which implies that $\varphi^*(\omega \wedge \xi) = (\varphi^*\omega) \wedge (\varphi^*\xi)$. Let $\omega \in \Omega^k(N)$ and $\xi \in \Omega^r(N)$.

$$\begin{aligned}
&(\varphi^*(\omega \wedge \xi))_p(X_p^1, \ldots, X_p^k, X_p^{k+1}, \ldots, X_p^{k+r}) \\
&\qquad = \omega_{\varphi(p)}(J\varphi(X_p^1)_{\varphi(p)}, \ldots, J\varphi(X_p^k)_{\varphi(p)}, J\varphi(X_p^{k+1})_{\varphi(p)}, \ldots, J\varphi(X_p^{k+r})_{\varphi(p)}) \\
&\qquad = \omega_{\varphi(p)}(J\varphi(X_p^1)_{\varphi(p)}, \ldots, J\varphi(X_p^k)_{\varphi(p)}) \cdot \xi_{\varphi(p)}(J\varphi(X_p^{k+1})_{\varphi(p)}, \ldots, J\varphi(X_p^{k+r})_{\varphi(p)}) \\
&\qquad = (\varphi^*\omega)_p(X_p^1, \ldots, X_p^k) \cdot (\varphi^*\xi)_p(X_p^{k+1}, \ldots, X_p^{k+r}) \\
&\qquad = (\varphi^*\omega)_p \wedge (\varphi^*\xi)_p(X_p^1, \ldots, X_p^k, X_p^{k+1}, \ldots, X_p^{k+r}).
\end{aligned}$$

Also, note that for $f \in \Omega^0(N)$,

$$(\varphi^*(f \cdot \omega))_p(X_p) = (f \cdot \omega)_{\varphi(p)}(J\varphi(X_p)_{\varphi(p)}) = (f \circ \varphi(p)) \cdot (\varphi^*\omega)_p(X_p) = (\varphi^* f)_p \cdot (\varphi^*\omega)_p(X_p).$$

Hence, φ^* is a graded algebra homomorphism. Now, we shall show that $\varphi^*(d\omega) = d(\varphi^*\omega)$. Since we made d by patching local derivations, it is sufficient to show that $\varphi^*(d\omega) = d(\varphi^*\omega)$ locally. Moreover, it is also sufficient to show this for bases. Let $U \subset M$ and $V \subset N$ be coordinate patches with $(x_1, \ldots, x_m)$ and $(y_1, \ldots, y_n)$, respectively. Assume that $\varphi : U \to V$. Note that $d|_{\Omega^0(N)}$ is ordinary differential. That is, if $g \in \Omega^0(N)$, then we have $dg(Y_q) = Y_q(g)$ for $Y \in \mathfrak{X}(N)$. So, we have

$$\begin{aligned}
\varphi^*(dg(y_1, \ldots, y_n))_p(X_p) &= dg_{\varphi(p)}(J\varphi(X_p))_{\varphi(p)} = (J\varphi(X_p))_{\varphi(p)}(g) = (X_p)(g \circ \varphi) \\
&= (X_p)(\varphi^* g) = (d(\varphi^* g))_p(X_p).
\end{aligned}$$

That is, for $g \in \Omega^0(N)$ we have $\varphi^*(dg) = d(\varphi^* g)$. In general, if we think $y_i \in \Omega^0(N)$ for $i = 1, \ldots, n$, then using the fact that φ^* is a graded algebra homomorphism, we have

$$\begin{aligned}
\varphi^*(d(f(y_1, \ldots, y_n)dy_{i_1} \wedge \cdots \wedge dy_{i_k})) &= \varphi^*(df \wedge dy_{i_1} \wedge \cdots \wedge y_{i_k}) \\
&= \varphi^*(df) \wedge \varphi^*(dy_{i_1}) \wedge \cdots \wedge \varphi^*(dy_{i_k}) \\
&= d(\varphi^* f) \wedge d(\varphi^* y_{i_1}) \wedge \cdots \wedge d(\varphi^* dy_{i_k}) \\
&= d(\varphi^* f \wedge d(\varphi^* y_{i_1}) \wedge \cdots \wedge d(\varphi^* y_{i_k})) \\
&= d(\varphi^* f \wedge \varphi^*(dy_{i_1}) \wedge \cdots \wedge \varphi^*(dy_{i_k})) \\
&= d(\varphi^*(f \wedge dy_{i_1} \wedge \cdots \wedge dy_{i_k})).
\end{aligned}$$

□

Recall that if $X, Y \in \mathfrak{X}(M)$ and $\xi_t \in \mathbf{Diff}(M)$ is a global flow of X, we have

$$\mathcal{L}_X(Y) \stackrel{def}{=} \frac{d}{dt}|_{t=0}(\xi_{-t} \cdot Y).$$

Now we define for $\omega \in \Omega^*(M)$,

$$\mathcal{L}_X\omega \stackrel{def}{=} \frac{d}{dt}|_{t=0}(\xi_{-t} \cdot \omega) \text{ where } \xi_{-t} \cdot \omega \stackrel{def}{=} \xi_t^*\omega \text{ the pullback of } \omega.$$

Note that ξ_t^* is a path of 1-forms. It is easy to see that $\mathcal{L}_X$ is a derivation on $\Omega^*(M)$: Since $\wedge$ is bilinear and ξ_0 is the identity map, we have

$$\mathcal{L}_X(\alpha \wedge \beta) = \frac{d}{dt}|_{t=0}\xi_t^*(\alpha \wedge \beta) = \frac{d}{dt}|_{t=0}(\xi_t^*\alpha \wedge \xi_t^*\beta) = (\mathcal{L}_X\alpha) \wedge \beta + \alpha \wedge \mathcal{L}_X\beta.$$

Also, it is easy to see that the degree of $\mathcal{L}_X$ is 0:

$$\mathcal{L}_X : \Omega^p(M) \to \Omega^p(M).$$

In particular, for $f \in \Omega^0(M)$

$$\mathcal{L}_X f = \frac{d}{dt}|_{t=0}\xi_t^* f = \frac{d}{dt}|_{t=0} f \circ \xi_t = Xf.$$

Moreover, we have using the linearity of d

$$\mathcal{L}_X(d\omega) = \frac{d}{dt}|_{t=0}\xi_t^*(d\omega) = \frac{d}{dt}|_{t=0}d(\xi_t^*\omega) \stackrel{linearity}{=} d\frac{d}{dt}|_{t=0}\xi_t^*\omega = d(\mathcal{L}_X\omega).$$

Hence, we have

$$[d, \mathcal{L}_X] = d\mathcal{L}_X - \mathcal{L}_X d = 0.$$

Remark 2.2.2.2. *From the above, $X \in \mathfrak{X}(M)$ gives a derivation $\mathcal{L}_X \in Der_0(\Omega^*(M))$. Also, it is worth noting that $X \in \mathfrak{X}(M)$ also gives a derivation on $\mathfrak{X}(M)$ by the following ways: Define*

$$\mathbf{ad}(X) : \mathfrak{X}(M) \to \mathfrak{X}(M) \text{ by } \mathbf{ad}(X)(Y) = [X, Y].$$

$\mathbf{ad}(X)$ *is a derivation, since by the Jacobi identity, we have*

$$\mathbf{ad}(X)([Y,Z]) = [X,[Y,Z]] = [Y,[X,Z]] + [[X,Y],Z] = [Y,\mathbf{ad}(X)(Z)] + [\mathbf{ad}(X)(Y),Z].$$

Hence, we have

$$\begin{array}{ccc} \mathfrak{X}(M) & \xrightarrow{\mathbf{ad}} & Der_0(\mathfrak{X}(M)) \\ \mathcal{L}\Big\downarrow & & \\ Der_0(\Omega^*(M)). & & \end{array}$$

In general, if $\mathfrak{g}$ *is an abstract Lie algebra, for* $X \in \mathfrak{g}$, $\mathbf{ad}(X) : \mathfrak{g} \to \mathfrak{g}$ *by* $\mathbf{ad}(X)(Y) = [X,Y]$ *is a derivation of* $\mathfrak{g}$ *by the same proof of the above. That is, given* $\mathfrak{g}$*, there exists a Lie algebra homomorphism*

$$\mathbf{ad} : \mathfrak{g} \to Der(\mathfrak{g}).$$

2.2.3 The Cartan formula and Useful identities

Recall that ι_X is a unique derivation of degree -1 and d is a unique derivation of degree 1 on $\Omega^*(M)$. Since $Der(\Omega^*(M))$ forms a graded Lie algebra, we have $[d,\iota_X] = d \circ \iota_X + \iota_X \circ d$ is a derivation of degree 0. We shall prove

Theorem 2.2.3.1 (Cartan Formula).

$$[d,\iota_X] = \mathcal{L}_X.$$

Note that the Cartan formula says that $\mathcal{L}_X$ is a chain homotopy. Hence, $\mathcal{L}_X$ induces an isomorphism on de Rham cohomology. Moreover, in general if $\mathfrak{g}$ is a Lie algebra, then it induces a differential graded algebra $\wedge^*\mathfrak{g}^*$. So, we have a Lie algebra cohomology

$$H^*(\wedge^*\mathfrak{g}^*) \stackrel{def}{=} H^*(\mathfrak{g}).$$

Note that if G is a compact Lie group, then we have

$$H^*(\mathfrak{g}) \cong H^*(G,\mathbb{R}).$$

Proof of Cartan Formula. Since $\Omega^*(M)$ is generated by $\Omega^0(M)$ and $d\Omega^0(M)$, it suffices to show this for $f \in \Omega^0(M)$ and df. First, Note $\iota_X f = 0$, since the degree of ι_X is -1. So,

$$(d \circ \iota_X + \iota_X \circ d)f = d \circ \iota_X f + \iota_X \circ df = \iota_X(df) = df(X) = Xf = \mathcal{L}_X f.$$

For $\omega = df \in d\Omega^0(M)$, we have by $[\mathcal{L}_X, d] = 0$, i.e., the commutativity of $\mathcal{L}_X$ and d,

$$(d \circ \iota_X + \iota_X \circ d)\omega = d \circ \iota_X df + \iota_X \circ d \circ df = d \circ \iota_X df = d(Xf) = d(\mathcal{L}_X f) = \mathcal{L}_X(df) = \mathcal{L}_X\omega.$$

□

Theorem 2.2.3.2.

$$[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_{[X,Y]}.$$

Proof. As usual, it suffices to show this for $f \in \Omega^0(M)$ and df. Clearly,

$$[\mathcal{L}_X, \mathcal{L}_Y]f = \mathcal{L}_X\mathcal{L}_Y f - \mathcal{L}_Y\mathcal{L}_X f = X(Yf) - Y(Xf) = (XY - YX)f = \mathcal{L}_{[X,Y]}f.$$

For $\omega = df \in d\Omega^0(M)$, we have by the commutativity of $\mathcal{L}_X$ and d,

$$[\mathcal{L}_X, \mathcal{L}_Y]df = \mathcal{L}_X\mathcal{L}_Y df - \mathcal{L}_Y\mathcal{L}_X df = d(\mathcal{L}_X\mathcal{L}_Y f - \mathcal{L}_Y\mathcal{L}_X f) = d\mathcal{L}_{[X,Y]}f = \mathcal{L}_{[X,Y]}df.$$

□

Theorem 2.2.3.3. *Let $\omega \in \Omega^1(M)$ and $X, Y \in \mathfrak{X}(M)$. Then we have*

$$(d\omega)(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y]).$$

Before we prove Theorem 2.2.3.3, we give some remarks: We know that $\omega \in \Omega^k(M)$ defines a k-multilinear alternating $C^\infty(M)$:

$$\omega : \underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{k} \to C^\infty(M) \text{ by } (Y_1, \ldots, Y_k) \mapsto \omega(Y_1, \ldots, Y_k).$$

Moreover, k-multilinearity over $\mathbb{R}$ gives k-multilinearity over $C^\infty(M)$. That is,

$$\omega(fY_1, \ldots, Y_k) = \omega(Y_1, \ldots, fY_i, \ldots, Y_k) = f\omega(Y_1, \ldots, Y_k).$$

Note that smoothness comes from the left-hand side. Equivalently speaking, since we can think $\otimes^k TM$ and $\mathbb{R}$ as vector bundles, in the sense of Definition 2.2.2.2, we can say $\omega \in \Gamma(\mathbf{Hom}(\otimes^k TM, \mathbb{R}))$. However, note that since

$$[fX, gY] = f \cdot X(g)Y - g \cdot Y(f)X + f \cdot g[X, Y],$$

We can see that $[,]$ is not a section of $\mathbf{Hom}(\otimes^2 TM, TM)$, i.e., $[,] \notin \Gamma(\mathbf{Hom}(\otimes^2 TM, TM))$. In this consideration, we have

Definition 2.2.3.1. *Let E and F be smooth vector bundle over a smooth manifold M. we say ψ is tensorial if $\psi \in \Gamma(\mathbf{Hom}(E, F))$.*

Example 2.2.3.1. *If $\omega \in \Omega^k(M)$, then $\iota_X\omega$ is a $(k-1)$-form and $d\omega$ is a $(k+1)$-from. So, these are tensorial.*

Proof. We give ad hoc proofs of these. Clearly,

$$(\iota_X\omega)(fY_1, \ldots, Y_{k-1}) = \omega(X, fY_1, \ldots, Y_{k-1}) = f\omega(X, Y_1, \ldots, Y_{k-1}).$$

Also, using Theorem 2.2.3.3, we have for 1-form ω,

$$\begin{aligned} d\omega(fX, Y) &= (fX)(\omega(Y)) - Y(\omega(fX)) - \omega([fX, Y]) = f(X(\omega(Y))) - Y(f\omega(X)) - \omega([fX, Y]) \\ &= f(X(\omega(Y))) - Y(f\omega(X)) - \omega(f[X, Y] - (Yf)X) \\ &= f(X(\omega(Y))) - (Yf)(\omega(X)) - fY(\omega(X)) - f\omega([X, Y]) + (Yf)(\omega X) \\ &= fd\omega(X, Y). \end{aligned}$$

□

Lemma 2.2.3.1. *Let $\varphi \in \mathbf{Diff}(M)$ and $\omega \in \Omega^k(M)$. Then*

$$\varphi \cdot (\omega(Y_1, \dots, Y_k)) = (\varphi \cdot \omega)(\varphi \cdot Y_1, \dots, \varphi \cdot Y_k).$$

Proof. When $k = 0$, clearly we have $\varphi \cdot (fg) = (\varphi \cdot f)(\varphi \cdot g)$ for $f, g \in C^\infty(M)$. So, now assume ω is a 1-form. Without loss of generality, we can assume that $\omega = df$ for some $f \in C^\infty(M)$. Note that above Remark 2.1.5.2 we showed that $\varphi \cdot (Xf) = (\varphi \cdot X)(\varphi \cdot f)$ and in Theorem 2.2.2.3 we showed that d and φ^* commute each other.

$$\begin{aligned}\varphi \cdot (\omega(X)) = \varphi \cdot (df(X)) = \varphi \cdot (Xf) = (\varphi \cdot X)(\varphi \cdot f) = (d(\varphi \cdot f))(\varphi \cdot X)\\ = (d((\varphi^{-1})^* f))(\varphi \cdot X) = ((\varphi^{-1})^* df)(\varphi \cdot X) = (\varphi \cdot \omega)(\varphi \cdot X).\end{aligned}$$

In general, note that the action of a k-form on k vector fields can be decomposed into the action of $k-1$-forms on $k-1$ vector fields and the action of 1 forms on a vector fields. So, using induction hypothesis and the cases when $k = 0$ and $k = 1$, it is easy to see that it is true. □

Lemma 2.2.3.2. *Let $\xi_t \in \mathbf{Diff}(M)$. Then we have for $\omega \in \Omega^k(M)$*

$$\begin{aligned}&\frac{d}{dt}|_{t=0}\big((\xi_{-t} \cdot \omega)(\xi_{-t} \cdot Y_1, \dots, \xi_{-t} \cdot Y_k)\big) = (\frac{d}{dt}|_{t=0}(\xi_{-t} \cdot \omega))(\xi_{-t} \cdot Y_1, \dots, \xi_{-t} \cdot Y_k)|_{t=0}\\ &\quad + \sum_{i=1}^{k} (\xi_{-t} \cdot \omega))|_{t=0}(\xi_{-t} \cdot Y_1|_{t=0}, \dots, \xi_{-t} \cdot Y_{i-1}|_{t=0}, \frac{d}{dt}|_{t=0}\xi_{-t}Y_i, \xi_{-t} \cdot Y_{i+1}|_{t=0}, \dots, \xi_{-t} \cdot Y_k|_{t=0}).\end{aligned}$$

Proof. The main point is that the action of a k-from on k vector fields is k-multilinear over $\mathbb{R}$. So, it is sufficient to prove this for 1-form. Recall the proof of Theorem 2.1.6.1.

$$\begin{aligned}\frac{d}{dt}|_{t=0}(\xi_t \cdot \omega)(\xi_t \cdot X) &= -\lim_{t\to 0} \frac{(\xi_t \cdot \omega)(\xi_t \cdot X) - \omega X}{t}\\ &= \lim_{t\to 0} \frac{(\xi_t \cdot \omega)(\xi_t \cdot X) - (\xi_t \cdot \omega)(X) + (\xi_t \cdot \omega)(X) - \omega X}{t}\\ &= \lim_{t\to 0} \frac{1}{t}(\xi_t \cdot \omega)((\xi_t \cdot X) - X) + \lim_{t\to 0} \frac{(\xi_t \cdot \omega - \omega)(X)}{t}\\ &= \lim_{t\to 0}(\xi_t \cdot \omega)(\frac{(\xi_t \cdot X) - X}{t}) + \lim_{t\to 0} \frac{(\xi_t \cdot \omega - \omega)(X)}{t}\\ &= \big((\xi_t \cdot \omega)|_{t=0}\big)\frac{d}{dt}|_{t=0}(\xi_t \cdot X) + \frac{d}{dt}|_{t=0}(\xi_t \cdot \omega)\big((\xi_t \cdot X)|_{t=0}\big).\end{aligned}$$

□

From these lemmas, we have the following lemma. Especially, we give two proofs of the following:

Lemma 2.2.3.3. *For $\omega \in \Omega^k(M)$, we have*

$$\mathcal{L}_X(\omega(Y_1, \dots, Y_k)) = (\mathcal{L}_X\omega)(Y_1, \dots, Y_k) + \sum_{i=1}^{k} \omega(Y_1, \dots, Y_{i-1}, [X, Y_i], Y_{i+1}, \dots, Y_k).$$

Proof 1. Consider a function $\omega(Y_1, \ldots, Y_k)$ and its Lie derivative: Let ξ_t be the global flow of X. Note that ξ_0 is the identity. Using Lemma 2.2.3.1 and Lemma 2.2.3.2,

$$\begin{aligned}
\mathcal{L}_X(\omega(Y_1, \ldots, Y_k)) &= \frac{d}{dt}|_{t=0}\xi_{-t} \cdot \omega(Y_1, \ldots, Y_k) = \frac{d}{dt}|_{t=0}(\xi_{-t} \cdot \omega)(\xi_{-t} \cdot Y_1, \ldots, \xi_{-t} \cdot Y_k) \\
&= (\frac{d}{dt}|_{t=0}(\xi_{-t} \cdot \omega))(Y_1, \ldots, Y_k) + \sum_{i=1}^{k} \omega(Y_1, \ldots, Y_{i-1}, \frac{d}{dt}|_{t=0}\xi_{-t}Y_i, Y_{i+1}, \ldots, Y_k) \\
&= (\mathcal{L}_X\omega)(Y_1, \ldots, Y_k) + \sum_{i=1}^{k} \omega(Y_1, \ldots, Y_{i-1}, \mathcal{L}_X Y_i, Y_{i+1}, \ldots, Y_k).
\end{aligned}$$

□

Proof 2. First, we show the following: Since $\mathcal{L}_X$ and ι_Y are derivations on $\Omega^*(M)$ for $X, Y \in \mathfrak{X}(M)$, we have a derivation $[\mathcal{L}_X, \iota_Y]$. We shall show

$$[\mathcal{L}_X, \iota_Y] = \iota_{[X,Y]}.$$

As usual, it suffices to show this for 1-forms and 0-forms. Since $[\mathcal{L}_X, \iota_Y]$ and $\iota_{[X,Y]}$ are degree -1, they give the same action on $\Omega^0(M)$, namely, a zero derivation. So, it's done. Now, assume $\omega = df$ for $f \in \Omega^0(M)$. The main ingredient of the proof will be $\iota_X \circ \iota_Y = 0$ on $\Omega^0(M)$ and $\Omega^1(M)$, and $\mathcal{L}_X = \iota_X \circ d + d \circ \iota_X$. We have

$$\begin{aligned}
\iota_{[X,Y]}df &= (\mathcal{L}_{[X,Y]} - d \circ \iota_{[X,Y]})f = \mathcal{L}_{[X,Y]}f = \mathcal{L}_X \circ (\iota_Y \circ d + d \circ \iota_Y)f - \mathcal{L}_Y \circ (\iota_X \circ d + d \circ \iota_X)f \\
&= (\mathcal{L}_X \circ \iota_Y \circ d)f - (\mathcal{L}_Y \circ \iota_X \circ d)f = (\mathcal{L}_X \circ \iota_Y - \mathcal{L}_Y \circ \iota_X) \circ df \\
&= (\mathcal{L}_X \circ \iota_Y - \iota_Y \circ \mathcal{L}_X + \iota_Y \circ \mathcal{L}_X - \mathcal{L}_Y \circ \iota_X)df = [\mathcal{L}_X, \iota_Y]df + (\iota_Y \circ \mathcal{L}_X - \mathcal{L}_Y \circ \iota_X)df \\
&= [\mathcal{L}_X, \iota_Y]df + (\iota_Y \circ (\iota_X \circ d + d \circ \iota_X) - (\iota_Y \circ d + d \circ \iota_Y) \circ \iota_X)df \\
&= [\mathcal{L}_X, \iota_Y]df + (\iota_Y \circ d \circ \iota_X)df - (\iota_Y \circ d \circ \iota_X)df = [\mathcal{L}_X, \iota_Y]df.
\end{aligned}$$

From this, now we have

$$\begin{aligned}
(\mathcal{L}_X\omega)(Y_1, \ldots, Y_k) &= (\iota_{Y_k} \circ \cdots \circ \iota_{Y_1} \circ \mathcal{L}_X)\omega = (\iota_{Y_k} \circ \cdots \circ \iota_{Y_2} \circ (\mathcal{L}_X \circ \iota_{Y_1} - \iota_{[X,Y_1]})\omega \\
&= (\iota_{Y_k} \circ \cdots \circ \iota_{Y_2} \circ \mathcal{L}_X \circ \iota_{Y_1})\omega - (\iota_{Y_k} \circ \cdots \circ \iota_{Y_2} \circ \iota_{[X,Y_1]})\omega \\
&= (\iota_{Y_k} \circ \cdots \circ \iota_{Y_3} \circ (\mathcal{L}_X \circ \iota_{Y_2} - \iota_{[X,Y_2]}) \circ \iota_{Y_1})\omega - (\iota_{Y_k} \circ \cdots \circ \iota_{Y_2} \circ \iota_{[X,Y_1]})\omega \\
&= (\iota_{Y_k} \circ \cdots \circ \iota_{Y_3} \circ \mathcal{L}_X \circ \iota_{Y_2} \circ \iota_{Y_1})\omega - \omega([X, Y_1], Y_2, \ldots, Y_k) - \omega(Y_1, [X, Y_2], Y_3, \ldots, Y_k) \\
&\vdots \\
&= (\mathcal{L}_X \circ \iota_{Y_k} \circ \cdots \circ \iota_{Y_1})\omega - \sum_{i=1}^{k} \omega(Y_1, \ldots, Y_{i-1}, [X, Y_i], Y_{i+1}, \ldots, Y_k) \\
&= \mathcal{L}_X(\omega(Y_1, \ldots, Y_k)) - \sum_{i=1}^{k} \omega(Y_1, \ldots, Y_{i-1}, [X, Y_i], Y_{i+1}, \ldots, Y_k)
\end{aligned}$$

□

Now we prove Theorem 2.2.3.3:

Proof of Theorem 2.2.3.3. Let $\omega \in \Omega^1(M)$ and $X, Y \in \mathfrak{X}(M)$. Note that for $f \in C^\infty(M)$, we have $df(X) = Xf$. In particular, $(d(\omega X))Y = Y(\omega(X))$. Using Lemma 2.2.3.3,

$$\begin{aligned} X(\omega(Y)) &= \mathcal{L}_X(\omega(Y)) = (\mathcal{L}_X\omega)(Y) + \omega([X,Y]) = ((\iota_X \circ d + d \circ \iota_X)\omega)(Y) + \omega([X,Y]) \\ &= (\iota_X d\omega)Y + (d\omega(X))(Y) + \omega([X,Y]) = (d\omega)(X,Y) + Y(\omega(X)) + \omega([X,Y]). \end{aligned}$$

□

In general we have

Theorem 2.2.3.4.

$$(d\omega)(X_0, \ldots, X_k) = \sum_{i=0}^{k}(-1)^i X_i\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k) + \sum_{0 \le i < j \le k}(-1)^{i+j}\omega([X_i, X_j], X_0, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k).$$

Proof. We shall prove this by induction. When $k = 0, 1$, we are already done. Now suppose $k = n$. The main ingredients of this proof are of course Lemma 2.2.3.3 and the identity $[d, \iota_X] = \mathcal{L}_X$.

$$\begin{aligned} (d\omega)(X_0, \ldots, X_n) &= (\iota_{X_0} \circ d\omega)(X_1, \ldots, X_k) = (\mathcal{L}_{X_0}\omega)(X_1, \ldots, X_k) - (d \circ \iota_{X_1}\omega)(X_0, \ldots, X_k) \\ &= \mathcal{L}_{X_0}(\omega(X_1, \ldots, X_k)) - \sum_{i=1}^{k}\omega(X_1, \ldots, X_{i-1}, [X_0, X_i], X_{i+1}, \ldots, X_k) \\ &- (d \circ \iota_{X_0}\omega)(X_1, \ldots, X_k) \\ &= X_0(\omega(X_1, \ldots, X_k)) - \sum_{i=1}^{k}\omega(X_1, \ldots, X_{i-1}, [X_0, X_i], X_{i+1}, \ldots, X_k) \\ &- \sum_{i=1}^{k}(-1)^{i-1}X_i(\iota_{X_0}\omega(X_1, \ldots, \widehat{X_i}, \ldots, X_k)) \\ &- \sum_{1 \le i < j \le k}(-1)^{i+j}(\iota_{X_0}\omega([X_i, X_j], X_1, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k)) \\ &= \sum_{i=0}^{k}(-1)^i X_i\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k)) \\ &- \sum_{i=1}^{k}(-1)^{i-1}\omega([X_0, X_i], X_1, \ldots, X_{i-1}, \widehat{X_i}, X_{i+1}, \ldots, X_k) \\ &+ \sum_{1 \le i < j \le k}(-1)^{i+j}\omega([X_i, X_j], X_0, X_1, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k)) \\ &= \sum_{i=0}^{k}(-1)^i X_i\omega(X_0, \ldots, \widehat{X_i}, \ldots, Y_k) \\ &+ \sum_{0 \le i < j \le k}(-1)^{i+j}\omega([X_i, X_j], X_0, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k). \end{aligned}$$

□

Example 2.2.3.2. *Let* $\omega = dz - xdy$ *on* $\mathbb{R}^3$ *Then* $d\omega(\partial x, \partial y) = -1$ *and* $d\omega(\partial x, \partial z) = 0$.

Proof. Since $d\omega = -dx \wedge dy$, by definition of a wedge product we have

$$d\omega = -(dx \otimes dy - dy \otimes dx).$$

Hence, $d\omega(\partial x, \partial y) = -1$ and $d\omega(\partial x, \partial z) = 0$. Since $[\partial x, \partial y] = [\partial x, \partial z] = 0$, by Theorem 2.2.3.3, it is also easy to see that $d\omega(\partial x, \partial y) = -1$ and $d\omega(\partial x, \partial z) = 0$. □

2.2.4 Differential form version of Frobenius's Theorem

Usually, working on differentiable forms is more preferable than working on vector fields, since differentiable forms has pullbacks and exterior algebras, etc. Also, for example, recalling Exercise 2.1.8.6, to find solution curves of

$$\frac{dy}{dx} = f(x, y),$$

we just look at the kernel of $dy - f(x, y)dx$. Now we shall give the dual version of Frobenius theorem. First of all we need a definition:

Definition 2.2.4.1 (Annihilation Ideal). *Let* E *be a* k*-plane field over* M.

$$\mathbf{Ann}^k(E) = \{\omega \in \Omega^k(M) \mid \omega(X_1, \ldots, X_k) = 0 \textit{ for all } X_i \in \Gamma(E)\}.$$

Also,

$$\mathbf{Ann}(E) = \bigoplus_{k \geq 0} \mathbf{Ann}^k(E).$$

Clearly, $\mathbf{Ann}(E)$ is an graded ideal in $\Omega^*(M)$. Moreover, since $\Omega^*(M)$ is generated by $\Omega^0(M)$ and $\Omega^1(M)$, obviously, $\mathbf{Ann}(E)$ is generated by $\mathbf{Ann}^1(E) = \mathbf{Ann}(E) \bigcap \Omega^1(M)$. Also, it is worth remarking the following: Let

$$\mathfrak{T} = \{X \in \mathfrak{X}(M) \mid \omega(X) = 0 \text{ for all } \omega \in \mathbf{Ann}^1(E)\}.$$

Clearly, $\Gamma(E) \subseteq \mathfrak{T}$. Now, suppose that $X \in \mathfrak{T}$. We know that by the local triviality of a vector bundle, the duals of $m - k$ vector fields in the codimensional space of E generate $\mathbf{Ann}^1(E)$ locally. So, X is locally in $\Gamma(E)$. Using partition of unity, we conclude that X is actually in $\Gamma(E)$. Hence,

$$\Gamma(E) = \{X \in \mathfrak{X}(M) \mid \omega(X) = 0 \text{ for all } \omega \in \mathbf{Ann}^1(E)\}.$$

Definition 2.2.4.2 (Differential ideal). *We say* $\mathbf{Ann}(E)$ *is d-stable or a differential ideal if*

$$d(\mathbf{Ann}(E)) \subseteq \mathbf{Ann}(E).$$

Since $\mathbf{Ann}(E)$ is generated by $\mathbf{Ann}^1(E) = \mathbf{Ann}(E) \bigcap \Omega^1(M)$, in order to show $\mathbf{Ann}(E)$ is a differential ideal, it suffices to show that

$$d(\mathbf{Ann}^1(E)) \subseteq \mathbf{Ann}^2(E).$$

The dual version of Frobenius theorem can be stated by the following way:

Theorem 2.2.4.1 (Frobenius's Integrability Theorem). *A k-plane field E on M^m is integrable if and only if $\mathbf{Ann}(E)$ is a differential ideal.*

The main point is that locally $m-k$ vector fields in the codimensional space of E generate a differential ideal $\mathbf{Ann}(M)$.

Proof. By the Frobenius theorem on vector fields, it suffices to show that $\Gamma(E)$ is a Lie subalgebra of $\mathfrak{X}(M)$ if and only if $\mathbf{Ann}(E)$ is a differential ideal. Suppose that $\Gamma(E)$is a Lie subalgebra of $\mathfrak{X}(M)$ and $\omega \in \mathbf{Ann}^1(E)$. We will show that $d\omega \in \mathbf{Ann}^2(E)$. Let $X, Y \in \Gamma(E)$. By Theorem 2.2.3.3,

$$(d\omega)(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]).$$

Since $[X,Y] \in \Gamma(E)$, we have $(d\omega)(X,Y) = 0$. Hence, $d\omega \in \mathbf{Ann}^2(E)$. Conversely, let $X, Y \in \Gamma(E)$. Since $\mathbf{Ann}(E)$ is a differential ideal, we have $(d\omega)(X,Y) = 0$. So, by $(d\omega)(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y])$ we have

$$\omega([X,Y]) = 0 \text{ for all } \omega \in \mathbf{Ann}^1(E).$$

Since $\Gamma(E) = \{X \in \mathfrak{X}(M) \mid \omega(X) = 0 \text{ for all } \omega \in \mathbf{Ann}^1(E)\}$ by the previous discussion, we conclude that $[X,Y] \in \Gamma(E)$. Hence it is a Lie subalgebra. □

Example 2.2.4.1. *Recalling Exercise 2.1.8.6, $E = \mathbf{span}\{\partial x, \partial y + x\partial z\}$ is not integrable, since*

$$d\omega(\partial x, \partial y + x\partial z) = -\omega([\partial x, \partial y + x\partial z]) = -(dz - xdy)(\partial z) = -1.$$

That is, $d\omega \notin \mathbf{Ann}(E)$.

Suppose that $S^s \xrightarrow{f} M^m$ be a submanifold of S^s. Since f_* is one-to-one, we have the cotangent bundle N_S of S in M:

$$N_S \overset{def}{=} \frac{TM|_S}{TS}.$$

So, we also have the dual N_S^* of N_S, which is called the conormal bundle of S. It is easy to see that

$$\Gamma(N_S^*) = \{\sigma \in \Omega^1(M) \mid f^*(\sigma) = 0\}.$$

Let $v \in T_pS$. If $\sigma \in \Gamma(N_S^*)$, $0 = f^*(\sigma)(v) = \sigma(f_*(v))$ for all $v \in T_pS$. So,

$$\Gamma(N_S^*) \subseteq \mathbf{Ann}^1(f_*(TS)).$$

Also, if $\sigma \in \mathbf{Ann}^1(f_*(TS))$, then $0 = \sigma(f_*(v)) = f^*(\sigma)(v)$ for all $v \in T_pS$. That is, $f^*(\sigma) = 0$. Hence,

$$\Gamma(N_S^*) = \mathbf{Ann}^1(f_*(TS)).$$

From this, we deduce that a submanifold S has a unique ideal, which is generated by σ_i such that $f^*(\sigma_i) = 0$.

Example 2.2.4.2. *Let $M^m \xrightarrow{f} N^n$ be a smooth map, e.g., N might be a Lie group, e.g. $M^2 \to$ $\mathbf{SO}(3)$. Define*

$$\mathsf{graph}(f) \overset{def}{=} \{(x, y) \in M \times N \mid f(x) = y\}.$$

First, note that $\mathsf{graph}(f)$ is a submanifold of $M \times N$: Let $F(m) = (m, f(m))$. Clearly, it is smooth and one-to-one. Moreover, clearly, F^* is one-to-one. So, we have a submanifold

$$M \xrightarrow{F} \mathsf{graph}(f) \subseteq M \times N.$$

So, the previous discussion says that $\mathsf{graph}(f)$ has a unique ideal $\mathcal{I}$ of $\Omega^*(M \times N)$. Let π_M and π_N be projections:

$$\begin{array}{ccc} M \times N & \xrightarrow{\pi_N} & N \\ {\scriptstyle \pi_M}\downarrow & \nearrow_{f} & \\ M & & \end{array}$$

Let $\omega_1, \ldots, \omega_d \in \Omega^1(N)$ be a basis. By f, we have $f^*\omega_i \in \Omega^1(M)$ for $i = 1, \ldots, d$. Define 1-forms $\alpha_i \in \Omega^1(M \times N)$ for $i = 1, \ldots, d$ by

$$\alpha_i = \pi_N^*\omega_i - \pi_M^*(f^*\omega_i).$$

Let $\mathcal{B}$ be the ideal generated by α_i. Then $\mathcal{B} = \mathcal{I}$.

Proof. By the previous remarks,

$$\mathcal{I} = \{\omega \in \Omega^1(M \times N) \mid F^*(\omega) = 0\}.$$

Since using $\pi_N \circ F = f$ and $\pi_M \circ F = id$,

$$\begin{aligned} F^*(\alpha_i) &= F^*(\pi_N^*\omega_i - \pi_M^*(f^*\omega_i)) = F^* \circ \pi_N^*(\omega_i) - F^* \circ \pi_M^*(f^*\omega_i) \\ &= (\pi_N \circ F)^*(\omega_i) - (\pi_M \circ F)^*(f^*\omega_i) = f^*(\omega_i) - (id)^*(f^*\omega_i) = 0, \end{aligned}$$

we have $\alpha_i \in \mathcal{I}$. Suppose $\omega \in \mathcal{I}$. So, $F^*(\omega) = 0$ Since $T(M \times N) = TM \times TN$, it is easy to see that $\Omega^1(M \times N) = \Omega^1(M) \times \Omega^1(N)$. So, without loss of generality, we can think $\omega = (\omega_M, \omega_N)$ where $\omega_M \in \Omega(M)$ and $\omega_N \in \Omega(N)$. So, for all $v \in T_m(M)$, we have

$$0 = (F^*(\omega))(v) = \omega(F_*(v)).$$

Hence, by identifying $\omega = (\omega_M, \omega_N)$, we have

$$\begin{aligned} (F^*(\omega))(v) &= \omega(F_*(v)) = \omega(id_*, f_*)(v) = \omega(v, f_*v) \\ &= (\omega_M, \omega_N)(v, f_*v) = (\omega_M v, \omega_N f_* v) = (0, 0). \end{aligned}$$

So, that $\omega_M v = 0$ for all $v \in T_m(M)$ implies that $\omega_M = 0$. Hence, we can identifying ω as ω_N. That is,

$$\pi_N^*\omega_N - \pi_M^*(f^*\omega_N) = \omega \in \mathcal{B}.$$

In particular, by Frobenius theorem, $\mathcal{B} = \mathcal{I}$ is a differential ideal, which annihilates $\mathsf{graph}(f)$. □

We give a converse of Exercise 2.2.4.2.

Example 2.2.4.3. *Suppose that given $\eta_1, \ldots, \eta_d \in \Omega^1(M)$ and a basis $\omega_1, \ldots, \omega_d \in \Omega^1(N)$, the 1-forms*

$$\alpha_i = \pi_N^* \omega_i - \pi_M^* \eta_i$$

generate a differential ideal $\mathcal{B}$ in $\Omega^1(M \times N)$. Then for all $(m, n) \in M \times N$, there exists a neighborhood $U \subseteq M$ of m and a smooth map $f : U \to N$ such that $f(m) = n$ and $f^*\omega_i = \eta_i|_U$. Furthermore, f is unique.

Proof. Since $\mathcal{B}$ is a differential ideal, by the Frobenius theorem, given (m_0, n_0) there exists a maximal submanifold, i.e., leaf, S of (m_0, n_0) in $M \times N$.

$$\begin{array}{ccc} M \times N & \xrightarrow{\pi_N} & N \\ {\scriptstyle \pi_M}\big\downarrow & & \\ M & & \end{array}$$

First, note that $\pi_M|_{TS}$ is one to one: Suppose that $v \in T_pS \subseteq T_p(M \times N)$ and $(\pi_M)_*(v) = 0$. So, we have for $i = 1, \ldots, d$

$$0 = (\pi_N^* \omega_i - \pi_M^* \eta_i)v = \omega_i(\pi_N)_*(v) - \eta_i(\pi_M)_*(v) = \omega_i(\pi_N)_*(v).$$

Since $\omega_1, \ldots, \omega_d \in \Omega^1(N)$ is a basis, we have $(\pi_N)_*(v) = 0$. Hence, the fact $(\pi_M)_*(v) = 0$ and $(\pi_N)_*(v) = 0$ implies that $v = 0$. Hence, $(\pi_M)_*$ is one-to-one. So, by the inverse function theorem, $\pi_M|_S$ is a local diffeomorphism. Hence, there exists an open neighborhood set $V \subseteq S$ of (m_0, n_0) such that

$$\pi_M : V \xrightarrow{\cong} U \subseteq M.$$

Define

$$f = \pi_N \circ (\pi_M^{-1}|_U) : U \to N.$$

Since π_M and π_N are projections, clearly, $f(m_0) = n_0$ and f is smooth. Moreover, $\mathsf{graph}(f) \overset{def}{=} \{(x, y) \in M \times N \mid f(x) = y\}$ is a submanifold of S. Furthermore, for all $v \in T_uM$ for any $u \in U$,

$$\begin{aligned} 0 = \alpha_i((\pi_M^{-1}|_U)_* v) &= ((\pi_M^{-1}|_U)^* \alpha_i)(v) = ((\pi_M^{-1}|_U)^* \circ (\pi_N^* \omega_i - \pi_M^* \eta_i))(v) \\ &= ((\pi_N^* \circ \pi_M^{-1}|_U)^* \omega_i - (id)^* \eta_i)(v) = (f^*(\omega_i) - \eta_i)(v). \end{aligned}$$

Hence, $f^*\omega_i = \eta_i|_U$. Note that f is unique by the following reasons: If $\widetilde{f}$ is another function satisfying all the required properties, by Exercise 2.2.4.2, $\mathsf{graph}(\widetilde{f})$ is a submanifold of $M \times N$ with a the same differential ideal as $\mathsf{graph}(f)$. Hence, the uniqueness of a maximal integral submanifold, leaf, gives the uniqueness of $f = \widetilde{f}$. $\square$

Example 2.2.4.4. *Let $x_1, \ldots, x_m$ be a local coordinate chart of M and $y_1, \ldots, y_n$ a local coordinate chart of N. Suppose that we are given a P.D.E. system: For $i = 1, \ldots, n$ and $j = 1, \ldots, m$,*

$$\frac{\partial y_i}{\partial x_j} = \varphi_{ij}(x_1, \ldots, x_m, y_1, \ldots, y_n).$$

So, we have for $i = 1, \ldots, n$ and $j = 1, \ldots, m$,

$$\omega_{ij} = dy_i - \varphi_{ij} dx_j.$$

If the ideal $\mathcal{I}_k$, generated by ω_{kj} for $j = 1, \ldots, m$, is a differential ideal, then by Example 2.2.4.3, we have

$$F_k(x_i, \ldots, x_m) = y_k.$$

So, if the ideal $\mathcal{I}$, generated by ω_{ij} for $i = 1, \ldots, n$ and $j = 1, \ldots, m$, is a differential ideal, then by Example 2.2.4.3, this P.D.E. system is solvable. That is, we have

$$F_i(x_i, \ldots, x_m) = y_i \text{ for } i = 1, \ldots, n.$$

In particular, if $M = \mathbb{R}$, then for $i = 1, \ldots, n$,

$$\frac{dy_i}{dt} = \varphi_i(t, y_1, \ldots, y_n).$$

So, we have for $i = 1, \ldots, n$,

$$\omega_i = dy_i - \varphi_i dt.$$

Since

$$d\omega_i = -\sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j} dt \wedge dy_j = -\sum_{j=1}^{n} (\frac{\partial \varphi_i}{\partial y_j} dt \wedge (dy_j - \varphi_j dt)) = -\sum_{j=1}^{n} (\frac{\partial \varphi_i}{\partial y_j} dt \wedge \omega_j),$$

we deduce that the ideal $\mathcal{I}$, generated by ω_i for $i = 1, \ldots, n$, is always a differential ideal. From this, we have then by Example 2.2.4.3, every first order O.D.E. system is solvable. Equivalently, every line field is integrable.

Chapter 3

General Theory of Lie Groups and Lie Algebras

3.1 Lie Algebras of Lie Groups

3.1.1 Adjoint representations

Let G be a Lie group. The Lie algebra of G is

$$\mathfrak{g} = \{X \in \mathfrak{X}(G) \mid l_g \cdot X = X\} \stackrel{def}{=} \mathfrak{L}\mathfrak{X}(G).$$

Recall that in Exercise 2.1.8.2, we showed that the evaluation map is an isomorphism:

$$\epsilon : \mathfrak{L}\mathfrak{X}(G) \to T_e(G) \text{ by } \epsilon(X) = X_e.$$

That is, X is completely determined by its value at e and each element of T_eG gives a unique left-invariant vector field. So, if $X \in \mathfrak{L}\mathfrak{X}(G)$, then

$$X_g = dl_g(X_e).$$

Let $i : G \to G$ be an inversion map by $i(g) = g^{-1}$. So, $i_* : \mathfrak{X}(G) \to \mathfrak{X}(G)$ is an isomorphism. Since $i \circ l_g(h) = r_{g^{-1}} \circ i(h)$, for $X \in \mathfrak{L}\mathfrak{X}(G)$ we have

$$r_{g^{-1}} \cdot (i_*(X)) = r_{g^{-1}} \cdot (i \cdot X) = (r_{g^{-1}} \circ i) \cdot X = (i \circ l_g) \cdot X = i \cdot (l_g \cdot X) = i \cdot X = i_*(X).$$

Hence, letting $\mathfrak{R}\mathfrak{X}(G)$ be the set of right-invariant vector fields, we have a well-defined map

$$i_* : \mathfrak{L}\mathfrak{X}(G) \to \mathfrak{R}\mathfrak{X}(G).$$

Since $i_* : \mathfrak{X}(G) \to \mathfrak{X}(G)$ is already an isomorphism, we conclude that $\mathfrak{L}\mathfrak{X}(G) \cong \mathfrak{R}\mathfrak{X}(G)$. Note that in Exercise 2.1.8.3, we already showed that they are isomorphic as Lie algebras. Suppose that G is an n-dimensional Lie group. In Theorem 2.1.8.3, we showed that $\mathfrak{X}(G)$ consists of n linearly independent left-invariant vector fields X_i for $i = 1, \ldots, n$. Let

$$\omega_i = (X_i)^* \text{ for } i = 1, \ldots, n.$$

Since $\{X_{j,g}\}_{j=1,\ldots,n}$ is a basis of T_gG and

$$(l_g^*\omega_i)_e(X_{j,e}) = \omega_{i,g} \circ (l_g)_*(X_{j,e}) = \omega_{i,g}(X_{j,g}) \Longrightarrow (l_g^*\omega_i)_e = \omega_{i,e},$$

we conclude that $(l_g^*\omega_i)_h = \omega_{i,h}$ for all $h \in G$, so that ω_i is a left-invariant 1-form for $i = 1, \ldots, n$, i.e., $l_g^*\omega_i = \omega_i$. Moreover, $(l_g)^*$ is a graded algebra homomorphism, we have

$$(l_g)^*(\omega_1 \wedge \cdots \wedge \omega_n) = (l_g)^*\omega_i \wedge \cdots \wedge (l_g)^*\omega_n = \omega_1 \wedge \cdots \wedge \omega_n.$$

Hence, we prove every Lie group has a left-invariant volume form. In particular, a Lie group is always orientable. Since a similar proof would give the existence of a right-invariant volume form, we have

Theorem 3.1.1.1. *If G is a Lie group, the Radon measure comes from the volume form. Moreover, there exists a left-invariant volume form. Also, there exists a right-invariant volume form. Note that in general, theory are not equal. Furthermore, if G is compact, there exists a unique left-invariant volume form with total volume 1.*

In general, we have

Theorem 3.1.1.2 (Haar). *Every locally compact topological group admits a left-invariant Radon measure*

Proof. See Ann. of Math. *34*, 1933, p.147. ☐

Example 3.1.1.1. *Let $\omega \in \Omega^n(M^n)$ be a volume form. So, $\omega_p \neq 0$ for all $p \in M$. For $A \subseteq M$, the measure $\mu(A)$ is given by*

$$\mu(A) = \int_M \chi_A \omega = \int_A \omega.$$

Note that by the previous remarks, we also have the evaluation map as an isomorphism, letting $\mathfrak{L}\Omega^k(G)$ be the set of left-invariant k-forms,

$$\epsilon : \mathfrak{L}\Omega^k(G) \to \wedge^k T_e^*G = \wedge^k \mathfrak{g}^*.$$

Since left-multiplications commute right-multiplications, right-multiplications r_g act on left-invariant vector fields: Let $X \in \mathfrak{LX}(G)$ and $h, g \in G$. We have

$$(l_h)_*((r_g)_*X) = ((l_h)_* \circ (r_g)_*)X = (l_h \circ r_g)_*X = (r_g \circ l_h)_*X = (r_g)_*((l_h)_*X) = (r_g)_*X.$$

So, we define the adjoint representation:

$$\mathbf{Ad} : G \to \mathbf{Aut}(\mathfrak{g}) \text{ so that } \mathbf{Ad}(g)(X) = (r_{g^{-1}})_*X \text{ where } X \in \mathfrak{g} = \mathfrak{LX}(G).$$

Note that if we identify $\mathfrak{g}$ with T_eG rather than $\mathfrak{LX}(G)$, we have for $X \in \mathfrak{LX}(G)$,

$$(\mathbf{Ad}(g)(X))_e = (r_{g^{-1}})_*(X_g) = (r_{g^{-1}})_*((l_g)_*X_e) = (r_{g^{-1}} \circ l_g)_*(X_e).$$

That is, letting $r_{g^{-1}} \circ l_g = \iota_g$, conjugation by g, we have

$$\mathbf{Ad}(g) = (\iota_g)_* \text{ on } T_eG = \mathfrak{g}.$$

For a reference, we give

Remark 3.1.1.1.

$$G \xrightarrow{\iota} \mathbf{Aut}(G) \text{ by } \iota(g)(h) = ghg^{-1}$$
$$G \xrightarrow{\mathbf{Ad}} \mathbf{Aut}(\mathfrak{g}) = \mathbf{Aut}(T_eG) \cong \mathbf{GL}(n,\mathbb{R}) \text{ by } \mathbf{Ad}(g)(v) = (\iota_g)_*(v)$$
$$G \xrightarrow{\mathbf{Ad}} \mathbf{Aut}(\mathfrak{g}) = \mathbf{Aut}(\mathfrak{LX}(G)) \cong \mathbf{GL}(n,\mathbb{R}) \text{ by } \mathbf{Ad}(g)(X) = (r_{g^{-1}})_*(X)$$
$$\mathfrak{LX}(g) = \mathfrak{g} \xrightarrow{\mathbf{ad}} Der(\mathfrak{g}) \cong \mathbf{M}_n(\mathbb{R}) \text{ by } \mathbf{ad}(X)Y = [X,Y].$$

Note that derivations are infinitesimal automorphisms: Since $T_{id}\mathbf{GL}(n,\mathbb{R}) \cong \mathbf{M}_n(\mathbb{R}) = \mathbf{End}(T_eG)$, *we have a commutative diagram*

$$\begin{array}{ccc} G & \xrightarrow{\mathbf{Ad}} & \mathbf{Aut}(\mathfrak{g}) = \mathbf{Aut}(T_eG) \cong \mathbf{GL}(n,\mathbb{R}) \\ {\scriptstyle\exp}\uparrow & & {\scriptstyle\exp}\uparrow \\ \mathfrak{g} \cong T_eG & \xrightarrow{(\mathbf{Ad})_*=\mathbf{ad}} & Der(\mathfrak{g}) \cong \mathbf{End}(\mathfrak{g}) \cong \mathbf{End}(T_eG) \cong T_{id}\mathbf{GL}(n,\mathbb{R}) = \mathfrak{gl}(n,\mathbb{R}). \end{array}$$

Exercise 3.1.1.1. *Show that* $\mathbf{Ad}(g)$ *is a Lie algebra automorphism.*

Proof. It suffices to show that $\mathbf{Ad}(g)[X_1,X_2] = [\mathbf{Ad}(g)X_1, \mathbf{Ad}(g)X_2] = [(r_{g^{-1}})_*(X_1), (r_{g^{-1}})_*(X_2)]$. Note that $(r_{g^{-1}})_*(X_i)$ and X_i are $(r_{g^{-1}})$-related for $i = 1,2$. So, by Exercise 2.1.8.1, we conclude that $[X_1,X_2]$ and $[(r_{g^{-1}})_*(X_1), (r_{g^{-1}})_*(X_2)]$ are $(r_{g^{-1}})$-related. That is,

$$\mathbf{Ad}(g)[X_1,X_2] = (r_{g^{-1}})_*[X_1,X_2] = [(r_{g^{-1}})_*(X_1), (r_{g^{-1}})_*(X_2)].$$

□

Note Ado states that every Lie algebra has a faithful representation by a matrix group, some $\mathfrak{gl}(m,\mathbb{R})$. However, from the example of the Heisenberg group, we know that not every Lie group has a faithful representation by a matrix group. Since $(\iota_g)_* = \mathbf{Ad}(g)$, it is easy to see that the center of G is the kernel of $\mathbf{Ad}$, i.e.,

$$\mathcal{Z}(G) = \mathbf{ker}(\mathbf{Ad}).$$

Hence, $\mathbf{Ad}$ is not a faithful representation if $\mathcal{Z}(G) \neq id$. Also, clearly,

$$\mathbf{Ad}(G) \cong G/\mathcal{Z}(G).$$

From this we have

Definition 3.1.1.1. *We say* G *is adjoint if* $\mathbf{Ad} : G \to \mathbf{Aut}(\mathfrak{g})$ *is faithful, equivalently,* $\mathcal{Z}(G) = id$.

Remark 3.1.1.2. g *is semisimple if and only if* $\mathcal{Z}(G)$ *is a finite group.*

Example 3.1.1.2. *Since the center of* $\mathbf{SL}(2,\mathbb{R}) = \{\pm I_2\}$, *we have*

$$\mathbf{Ad}(\mathbf{SL}(2,\mathbb{R})) \cong \mathbf{SL}(2,\mathbb{R})/\{\pm I_2\} = \mathbf{PSL}(2,\mathbb{R}).$$

Also,

$$\mathbf{Ad}(\mathbf{SU}(2)) \cong \mathbf{SU}(2)/\{\pm I_2\} = \mathbf{PU}(2) \cong \mathbf{SO}(3).$$

Example 3.1.1.3. *Let $g \in G$ where G is a n-dimensional Lie group. Since $\mathbf{Ad}(g) \in \mathbf{Aut}(\mathfrak{g})$, we have*

$$(r_{g^{-1}})_* = \mathbf{Ad}(g) : T_eG \to T_eG.$$

Let $v_1, \ldots, v_n$ be a basis of T_eG. So, we have

$$\mathbf{Ad}(g)(v_i) = \sum_{j=1}^{n} a_{ij} v_j.$$

Since we know that $\mathfrak{L}\Omega^n(G) \cong \wedge^n T_e^* G$, by the fact that G acts on $\mathfrak{L}\Omega^n(G)$ by right multiplications, so that $(r_{g^{-1}})^*$ is a graded Lie algebra homomorphism ,and using the dual basis v_i^* for $i = 1, \ldots, n$

$$\begin{aligned}(r_{g^{-1}})^*(v_1^* \wedge \cdots \wedge v_n^*) &= (\mathbf{Ad}(g))^*(v_1^* \wedge \cdots \wedge v_n^*) \\ &= (\sum_{j=1}^{n} a_{1j} v_j^*) \wedge \cdots \wedge (\sum_{j=1}^{n} a_{nj} v_j^*) \\ &= \det((\mathbf{Ad}(g))^*) v_1^* \wedge \cdots \wedge v_n^* \\ &= \det(\mathbf{Ad}(g)) v_1^* \wedge \cdots \wedge v_n^*.\end{aligned}$$

So, defining

$$G \xrightarrow{\mathbf{Ad}} \mathbf{Aut}(\mathfrak{g}) \xrightarrow{\det} \mathbb{R}^\times, \text{ i.e., a modular function } \det \circ \mathbf{Ad} \text{ on } G,$$

we conclude that $\det \circ \mathbf{Ad} \equiv 1$, i.e., G is unimodular, if and only if a left-invariant volume form is right-invariant. That is, G is unimodular if and only if there exists a bi-invariant volume form on G.

Example 3.1.1.4. *Let*

$$G = \mathbf{Aff}_+(\mathbb{R}) = \{\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \mid x, y \in \mathbb{R} \text{ and } y > 0\}.$$

G is a connected open multiplicative subgroup of $\mathbf{GL}(2, \mathbb{R})$. So, it is a Lie group. Identify (x, y) with $\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix}$. So, the identity of G becomes $(0, 1)$. Note that

$$l_{(a,b)}(x, y) = (a + bx, by).$$

Since left-invariant vector fields on G is completely determined by elements of $T_{(0,1)}G$ by the equation

$$\begin{aligned}dl_{(a,b)}(1, 0) = dl_{(a,b)}(\frac{\partial}{\partial x}) = (\frac{\partial}{\partial x})(a + bx, by) = (b, 0) = b\frac{\partial}{\partial x} \\ dl_{(a,b)}(0, 1) = dl_{(a,b)}(\frac{\partial}{\partial y}) = (\frac{\partial}{\partial y})(a + bx, by) = (0, b) = b\frac{\partial}{\partial y},\end{aligned}$$

we conclude that

$$X_{(x,y)} = y\frac{\partial}{\partial x} \text{ and } Y_{(x,y)} = y\frac{\partial}{\partial y}$$

are left-invariant vector fields on G. Note that $[y\partial y, y\partial x] = y\partial x$ says that $\mathfrak{g}$ is a 2-dimensional non abelian Lie algebra. Now look at $\mathfrak{L}\Omega^1(G) = \mathfrak{g}^*$. Clearly, the basis of $\mathfrak{L}\Omega^1(G)$ is $\{y^{-1}dx, y^{-1}dy\}$. Notice that since $y : G \to \mathbb{R}_+$, we have

$$\log y : G \to \mathbb{R}.$$

In particular, $d(\log y) = y^{-1}dy$. A left-invariant volume form is

$$y^{-1}dx \wedge y^{-1}dy = y^{-2}dx \wedge dy, \text{ which is everywhere nonzero.}$$

Now, we shall do the exactly the same things for right-invariant vector fields. So, we have right-invariant vector fields

$$\frac{\partial}{\partial x} \text{ and } x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}.$$

The basis of $\mathfrak{R}\Omega^1(G)$ is $\{dx, y^{-1}dy\}$ and a right-invariant volume form is

$$dx \wedge (y^{-1}dy) = y^{-1}dx \wedge dy.$$

From this, we observe that comparing the flows, the dynamics and behaviors of left-invariant vector fields and right-invariant vector fields are very different even though the Lie algebras are isomorphic.

Exercise 3.1.1.2. *Compute adjoint representations.*

Proof. Note that $\mathbf{Ad}(a,b) = (\iota_{(a,b)})_*$. Since

$$\begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{b} & -\frac{a}{b} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} y & bx - ay + a \\ 0 & 1 \end{pmatrix},$$

we have

$$d\iota_{(a,b)}(1,0) = d\iota_{(a,b)}(\frac{\partial}{\partial x}) = (\frac{\partial}{\partial x})(bx - ay + a, y) = (b, 0)$$
$$dl_{(a,b)}(0,1) = dl_{(a,b)}(\frac{\partial}{\partial y}) = (\frac{\partial}{\partial y})(bx - ay + a, y) = (-a, 1).$$

Hence,

$$\mathbf{Ad}(a,b) = (\iota_{(a,b)})_* = \begin{pmatrix} b & -a \\ 0 & 1 \end{pmatrix}.$$

□

3.1.2 Maurer-Cartan Forms

Let $\mathfrak{g}$ be the Lie algebra of left-invariant vector fields on a Lie group G and $\mathfrak{g}^*$ be the dual vector space of $\mathfrak{g}$. By letting $T : \mathbf{Aut}(\mathfrak{g}) \to \mathbf{Aut}(\mathfrak{g}^*)$ by $T(f) = f^T$ the transpose of f, we have a coadjoint map $\mathbf{Ad}^*$ from the following commutative diagram:

$$\begin{array}{ccc} G & \xrightarrow{\mathbf{Ad}^*} & \mathbf{Aut}(\mathfrak{g}^*) \\ {\scriptstyle \mathbf{Ad}}\downarrow & T \nearrow & \\ \mathbf{Aut}(\mathfrak{g}) & & \end{array}$$

Definition 3.1.2.1 (Symplectic Structure). *Let M^{2n} be a 2n-dimensional smooth manifold and $\omega \in T^*M \otimes T^*M$ be a nondegenerate bilinear form:*

$$\omega : TM \times TM \to \mathbb{R}.$$

Note that ω is called a Riemannian metric if it is nondegenerate and ***symmetric****. If ω is nondegenerate and* ***skew-symmetric*** *bilinear form satisfying*

$$d\omega = 0,$$

we call ω a ***symplectic structure*** *on M.*

Definition 3.1.2.2 (Almost Symplectic Structures). *Let $\omega = \sum_{i,j}^{2n} a_{ij} dx_i \wedge dx_j$ be a 2-form on smooth manifold M^{2n}. We say ω is an almost symplectic structure if*

$$\det((a_{ij})_{2n\times 2n)}) \neq 0.$$

That is, ω is nondegenerate.

Note that nondegeneracy of $\omega \in \Omega^2(M)$ implies that $0 \neq \underbrace{\omega \wedge \cdots \wedge \omega}_{n} \in \Omega^{2n}(M)$ is the volume form and the Darboux's theorem as some disguise of the Frobenius theorem tells us that $d\omega = 0$ implies that ω is locally nothing but the standard symplectic structure $dx_1 \wedge dy_1 + \cdots + dx_n \wedge dy_n$ of Euclidean space $\mathbb{R}^{2n}$.

Exercise 3.1.2.1. *Using a Lie algebra structure on $\mathfrak{g}$, construct a symplectic structure on an orbit*

$$G \cdot \psi = \{\mathbf{Ad}^*(g)(\psi) \mid g \in G\} \text{ where } \psi \in \mathfrak{g}^*.$$

Sketch of Proof. Since $\psi \in \mathfrak{g}^*$, we have $d\psi \in \wedge^2 \mathfrak{g}^*$ and $d \circ d\psi = 0$. Let

$$\mathfrak{h} = \{X \in \mathfrak{g} \mid d\psi(X, Y) = 0 \text{ for all } Y \in \mathfrak{g}\}.$$

Using Theorem 2.2.3.4, it is easy to see that $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. So, there exists a closed Lie subgroup H of G corresponding to $\mathfrak{h}$, which we shall prove later. Now let

$$\pi : G \to G/H. \text{ Note that } G/H \text{ is a smooth manifold.}$$

Let $\pi^*(\widehat{d\psi}) = d\psi$. Since $d\psi$ is skew-symmetric, $\widehat{d\psi}$ is skew-symmetric. Since

$$0 = d \circ d\psi = d\pi^*(\widehat{d\psi}) = \pi^*(d \circ \widehat{d\psi}),$$

and π^* is a monomorphism, we have $d \circ \widehat{d\psi} = 0$. By the construction of G/H, it is easy to see that $\widehat{d\psi}$ is nondegenerate. That is, $\widehat{d\psi}$ is a symplectic structure of G/H. Now, letting $\pi^*(\widehat{\psi}) = \psi$, define

$$f_{mH} : G \to G/H \text{ by } f_{mH}(g) = gmH \text{ and}$$
$$F : G/H \to \mathfrak{g} \text{ by } F(gmH) = \mathbf{Ad}^*(g)(f^*_{mH}(\widehat{\psi})).$$

□

Remark 3.1.2.1. *From this, we deduce that the G-orbits in $\mathfrak{g}^*$ have invariant symplectic structures and $\dim G{\cdot}\psi$ is even. A Poisson manifold M is a smooth manifold in which $C^\infty(M)$ is a Lie algebra. Note that $\mathfrak{g}^*$ is an example of Poisson manifolds, since $C^\infty(\mathfrak{g}^*) = \mathfrak{g}$.*

Notation 3.1.2.1. *Let*

$$\Omega^k(G, \mathfrak{g}) \overset{def}{=} \mathfrak{L}\Omega^k(G) \otimes_{\mathbb{R}} \mathfrak{g}.$$

In other words, $\Omega^k(G, \mathfrak{g}) = \wedge^k \mathfrak{g}^* \otimes_{\mathbb{R}} \mathfrak{g}$. So, if $\xi \in \Omega^k(G, \mathfrak{g})$, then

$$\xi = \sum \omega_i \otimes_{\mathbb{R}} X_i \text{ where } \omega_i \in \mathfrak{L}\Omega^k(G) \text{ and } X_i \in \mathfrak{L}\mathfrak{X}(G).$$

Since

$$\xi = \sum \omega_i \otimes_{\mathbb{R}} X_i \in \Gamma(\mathbf{Hom}(\underbrace{TG \times \cdots \times TG}_{k}, \mathfrak{g})),$$

ξ is nothing but a $\mathfrak{g}$-valued left invariant k-form.

Now we define a derivation d of degree 1 on $\Omega^k(M, \mathfrak{g})$ by the following way:

$$d : \Omega^k(M, \mathfrak{g}) \to \Omega^{k+1}(M, \mathfrak{g}) \text{ by } d(\alpha \otimes X) = d\alpha \otimes X.$$

It is easy to see that d is a derivation. Also, we will define a graded algebra structure on $\Omega^*(M, \mathfrak{g})$, using Lie algebra structures of $\mathfrak{g}^*$ and $\mathfrak{g}$:

$$\Omega^{k_1}(M, \mathfrak{g}) \times \Omega^{k_2}(M, \mathfrak{g}) \xrightarrow{[,]} \Omega^{k_1+k_2}(M, \mathfrak{g}) \text{ by } [\alpha \otimes X, \beta \otimes Y] = (\alpha \wedge \beta) \otimes [X, Y].$$

Note that we can easily check that it defines a graded algebra on $\Omega^*(M, \mathfrak{g})$, which is not a graded Lie algebra by the next theorem.

Theorem 3.1.2.1 (Maurer-Cartan structure equation). *Let $X_1, \ldots, X_n$ be a basis of $\mathfrak{g}$ and $\omega_1, \ldots, \omega_n$ be the corresponding dual basis of $\mathfrak{g}^*$. Note that an element of $\mathfrak{g}^*$ is called a **Maurer-Cartan** form on M. Let $\omega = \sum_{k=1}^n \omega_k \otimes X_k$ and $[X_i, X_j] = \sum_{k=1}^n c_{ij}^k X_k$. We have*

$$d\omega_k = -\sum_{i<j} c_{ij}^k \omega_i \wedge \omega_j \text{ and } d\omega + \frac{1}{2}[\omega, \omega] = 0.$$

Proof. Note that

$$d(\omega_k \otimes X_k) = d\omega_k \otimes X_k = (\sum_{i<j} e_{ij}^k \omega_i \wedge \omega_j) \otimes X_k.$$

Consider $[X_i, X_j] = \sum_{s=1}^n c_{ij}^s X_s$. Even though we need the following property in this proof, we want to note that by the skew-symmetry and the Jacobi identity, we have

$$c_{ij}^s = -c_{ji}^s \text{ and } \sum_{r=1}^n c_{ij}^r c_{rk}^s + c_{jk}^r c_{ri}^s + c_{ki}^r c_{rj}^s = 0 \text{ for all } i, j, k, s = 1, \ldots, n.$$

Note that the second equation shows that a Lie algebra structure gives an algebraic variety. Now, let us get back to the proof. We will show that $c_{ij}^k = -e_{ij}^k$. By Theorem 2.2.3.3 and recalling

$\omega_i(X_j) = \delta_{ij}$, Kronecker's delta, so constant in any cases, we have

$$-e^k_{i_1j_1} = -\sum_{i<j} e^l_{ij}\omega_i \wedge \omega_j(X_{i_1}, X_{j_1}) = -d\omega_k(X_{i_1}, X_{j_1})$$
$$= -X_{i_1}(\omega_k(X_{j_1})) + X_{j_1}(\omega_k(X_{i_1})) + \omega_k([X_{i_1}, X_{j_1}])$$
$$= \omega_k([X_{i_1}, X_{j_1}]) = \omega_k(\sum_{s=1}^n c^s_{i_1j_1} X_s) = c^k_{i_1j_1}$$

Using $c^k_{ij} = -c^k_{ji}$ and the skew-symmetry of wedge products, we have

$$[\omega, \omega] = [\sum_{i=1}^n \omega_i \otimes X_i, \sum_{j=1}^n \omega_j \otimes X_j] = \sum_{i=1}^n\sum_{j=1}^n [\omega_i \otimes X_i, \omega_j \otimes X_j] = \sum_{i=1}^n\sum_{j=1}^n \big((\omega_i \wedge \omega_j) \otimes [X_i, X_j]\big)$$
$$= \sum_{i=1}^n\sum_{j=1}^n \big((\omega_i \wedge \omega_j) \otimes \sum_{k=1}^n c^k_{ij} X_k\big) = \sum_{k=1}^n \big(\sum_{i=1}^n\sum_{j=1}^n c^k_{ij}(\omega_i \wedge \omega_j) \otimes X_k\big) = \sum_{k=1}^n 2\sum_{i<j} c^k_{ij}(\omega_i \wedge \omega_j) \otimes X_k$$
$$= -2\sum_{k=1}^n\sum_{i<j} e^k_{ij}(\omega_i \wedge \omega_j) \otimes X_k = -2\sum_{k=1}^n (d\omega_k \otimes X_k) = -2d(\sum_{k=1}^n \omega_k \otimes X_k) = -2d\omega.$$

□

Remark 3.1.2.2. *Letting $X_1, \ldots, X_n$ be a basis of $\mathfrak{g}$ and $\omega_1, \ldots, \omega_n$ be the corresponding dual basis of $\mathfrak{g}^*$, we have the Maurer-Cartan form*

$$\omega = \sum_{k=1}^n \omega_k \otimes X_k.$$

So, we have $\omega(X_j) = \sum_{k=1}^n \omega_k(X_j) \otimes X_k = X_j$. In this consideration, we have another description of the Maurer-Cartan form: Since an element of $\mathfrak{g}$ is completely determined by T_eG, we have

$$\mathbf{id} : T_eG \to \mathfrak{g} \text{ by } \mathbf{id}(v) = dl_g(v).$$

Note that $dl_g(v)$ is a vector field, i.e., $dl_g(v) \in T_gG$ for each $g \in G$. The Maurer-Cartan form is corresponding to the identity map **id**.

Note that $dl_{g^{-1}} : T_gG \to T_eG$ is an isomorphism. If $\{v_1, \ldots, v_n\}$ is a basis of T_eG, then we have n vector fields $dl_g(v_i)$ for $i = 1, \ldots, n$. Using the fact dl_g is an isomorphism for each $g \in G$, it is easy to see that those vector fields are linearly independent. Recalling Exercise 2.1.8.2, we conclude that every Lie group is parallelizable. Especially, we call dl_g a trivialization (resp. parallelization) of TG (resp. G). So in general we have there exists a 1-from $\omega \in \Omega(M, \mathfrak{g})$ such that $d\omega + \frac{1}{2}[\omega, \omega] = 0$ if and only if TM is parallelizable. Also, in the case of a Lie group G, if the Lie algebra $\mathfrak{g}$ is abelian, we have $d\omega = 0$ by the equation $d\omega + \frac{1}{2}[\omega, \omega] = 0$. So, if $d\omega \neq 0$, then $\mathfrak{g}$ can not be abelian. That is, the Maurer-Cartan equation shows that whether or not $\mathfrak{g}$ is abelian. Suppose that we have a Lie group homomorphism $\varphi : G \to H$ between two Lie groups, G, H. Recalling Exercise 2.1.8.1, we note that what Exercise 2.1.8.1 tells us is that $d\varphi$ is a Lie algebra homomorphism: Recall that we say $X \in \mathfrak{LX}(G)$ is φ-related to $Y \in \mathfrak{LX}(H)$ if

$$d\varphi(X_g) = Y_{\varphi(g)}.$$

The claim is that X is φ-related to $\Phi(X)$:

$$\begin{array}{ccc} \mathfrak{g} & = & T_eG \\ \Phi \downarrow & & \downarrow (d\varphi)_e \\ \mathfrak{h} & = & T_eH \end{array}$$

Note that $\varphi(e) = e$ implies that $(\Phi(X))_e = d\varphi(X_e)$. Also, since φ is a homomorphism, we have

$$l_{\varphi(g)} \circ \varphi(a) = \varphi(g) \cdot \varphi(a) = \varphi(ga) = \varphi \circ l_g(a).$$

So, we also have $dl_{\varphi(g)} \circ d\varphi = d\varphi \circ dl_g$. Hence,

$$\Phi(X)_{\varphi(g)} = dl_{\varphi(g)}(\Phi(X)_e) = dl_{\varphi(g)}(d\varphi(X_e)) = d\varphi(dl_g(X_e)) = d\varphi(X_g).$$

That is, X is φ-related to $\Phi(X)$. Therefore, Exercise 2.1.8.1 tells us that $(d\varphi)_e$ is a Lie algebra homomorphism and so is Φ. However, what is remarkable is that the converse is also true. That is, every Lie algebra homomorphism gives rise to a Lie group homomorphism. See the Warner's book p. 94.

3.1.3 Almost Complex structures

Let E be a k-plane field of a tangent bundle TM of a smooth manifold M. Define a projection

$$\pi : \mathfrak{X}(M) \overset{def}{=} \Gamma(TM) \to \Gamma(TM/E).$$

Of course, TM/E is called a normal bundle of a foliation $\mathcal{F}$ if E is integrable. Using the projection, we now define

$$\Omega_E : \Gamma(E) \times \Gamma(E) \to \Gamma(TM/E) \text{ by } \Omega_E(X, Y) = \pi([X, Y]).$$

It is easy to see that the Frobenius theorem can be reformulated as follows:

$$\Omega_E \equiv 0 \text{ if and only if } E \text{ is integrable.}$$

Note that Ω_E is $C^\infty(M)$-bilinear:

$$\Omega_E(fX, gY) = \pi([fX, gY]) = \pi(f(Xg)Y + fg[X, Y] - g(Yf)X) = fg\pi([X, Y]) = fg\Omega_E(X, Y).$$

Definition 3.1.3.1 (Complex Structures on a vector space). *Let V be a real vector space. A complex structure J on V is an $\mathbb{R}$-linear isomorphism such that $J^2 = -id$.*

Note that if V has a complex structure, it is obvious that V can not have eigenvalues of J. That is, V only has eigenpaires of J. Hence , V is necessarily even dimensional.

Definition 3.1.3.2 (Complex Structure on a manifold). *A complex structure (holomorphic structure) on M is an atlas of holomorphic patches. To avoid confusion, we will call this a holomorphic structure on M.*

Exercise 3.1.3.1. *Show that there is one-to-one correspondence between the set of complex structures of $\mathbb{R}^2$ and $\mathbf{SL}(2, \mathbb{R})/\mathbf{SO}(2)$.*

Proof. Note that by Example 1.2.2.1, there is the standard identification of $\mathbf{GL}(1,\mathbb{C}) \hookrightarrow \mathbf{GL}(2,\mathbb{R})$. Moreover, we have

$$\mathbf{GL}(1,\mathbb{C}) \cong \{A \in \mathbf{GL}(2,\mathbb{R}) \mid A^{-1} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\}.$$

So, letting $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, which is called the standard complex structure on $\mathbb{R}^2$, we have a one-to-one correspondence

$$\mathbf{GL}(2,\mathbb{R})/\mathbf{GL}(1,\mathbb{C}) \leftrightarrow \text{ the set of complex strucures on } \mathbb{R}^2 \text{ by } [A] \mapsto A^{-1}JA.$$

Note that $(A^{-1}JA)^2 = -id$. It is easy to see that $\mathbf{GL}(2,\mathbb{R})/\mathbf{GL}(1,\mathbb{C}) \cong \mathbf{SL}(2,\mathbb{R})/\mathbf{SO}(2)$, since $id_2, -id_2 \in \mathbf{GL}(1,\mathbb{C})$. □

Definition 3.1.3.3 (Almost Complex Structures). *An almost complex structure on M is a complex structure on TM. That is, an almost complex structure J is a smooth section of TM, i.e., $J \in \mathfrak{X}(M)$ such that $(J_m)^2 = -id$ where J_m is the restriction on T_mM.*

Definition 3.1.3.4 (Almost Complex manifold). *If M^{2n} admits an almost complex structure, it is called an almost complex manifold.*

Example 3.1.3.1. *If we think an n-dimensional complex manifold as a $2n$-dimensional real smooth manifold, then $\sqrt{-1}$, which comes from a holomorphic structure of M, gives an almost complex structure on the $2n$-dimensional real smooth manifold. Hence we conclude that a complex manifold induces an almost complex structure on its underlying smooth manifold.*

Remark 3.1.3.1. *Obviously, an almost complex structure gives a complex structure on each tangent space T_mM. However, even though each tangent space T_mM having a complex structure does not guarantee that M has an almost complex structure. Clearly, the obstruction is the smoothness of patching process. That is a problem of integrability. Newlander and Nirenberg proved that integrable almost complex manifold has a unique holomorphic coordinate patches, which is induced form the almost complex structure. Of course, we have not defined what an integrable almost complex structure means. This is the content of the next discussion.*

3.1.4 Integrable almost Complex Structure

Let M^{2n} be a $2n$-dimensional real smooth manifold. Suppose that T_mM has a complex structure J_m. By complexifying the tangent space, we have $2n$-dimensional complex vector space $T_mM \otimes_{\mathbb{R}} \mathbb{C}$. Extend J_m to $T_m \otimes_{\mathbb{R}} \mathbb{C}$ in the following ways: For $c \in \mathbb{C}$ and $v \in T_mM$,

$$\widetilde{J_m}(v \otimes c) = J_m(v) \otimes c.$$

So, $\widetilde{J_m}$ has eigenvalues $i, -i$. Let the eigenspaces corresponding to $i, -i$ be

$$T_m^{1,0}M = \{w \in T_mM \otimes_{\mathbb{R}} \mathbb{C} \mid \widetilde{J_m}(w) = iw\}$$
$$T_m^{0,1}M = \{w \in T_mM \otimes_{\mathbb{R}} \mathbb{C} \mid \widetilde{J_m}(w) = -iw\}$$

Note that $T_m^{1,0}M$ and $T_m^{0,1}M$ are n-dimensional complex vector spaces. So, identifying $T_m^{1,0}M$ and $T_m^{0,1}M$ as $2n$-dimensional real vector spaces, respectively, we have

$$T_m^{1,0}M \cong_{\mathbb{R}} T_m^{0,1}M \cong_{\mathbb{R}} T_mM \text{ as } 2n\text{-dimensional real vector spaces.}$$

So, now if we assume that TM has an almost complex structure J as a smooth section, then we have

$$T^{1,0}M = \{w \in TM \otimes_{\mathbb{R}} \mathbb{C} \mid \widetilde{J}(w) = iw\}$$
$$T^{0,1}M = \{w \in TM \otimes_{\mathbb{R}} \mathbb{C} \mid \widetilde{J}(w) = -iw\}.$$

Definition 3.1.4.1 (Integrability of J). *We say that an almost complex structure J is integrable if $\Gamma(T^{1,0}M)$ and $\Gamma(T^{0,1}M)$ are complex Lie subalgebras of $\Gamma(TM \otimes_{\mathbb{R}} \mathbb{C})$ where Γ means the set of smooth sections.*

Now, we show a **differential form version of integrability of** J, since as always forms are much more powerful tools than vector fields.

Notation 3.1.4.1. *Let M be a smooth $2n$-dimensional real manifold with an almost complex structure J. Now, we fix notations in this subsection as follows:*

$$\mathsf{E}^r(M) = \Gamma(\wedge^r T^*M) \text{ and } \mathcal{E}^r(M) = \Gamma(\wedge^r (TM \otimes_{\mathbb{R}} \mathbb{C})^*).$$

Note that $\mathcal{E}^r(M)$ is nothing but the set of complex valued $C^\infty(M)$ r-forms. Since

$$\mathbf{Hom}(TM \otimes_{\mathbb{R}} \mathbb{C}, \mathbb{C}) = (TM \otimes_{\mathbb{R}} \mathbb{C})^* = (T^{1,0}M)^* \oplus (T^{0,1}M)^*,$$

we also have

$$\mathcal{E}^{p,q}(M) = \Gamma(\wedge^p (T^{1,0}M)^*) \oplus \Gamma(\wedge^q (T^{0,1}M)^*). \text{ So, } \mathcal{E}^r(M) = \sum_{p+q=r} \mathcal{E}^{p,q}(M).$$

Since we have the exterior derivation d on $\mathsf{E}^r(M)$, we can extend d on $\mathcal{E}^r(M)$. So,

$$d : \mathcal{E}^{p,q} \to \sum_{r+s=p+q+1} \mathcal{E}^{r,s}(M).$$

Clearly, $d \circ d = 0$. Letting a projection $\pi_{p,q} : \mathcal{E}^{p+q} \to \mathcal{E}^{p,q}(M)$, we also define on $\mathcal{E}^{p,q}(M)$,

$$\partial = \pi_{p+1,q} \circ d \text{ and } \overline{\partial} = \pi_{p,q+1} \circ d.$$

Note that in this description we have

$$d\mathcal{E}^{1,0} = \partial\mathcal{E}^{1,0}(M) + \overline{\partial}\mathcal{E}^{1,0} + \pi_{0,2} \circ d\mathcal{E}^{1,0} \subseteq \sum_{p+q=2} \mathcal{E}^{p,q}(M)$$
$$d\mathcal{E}^{0,1} = \partial\mathcal{E}^{0,1}(M) + \overline{\partial}\mathcal{E}^{0,1} + \pi_{2,0} \circ d\mathcal{E}^{0,1} \subseteq \sum_{p+q=2} \mathcal{E}^{p,q}(M).$$

Theorem 3.1.4.1. *J is integrable if and only if $\pi_{0,2} \circ d\mathcal{E}^{1,0} \equiv 0$ and $\pi_{2,0} \circ d\mathcal{E}^{0,1} \equiv 0$.*

Proof. If J is integrable, then by definition $\Gamma(T^{1,0}M)$ and $\Gamma(T^{0,1}M)$ are complex Lie subalgebras of $\Gamma(TM \otimes_{\mathbb{R}} \mathbb{C})$. Let

$$\mathbf{Ann}^1(T^{1,0}M) = \{\omega \in \mathcal{E}^1(M) \mid \omega(X) = 0 \text{ for all } X \in \Gamma(T^{1,0}M)\}.$$

Clearly,

$$\mathbf{Ann}^1(T^{1,0}M) = \mathcal{E}^{0,1}.$$

Since Theorem 2.2.3.3 is also valid in this case, we have

$$d\omega \in \mathbf{Ann}^2(T^{1,0}M) = \mathcal{E}^{1,1}(M) \oplus \mathcal{E}^{0,2}.$$

Since $d\omega = \partial\omega + \overline{\partial}\omega + \pi_{2,0} \circ d\omega$, we conclude that $\pi_{2,0} \circ d\mathcal{E}^{0,1} \equiv 0$. By the same reasoning, we have $\pi_{0,2} \circ d\mathcal{E}^{1,0} \equiv 0$. Also, we get the converse by imitating the proof by backward. □

Since we always have $d = \partial + \overline{\partial}$ on $\mathcal{E}^0(M)$, and $\mathcal{E}^1(M)$ and $\mathcal{E}^0(M)$ generate $\mathcal{E}^*(M)$, by Theorem 3.1.4.1, we have the following definition.

Definition 3.1.4.2. *J is integrable if $d = \partial + \overline{\partial}$ on $\mathcal{E}^*(M)$.*

Note that if $d = \partial + \overline{\partial}$, then $d \circ d = 0$ implies that

$$\partial \circ \partial = 0 \text{ and } \overline{\partial} \circ \overline{\partial} = 0.$$

Example 3.1.4.1. *Let $M = \mathbb{R}^2$. Then TM has a basis $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$. Suppose an almost complex structure J is given by*

$$J(\frac{\partial}{\partial x}) = \frac{\partial}{\partial y} \text{ and } J(\frac{\partial}{\partial y}) = -\frac{\partial}{\partial x}.$$

Note that $J^2 = -id$. So, an element in 2-dimensional complex vector space $TM \otimes_{\mathbb{R}} \mathbb{C}$ is given by

$$(a+ib)\frac{\partial}{\partial x} + (c+id)\frac{\partial}{\partial y} \text{ where } a, b, c, d, \in \mathbb{R}.$$

Since $T^{1,0}M = \{w \in TM \otimes_{\mathbb{R}} \mathbb{C} \mid \widetilde{J}(w) = iw\}$, we deduce that $T^{1,0}(M)$ is 1-dimensional complex vector space generated by

$$\frac{1}{2}(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}), \text{ which is often denoted by } \frac{\partial}{\partial z}.$$

Similarly, we have $T^{0,1}(M)$ is 1-dimensional complex vector space generated by

$$\frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}), \text{ which is often denoted by } \frac{\partial}{\partial \overline{z}}.$$

Note that we call an element of $\mathbf{ker}\overline{\partial}$ a holomorphic function on $\mathbb{R}^2$ where

$$\mathcal{E}^0(\mathbb{R}^2) \xrightarrow{d=\partial+\overline{\partial}} \mathcal{E}^{1,0}(\mathbb{R}^2) \oplus \mathcal{E}^{0,1}(\mathbb{R}^2).$$

So, if f is holomorphic, then

$$0 = \overline{\partial} f = \frac{\partial f}{\partial \overline{z}} d\overline{z} \Longrightarrow \frac{1}{2}(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}) = 0.$$

By letting $f = g + ih$ where g, h are smooth real-valued functions, we have the Cauchy-Riemann equation.

Exercise 3.1.4.1. *Find a nonintegrable almost complex structure.*

Information. Note that J is integrable if and only if the Nijenhuis tensor $N(X,Y) = 2([JX,JY] - [X,Y] - J[X,JX] - J[JX,Y]) = 0$ for all $X, Y \in \mathfrak{X}(M)$. There is a nonintegrable almost complex structure on S^6, which is induced by the Cayley numbers. See the Kobayashi's book or Frölicher's paper " Zur Differetialgeometrie der komplexen Struckturen ", Math. Ann. 129 (1955), 50-95. □

Example 3.1.4.2. *Every even dimensional Lie group G admits a left-invariant almost complex structure.*

Proof. Note that we always have a complex structure on T_eG. Let J_e be a complex structure on T_eG. So, since $T_eG \cong \mathfrak{g}$, letting

$$J_g(X_g) = dl_g \circ J_e \circ dl_{g^{-1}}(X_g),$$

we have an almost complex structure J. By construction, $(dl_g)_e \circ J_e = J_g$, which means J is left-invariant. □

Note that a complex Lie group is a complex manifold with a holomorphic group operation. From example 3.1.3.1, we know that the complex structure of a complex Lie group induces an almost complex structure J.

Example 3.1.4.3. *If G is a complex manifold and J is the almost complex structure induced from $\sqrt{-1}$, then J is bi-invariant.*

Proof. Since J acts on the complex Lie algebra $\mathfrak{g}$ by multiplication by $\sqrt{-1}$ from $J^2 = -id$, we deduce that J is invariant under any element of $\mathbf{Aut}(\mathfrak{g})$ which includes dl_g and $\mathbf{Ad}(g) = dr_{g^{-1}}$. Hence, J is bi-invariant. Note that $\mathbf{Ad}(g) \circ J_e = J_e \circ \mathbf{Ad}(g)$. □

Note that one of necessary conditions that we can make a $2n$-dimensional real Lie group G into an n-dimensional complex Lie group is that G admits a bi-invariant (almost) complex structure.

3.1.5 Darboux Derivative

Let M be a smooth manifold and G be a Lie group. Suppose that $f : M \to G$ is a smooth map. So,we have

$$T_mM \xrightarrow{df} T_{f(m)}G \xrightarrow{dl_{(f(m))^{-1}}} T_eG = \mathfrak{g}.$$

Definition 3.1.5.1 (Darboux Derivatives). *We define the Darboux derivative Df of f to be*

$$(Df)_m \overset{def}{=} dl_{(f(m))^{-1}} \circ df.$$

Note that Df is nothing but a $\mathfrak{g}$-valued 1-form: Since $dl_{(f(m))^{-1}} \circ df : T_mM \to \mathfrak{g}$, for $v \in T_mM$ we have $(Df)_m(v) \in \mathfrak{g}$. We know that $f : M \to G$ induces $f^* : \Omega^1(G, \mathfrak{g}) \to \Omega^1(M, \mathfrak{g})$ by $f^*(\omega \otimes X) = (f^*\omega) \otimes X$.

Theorem 3.1.5.1. *Let $X_1, \ldots, X_n$ be a basis of $\mathfrak{g}$ and $\omega_1, \ldots, \omega_n$ the corresponding dual basis of $\mathfrak{g}^*$, we have*

$$f^*(\sum_{k=1}^{n} \omega_k \otimes X_k) = Df.$$

Proof. It suffices to show that

$$\sum_{k=1}^{n}(f^*\omega_k)_m \otimes (X_k)_e = (Df)_m.$$

Note that since ω_i is a left-invariant 1-form, we have

$$(\omega_i)_h = (dl_g^*\omega)_h \text{ for all } g, h \in G.$$

Let $v \in T_mM$. Then $(Df)_m(v) = dl_{(f(m))^{-1}} \circ df(v) = Y_e$ where Y is the unique left-invariant vector field corresponding $Y_{f(m)} = df(v)$. Let $Y_e = \sum_{k=1}^{n} a_k(X_k)_e$. So, $a_k = (\omega(Y))_e$. Hence, we also have

$$\begin{aligned}
\sum_{k=1}^{n}(f^*\omega_k)_m(v) \otimes (X_k)_e &= \sum_{k=1}^{n}((\omega_k)_{f(m)} \circ df(v)) \otimes (X_k)_e \\
&= \sum_{k=1}^{n}((l^*_{(f(m))^{-1}}\omega_k)_{f(m)} \circ df(v)) \otimes (X_k)_e \\
&= \sum_{k=1}^{n}((\omega_k)_{(f(m))^{-1}\cdot f(m)} \circ dl_{(f(m))^{-1}} \circ df(v)) \otimes (X_k)_e \\
&= \sum_{k=1}^{n}((\omega_k)_e \circ dl_{(f(m))^{-1}} \circ df(v)) \otimes (X_k)_e \\
&= \sum_{k=1}^{n}((\omega_k)_e(Y_e)) \otimes (X_k)_e \\
&= \sum_{k=1}^{n} a_k(X_k)_e = Y_e = (Df)_m(v).
\end{aligned}$$

□

Clearly, f^* commutes d:

$$\begin{array}{ccc}
\Omega^1(G,\mathfrak{g}) & \xrightarrow{f^*} & \Omega^1(M,\mathfrak{g}) \\
d\downarrow & & \downarrow d \\
\Omega^2(G,\mathfrak{g}) & \xrightarrow{f^*} & \Omega^2(M,\mathfrak{g})
\end{array}$$

Moreover, it is easy to see that f is a Lie algebra homomorphism:

$$\begin{array}{ccc}
\Omega^p(G,\mathfrak{g}) \times \Omega^q(G,\mathfrak{g}) & \xrightarrow{f^*} & \Omega^p(M,\mathfrak{g}) \times \Omega^q(M,\mathfrak{g}) \\
[,]\downarrow & & \downarrow [,] \\
\Omega^{p+q}(G,\mathfrak{g}) & \xrightarrow{f^*} & \Omega^{p+q}(M,\mathfrak{g})
\end{array}$$

Combining these with Theorem 3.1.5.1 and Maurer-Cartan structure equation, we have

$$\begin{aligned}
d(Df) + \frac{1}{2}[Df, Df] &= d(f^*(\sum_{k=1}^{n}\omega_k \otimes X_k)) + \frac{1}{2}[f^*(\sum_{k=1}^{n}\omega_k \otimes X_k), f^*(\sum_{k=1}^{n}\omega_k \otimes X_k)] \\
&= f^*(d(\sum_{k=1}^{n}\omega_k \otimes X_k) + \frac{1}{2}[\sum_{k=1}^{n}\omega_k \otimes X_k, \sum_{k=1}^{n}\omega_k \otimes X_k]) = 0.
\end{aligned}$$

That is, every Darboux derivative of a smooth map f satisfies the Maurer-Cartan structure equation.

Exercise 3.1.5.1.

$$D(l_g \circ f) = Df.$$

Proof.

$$\begin{aligned}(D(l_g \circ f))_m(v) &= dl_{(g\cdot f(m))^{-1}} \circ dl_g \circ df(v) \\ &= dl_{(f(m))^{-1}\cdot g^{-1}\cdot g} \circ df(v) = dl_{(f(m))^{-1}} \circ df(v) \\ &= (Df)_m(v).\end{aligned}$$

□

Amazingly, the converse is also true. Notice that this is similar to a local Poincaré lemma.

Theorem 3.1.5.2. *Let G be a Lie group with the Lie algebra $\mathfrak{g}$. Given $\eta \in \Omega^1(M, \mathfrak{g})$ such that $d\eta + \frac{1}{2}[\eta, \eta] = 0$, then for each point $p \in M$, there is a neighborhood U of p and a smooth map $f : U \to G$ such that*

$$Df = \eta|_U.$$

Proof. Note that as we already have seen, constructing a smooth map $f : W \to G$ is equivalent to finding a plane field on $W \times G$ having $\mathsf{graph}(f)$ as an integral submanifold of $W \times G$. Let π_W and π_G be projections from $W \times G$ to W and G, respectively, where W is a neighborhood of p. Letting $\omega = \sum_{k=1}^n \omega_k \otimes X_k$ the Maurer-Cartan form, we define recalling $\eta|_W \in \Omega^1(W, \mathfrak{g})$,

$$\Phi = \pi_G^*\omega - \pi_W^*(\eta|_W) \in \Omega^1(W \times G, \mathfrak{g}).$$

Let

$$E = \mathbf{ker}(\Phi) = \bigcap_{i=1}^n \mathbf{ker}(\Phi_i) \text{ where } \Phi = \sum_{i=1}^n \Phi_i \otimes X_i, \text{ i.e., } \Phi_i \in \mathfrak{L}\Omega^1(W \times G).$$

Note that if $Y \in \mathfrak{g}$, we have $Y = \sum_{k=1}^n a_k X_k$ where a_k are smooth. Since $\{X_1, \ldots, X_n\}$ is a basis, a_k are necessarily constant by the equation $a_k(e) \cdot dl_g(X_e) = dl_g((a_k(e)X_e) = a_k(g)X_g$. So, by the $\mathbb{R}$-linearity of tensor product, we can decompose

$$\Phi = \sum_{i=1}^n \Phi_i \otimes X_i, \text{ i.e., } \Phi_i \in \mathfrak{L}\Omega^1(W \times G).$$

The claim is that E is an integrable distribution of $T(W \times G) = TW \times TG$. By the Frobenius theorem, it suffices to show that $d\mathbf{Ann}^1(E) \subset \mathbf{Ann}^2(E)$. Note that $\mathbf{Ann}^1(E)$ is generated by Φ by construction. So, it is sufficient to show that $d\Phi \in \mathbf{Ann}^2(E)$. Then

$$\begin{aligned}d\Phi &= \pi_G^* d\omega - \pi_W^* d(\eta|_W) = -\frac{1}{2}(\pi_G^*[\omega, \omega] - \pi_W^*[\eta|_W, \eta|_W]) \\ &= -\frac{1}{2}([\pi_G^*\omega, \pi_G^*\omega] - [\pi_W^*(\eta|_W), \pi_W^*(\eta|_W)]) \\ &= -\frac{1}{2}([\pi_G^*\omega - \pi_W^*(\eta|_W), \pi_G^*\omega - \pi_W^*(\eta|_W)] + [\pi_G^*\omega - \pi_W^*(\eta|_W), \pi_W^*(\eta|_W)] + [\pi_W^*(\eta|_W), \pi_G^*\omega - \pi_W^*(\eta|_W)]) \\ &= -\frac{1}{2}([\Phi, \Phi] + [\Phi, \pi_W^*(\eta|_W)] + [\pi_W^*(\eta|_W), \Phi]).\end{aligned}$$

If $Y_1, Y_2 \in E$, then $\Phi(Y_1) = \Phi(Y_2) = 0$, since $E = \bigcap_{i=1}^n \mathbf{ker}(\Phi_i)$. Hence $d\Phi_i(Y_1, Y_2) = 0$. That is, $d\Phi \in \mathbf{Ann}^2(E)$. Now, the only remaining thing to show is that there exists a neighborhood $U \subseteq W$ of p such that

$$Df = \eta|_U.$$

Recalling Exercise 2.2.4.3, since $\Phi = \pi_G^*\omega - \pi_W^*(\eta|_W)$ generates the differential ideal $\mathbf{Ann}^1(E)$ in $\Omega^1(W \times G, \mathfrak{g})$, there do exist a neighborhood $U \subseteq W$ of p and a smooth map $f : U \to G$ such that $f^*\omega = \eta|_U$. By Theorem 3.1.5.1, we have

$$\eta|_U = f^*\omega = f^*(\sum_{k=1}^n \omega_k \otimes X_k) = Df.$$

□

3.1.6 When M is $\mathbf{GL}(n, \mathbb{R})$ and G is $\mathbf{GL}(n, \mathbb{R})$ as a representations

Recall that the exterior derivative d is defined on $\Omega^*(M)$ where M is a smooth manifold. Suppose that $g \in \mathbf{GL}(n, \mathbb{R})$ as a manifold. Clearly, we have a representation

$$g \mapsto \begin{pmatrix} x_{11}(g) & \cdots & x_{1,n}(g) \\ \vdots & \ddots & \vdots \\ x_{n1}(g) & \cdots & x_{nn}(g) \end{pmatrix}$$

We will denote this map by $\mathcal{X}$, which is nothing but the identity map, from $\mathbf{GL}(n, \mathbb{R})$ to $\mathbf{GL}(n, \mathbb{R})$:

$$\mathcal{X} : \mathbf{GL}(n, \mathbb{R}) \to \mathbf{GL}(n, \mathbb{R}),$$

So, regarding the coordinate functions x_{ij} as elements of $\Omega^0(\mathbf{GL}(n, \mathbb{R}))$, it makes sense to define $d\mathcal{X}$ by

$$d\mathcal{X}_g \overset{def}{=} \begin{pmatrix} dx_{11}(g) & \cdots & dx_{1,n}(g) \\ \vdots & \ddots & \vdots \\ dx_{n1}(g) & \cdots & dx_{nn}(g) \end{pmatrix}$$

Clearly, $d\mathcal{X}$ is an $(n \times n)$ matrix of 1-forms on $\mathbf{GL}(n, \mathbb{R})$, i.e., $d\mathcal{X}_g \in \mathbf{M}_n(\mathbb{R})$ and $dx_{ij} \in \Omega^1(\mathbf{GL}(n, \mathbb{R}))$. Since the Lie algebra of Lie group $\mathbf{GL}(n, \mathbb{R})$ is $\mathbf{M}_n(\mathbb{R})$;

$$\mathfrak{gl}(n, \mathbb{R}) = \mathbf{M}_n(\mathbb{R}),$$

we conclude that $d\mathcal{X}$ is a $\mathfrak{gl}(n, \mathbb{R})$-valued 1-form:

$$d\mathcal{X} \in \Omega^1(\mathbf{GL}(n, \mathbb{R}), \mathfrak{gl}(n, \mathbb{R})) = \Omega^1(\mathbf{GL}(n, \mathbb{R}), \mathbf{M}_n(\mathbb{R})).$$

Hence, we have

$$d\mathcal{X}_h : T_h(\mathbf{GL}(n, \mathbb{R})) \to \mathfrak{gl}(n, \mathbb{R}) \cong \mathbf{M}_n(\mathbb{R}).$$

An explicit formula is given by for $v \in T_h(\mathbf{GL}(n, \mathbb{R}))$,

$$(d\mathcal{X}_h)(v) = \begin{pmatrix} dx_{11}(h)(v) & \cdots & dx_{1,n}(h)(v) \\ \vdots & \ddots & \vdots \\ dx_{n1}(h)(v) & \cdots & dx_{nn}(h)(v) \end{pmatrix} = \begin{pmatrix} (v(x_{11}))(h) & \cdots & (v(x_{1,n}))(h) \\ \vdots & \ddots & \vdots \\ (v(x_{n1}))(h) & \cdots & (v(x_{nn}))(h) \end{pmatrix}$$

Note that we regard v as a differential operator so that $v(x_{ij}) \in C^\infty(\mathbf{GL}(n,\mathbb{R}))$. If we identify v with $n \times n$-matrix;

$$v = \begin{pmatrix} a_{11} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in \mathbf{M}_n(\mathbb{R}),$$

it is easy to see that $(v(x_{ij}))(g) = a_{ij}$, a constant function on $\mathbf{GL}(n,\mathbb{R})$. So, the above formula becomes

$$(d\mathcal{X}_h)(v) = v.$$

Actually, an easier way to see this is the following: We know that $\mathcal{X} : \mathbf{GL}(n,\mathbb{R}) \to \mathbf{GL}(n,\mathbb{R})$ is the identity map. So,

$$d\mathcal{X} : T(\mathbf{GL}(n,\mathbb{R})) \to T(\mathbf{GL}(n,\mathbb{R})) \text{ is the identity map.}$$

Hence,

$$d\mathcal{X}_h : T_h(\mathbf{GL}(n,\mathbb{R})) \to T_h(\mathbf{GL}(n,\mathbb{R})) \cong \mathbf{M}_n(\mathbb{R})$$

is the identity map. Now, we define $l_h^* d\mathcal{X}$: For $v \in T_p(\mathbf{GL}(n,\mathbb{R}))$,

$$(l_h^* d\mathcal{X})_p(v) = d\mathcal{X}_{hp}(dl_h(v)) = \begin{pmatrix} dx_{11}(hp)(dl_h(v)) & \cdots & dx_{1,n}(hp)(dl_h(v)) \\ \vdots & \ddots & \vdots \\ dx_{n1}(hp)(dl_h(v)) & \cdots & dx_{nn}(hp)(dl_h(v)) \end{pmatrix}.$$

Note that $dl_h(v) \in T_{hp}(\mathbf{GL}(n,\mathbb{R}))$ and $l_h^* : \Omega^1(\mathbf{GL}(n,\mathbb{R}), \mathbf{M}_n(\mathbb{R})) \to \Omega^1(\mathbf{GL}(n,\mathbb{R}), \mathbf{M}_n(\mathbb{R}))$. From this, it is easy to see that $d\mathcal{X}$ is not a left-invariant $\mathfrak{gl}(n,\mathbb{R})$-valued 1-form: Note that by the above discussion, we have

$$(d\mathcal{X}_p)(v) = v \text{ and } d\mathcal{X}_{hp}(dl_h(v)) = dl_h(v).$$

Since $dl_h(v) = \frac{d}{dt}|_{t=0} l_h(p + tv) = hv$, we have

$$(l_h^* d\mathcal{X})_p(v) = d\mathcal{X}_{hp}(dl_h(v)) = dl_h(v) = hv = h \cdot (d\mathcal{X})_p(v).$$

Notice that

$$h \cdot (d\mathcal{X})_p(v) = \begin{pmatrix} x_{11}(h) & \cdots & x_{1,n}(h) \\ \vdots & \ddots & \vdots \\ x_{n1}(h) & \cdots & x_{nn}(h) \end{pmatrix} \cdot \begin{pmatrix} (v(x_{11}))(p) & \cdots & (v(x_{1,n}))(p) \\ \vdots & \ddots & \vdots \\ (v(x_{n1}))(p) & \cdots & (v(x_{nn}))(p) \end{pmatrix}$$

Now, define a map

$$\mathcal{X}^{-1} : \mathbf{GL}(n,\mathbb{R}) \to \mathbf{GL}(n,\mathbb{R}) \text{ by } \mathcal{X}^{-1}(g) \mapsto \begin{pmatrix} x_{11}(g^{-1}) & \cdots & x_{1,n}(g^{-1}) \\ \vdots & \ddots & \vdots \\ x_{n1}(g^{-1}) & \cdots & x_{nn}(g^{-1}) \end{pmatrix}$$

Similar procedures shall give a $\mathfrak{gl}(n,\mathbb{R})$-valued 1-form $\mathcal{X}^{-1} \cdot d\mathcal{X} \in \Omega^1(\mathbf{GL}(n,\mathbb{R}), \mathbf{M}_n(\mathbb{R}))$. Note that for $g \in \mathbf{GL}(n\mathbb{R})$, we have

$$\mathcal{X}_g^{-1} \cdot (d\mathcal{X})_g = \begin{pmatrix} x_{11}(g^{-1}) & \cdots & x_{1,n}(g^{-1}) \\ \vdots & \ddots & \vdots \\ x_{n1}(g^{-1}) & \cdots & x_{nn}(g^{-1}) \end{pmatrix} \cdot \begin{pmatrix} dx_{11}(g) & \cdots & dx_{1,n}(g) \\ \vdots & \ddots & \vdots \\ dx_{n1}(g) & \cdots & dx_{nn}(g) \end{pmatrix} \in \mathbf{M}_n(\mathbb{R}).$$

For $v \in T_p(\mathbf{GL}(n,\mathbb{R}))$, we define

$$(l_h^*(\mathcal{X}^{-1} \cdot d\mathcal{X}))_p(v) = \mathcal{X}_{hp}^{-1} \cdot d\mathcal{X}_{hp}(dl_h(v)).$$

Theorem 3.1.6.1. *$\mathcal{X}^{-1} \cdot d\mathcal{X}$ is a left-invariant $\mathfrak{gl}(n,\mathbb{R})$-valued 1-form.*

Proof. Since $\mathcal{X}^{-1} : \mathbf{GL}(n,\mathbb{R}) \to \mathbf{GL}(n,\mathbb{R})$ is a homomorphism with $\mathcal{X}_{hp}^{-1} = \mathcal{X}_p^{-1} \cdot \mathcal{X}_h^{-1}$ and it is not hard to see that $\mathcal{X}_h^{-1} = \mathcal{X}_{h^{-1}} = h^{-1}$, we have for $v \in T_p(\mathbf{GL}(n,\mathbb{R}))$,

$$\begin{aligned}(l_h^*(\mathcal{X}^{-1} \cdot d\mathcal{X}))_p(v) &= \mathcal{X}_{hp}^{-1} \cdot d\mathcal{X}_{hp}(dl_h(v)) = \mathcal{X}_{hp}^{-1} \cdot h \cdot (d\mathcal{X})_p(v) \\ &= \mathcal{X}_p^{-1} \cdot \mathcal{X}_h^{-1} \cdot h \cdot (d\mathcal{X})_p(v) = \mathcal{X}_p^{-1} \cdot h^{-1} \cdot h \cdot (d\mathcal{X})_p(v) \\ &= \mathcal{X}_p^{-1} \cdot (d\mathcal{X})_p(v) = (\mathcal{X}^{-1} \cdot d\mathcal{X})_p(v).\end{aligned}$$

□

Example 3.1.6.1. *Let $\mathbb{R}_+$ be a connected component of $\mathbf{GL}(1,\mathbb{R})$. Note that we still have*

$$T_g(\mathbb{R}_+) = \mathbf{M}_1(\mathbb{R}) = \mathbb{R} \text{ for } g \in \mathbb{R}_+.$$

Let x be a coordinate function $x : \mathbb{R}_+ \to \mathbb{R}$, which is the identity map. Now, it is easy to see that

$$x^{-1}dx = d\log x.$$

It is obvious that dx is not left-invariant but $d\log x$ is left-invariant.

Remark 3.1.6.1. *Imitating all the proofs so far, we also conclude that*

$$(d\mathcal{X}) \cdot \mathcal{X}^{-1}$$

is a right-invariant Maurer-Cartan form.

Now we calculate $d\mathcal{X}^{-1}$. First note that

$$\mathcal{X}_g \cdot \mathcal{X}_g^{-1} = \begin{pmatrix} x_{11}(g) & \cdots & x_{1,n}(g) \\ \vdots & \ddots & \vdots \\ x_{n1}(g) & \cdots & x_{nn}(g) \end{pmatrix} \cdot \begin{pmatrix} x_{11}(g^{-1}) & \cdots & x_{1,n}(g^{-1}) \\ \vdots & \ddots & \vdots \\ x_{n1}(g^{-1}) & \cdots & x_{nn}(g^{-1}) \end{pmatrix} = \mathbf{Id}_{(n\times n)}.$$

Hence, we have for any $g \in \mathbf{GL}(n,\mathbb{R})$,

$$(d\mathcal{X})_g \cdot \mathcal{X}_g^{-1} + \mathcal{X}_g \cdot (d\mathcal{X}^{-1})_g = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = \mathbf{0}_{(n\times n)}.$$

Therefore,

$$\begin{aligned}(d\mathcal{X}^{-1})_g &= -\mathcal{X}_g^{-1} \cdot (d\mathcal{X})_g \cdot \mathcal{X}_g^{-1} \\ &= -\begin{pmatrix} x_{11}(g^{-1}) & \cdots & x_{1,n}(g^{-1}) \\ \vdots & \ddots & \vdots \\ x_{n1}(g^{-1}) & \cdots & x_{nn}(g^{-1}) \end{pmatrix} \cdot \begin{pmatrix} dx_{11}(g) & \cdots & dx_{1,n}(g) \\ \vdots & \ddots & \vdots \\ dx_{n1}(g) & \cdots & dx_{nn}(g) \end{pmatrix} \cdot \begin{pmatrix} x_{11}(g^{-1}) & \cdots & x_{1,n}(g^{-1}) \\ \vdots & \ddots & \vdots \\ x_{n1}(g^{-1}) & \cdots & x_{nn}(g^{-1}) \end{pmatrix}.\end{aligned}$$

Example 3.1.6.2. *Let $\mathbb{R}_+$ be a connected component of $\mathbf{GL}(1,\mathbb{R})$. Let x be a coordinate function and $x^{-1} : \mathbb{R}_+ \to \mathbb{R}$, which is the inversion map. Note that x^{-1} is well-defined on $\mathbb{R}_+$. Now, it is easy to see that*

$$dx^{-1} = -x^{-2}dx.$$

Remark 3.1.6.2. *We know that the identity map $\mathcal{X} : \mathbf{GL}(n\mathbb{R}) \to \mathbf{GL}(n,\mathbb{R})$ induces the identity map*

$$(d\mathcal{X})_g : T_g(\mathbf{GL}(n,\mathbb{R})) \to T_g(\mathbf{GL}(n,\mathbb{R})) \text{ for all } g \in \mathbf{GL}(n,\mathbb{R}).$$

However, the formula $(d\mathcal{X}^{-1})_g = -\mathcal{X}_g^{-1} \cdot (d\mathcal{X})_g \cdot \mathcal{X}_g^{-1}$ shows that the induced map $(d\mathcal{X}^{-1})_g$ on $T_g(\mathbf{GL}(n,\mathbb{R}))$ of the inversion map $\mathcal{X}^{-1} : \mathbf{GL}(n,\mathbb{R}) \to \mathbf{GL}(n,\mathbb{R})$ is not quite the identity. What is true is that $(d\mathcal{X}^{-1})_e = -id$ on $T_e(\mathbf{GL}(n,\mathbb{R})) = \mathfrak{gl}(n,\mathbb{R}) = \mathbf{M}_n(\mathbb{R})$ as an additive group when $g = e$ the identity. This is a general phenomenon. A law of composition of Lie groups induces a law of composition of their Lie algebras. The next two exercises shall give some clear view.

Exercise 3.1.6.1. *Show that*

$$(d\mathcal{X}^k)_g = \begin{cases} \sum_{i=1}^k \overbrace{\mathcal{X}_g \cdots \underbrace{(d\mathcal{X})_g}_{ith} \cdots \mathcal{X}_g}^{k} & k \text{ is a posistive integer} \\ \sum_{i=1}^k \overbrace{\mathcal{X}_g^{-1} \cdots \underbrace{(d\mathcal{X}^{-1})_g}_{ith} \cdots \mathcal{X}_g^{-1}}^{k} & k \text{ is a negative integer} \\ 0 & k = 0. \end{cases}$$

In particular. on $T_e(\mathbf{GL}(n,\mathbb{R})) = \mathfrak{gl}(n,\mathbb{R}) = \mathbf{M}_n(\mathbb{R})$, we have

$$(d\mathcal{X}^k)_e = k \cdot \mathbf{id}_{(n\times n)} \text{ for } k \in \mathbb{Z}.$$

Proof. This is nothing but the Leibniz rule of d. Also, note that

$$\mathcal{X}_e = (d\mathcal{X})_e = \mathcal{X}_e^{-1} = \mathbf{id}_{(n\times n)} \text{ and } (d\mathcal{X}^{-1})_e = -\mathbf{id}_{(n\times n)}.$$

□

Exercise 3.1.6.2. *Let $h_1 \in \mathbf{GL}(n,\mathbb{R})$ and define*

$$\mathcal{M} : \mathbf{GL}(n,\mathbb{R}) \to \mathbf{GL}(n,\mathbb{R}) \text{ by } \mathcal{M}(g) = \mathcal{X}(h_1) \cdot \mathcal{X}(g).$$

We have

$$(d\mathcal{M})_g : T_g(\mathbf{GL}(n,\mathbb{R})) \to T_{h_1 g}(\mathbf{GL}(n,\mathbb{R})) \text{ by } (d\mathcal{M})_g = \mathcal{X}_{h_1} \cdot (d\mathcal{X})_g.$$

Proof. Using the Leibniz rule of d and $(d\mathcal{X})_{h_1} = 0$ by the fact that $\mathcal{X}_{h_1}$ is constant as $g \in \mathbf{GL}(n,\mathbb{R})$ varies, it is done. □

The following example shall show that a law of composition of Lie groups induces a law of composition of their Lie algebras.

Example 3.1.6.3. *Define*

$$\mathcal{M} : \mathbf{GL}(n,\mathbb{R}) \times \mathbf{GL}(n,\mathbb{R}) \to \mathbf{GL}(n,\mathbb{R}) \text{ by } \mathcal{M}(g,h) = \mathcal{X}(g) \cdot \mathcal{X}(h) \cdot \mathcal{X}^{-1}(g) \cdot \mathcal{X}^{-1}(h).$$

We have

$$(d\mathcal{M})_{g,h} : T_g(\mathbf{GL}(n,\mathbb{R})) \times T_h(\mathbf{GL}(n,\mathbb{R})) \to T_{ghg^{-1}h^{-1}}(\mathbf{GL}(n,\mathbb{R})) \text{ by}$$

$$(d\mathcal{M})_{g,h} = (d\mathcal{X})_g \cdot \mathcal{X}_h \cdot \mathcal{X}_g^{-1} \cdot \mathcal{X}_h^{-1} + \mathcal{X}_g \cdot (d\mathcal{X})_h \cdot \mathcal{X}_g^{-1} \cdot \mathcal{X}_h^{-1} + \mathcal{X}_g \cdot \mathcal{X}_h \cdot (d\mathcal{X}^{-1})_g \cdot \mathcal{X}_h^{-1} + \mathcal{X}_g \cdot \mathcal{X}_h \cdot \mathcal{X}_g^{-1} \cdot (d\mathcal{X}^{-1})_h.$$

In particular, letting e be the identity, we have

$$(d\mathcal{M})_{e,h} : T_e(\mathbf{GL}(n,\mathbb{R})) \times T_h(\mathbf{GL}(n,\mathbb{R})) \to T_e(\mathbf{GL}(n,\mathbb{R})).$$

Suppose that $v_e \in T_e(\mathbf{GL}(n,\mathbb{R}))$ and $w_h \in T_h(\mathbf{GL}(n,\mathbb{R}))$. It is easy to see that from the above formula

$$(d\mathcal{M})_{e,h}(v_e, w_h) = (\mathbf{Id}_{n\times n})_e(v_e) + (\mathbf{Id}_{n\times n})_h(w_h) \cdot \mathcal{X}_h^{-1} - (\mathbf{Id}_{n\times n})_e(v_e) - (\mathbf{Id}_{n\times n})_h(w_h) \cdot \mathcal{X}_h^{-1} = \mathbf{0}_{(n\times n)}$$

In general, by abusing notation, we write

$$(d\mathcal{M})_{g,e} : T_g(\mathbf{GL}(n,\mathbb{R})) \times T_e(\mathbf{GL}(n,\mathbb{R})) \to T_e(\mathbf{GL}(n,\mathbb{R})) \text{ by } (d\mathcal{M})_{g,e} = 0$$
$$(d\mathcal{M})_{e,e} : T_e(\mathbf{GL}(n,\mathbb{R})) \times T_e(\mathbf{GL}(n,\mathbb{R})) \to T_e(\mathbf{GL}(n,\mathbb{R})) \text{ by } (d\mathcal{M})_{e,e} = 0.$$

The following example shall give some insight about what is really going on.

Example 3.1.6.4. *Define*

$$\mathcal{M} : \mathbf{GL}(n,\mathbb{R}) \times \mathbf{GL}(n,\mathbb{R}) \to \mathbf{GL}(n,\mathbb{R}) \text{ by } \mathcal{M}(g,h) = \mathcal{X}(g) \cdot \mathcal{X}(h).$$

This is nothing but a usual operation on a ***Lie group*** *as one can easily expect. We have*

$$(d\mathcal{M})_{g,h} : T_g(\mathbf{GL}(n,\mathbb{R})) \times T_h(\mathbf{GL}(n,\mathbb{R})) \to T_{gh}(\mathbf{GL}(n,\mathbb{R})) \text{ by}$$

$$(d\mathcal{M})_{g,h} = (d\mathcal{X})_g \cdot \mathcal{X}_h + \mathcal{X}_g \cdot (d\mathcal{X})_h.$$

In particular, letting e be the identity, we have

$$(d\mathcal{M})_{e,e} : T_e(\mathbf{GL}(n,\mathbb{R})) \times T_e(\mathbf{GL}(n,\mathbb{R})) \to T_e(\mathbf{GL}(n,\mathbb{R})).$$

Suppose that $v_e \in T_e(\mathbf{GL}(n,\mathbb{R}))$ and $w_e \in T_e(\mathbf{GL}(n,\mathbb{R}))$. It is easy to see that, reminding $d\mathcal{X}$ and $\mathcal{X}_e$ are the identities, from the above formula

$$(d\mathcal{M})_{e,e}(v_e, w_e) = (\mathbf{Id}_{n\times n})_e(v_e) + (\mathbf{Id}_{n\times n})_h(w_e) = v_e + w_e.$$

This is, the multiplication $\mathcal{M}$ on a ***Lie group*** *as a usual operation in a Lie group is transformed into the addition $(d\mathcal{M})_{e,e}$ on the* ***Lie algebra*** *as a usual operation in the Lie algebra.*

Now we show that $\mathcal{X}^{-1} \cdot (d\mathcal{X})$ satisfies the Maurer-Cartan equation.

Theorem 3.1.6.2. *Let $\omega = \mathcal{X}^{-1} \cdot (d\mathcal{X})$. Then*

$$d\omega + \frac{1}{2}[\omega, \omega] = 0.$$

Proof. Note that

$$d\omega = d(\mathcal{X}^{-1} \cdot (d\mathcal{X})) = (d\mathcal{X}^{-1}) \cdot d\mathcal{X} + \mathcal{X}^{-1} \cdot d(d\mathcal{X}) = -\mathcal{X}^{-1} \cdot (d\mathcal{X}) \cdot \mathcal{X}^{-1} \cdot (d\mathcal{X}).$$

We know that for $\alpha \otimes X, \beta \otimes Y \in \Omega^1(G, \mathfrak{g})$, we have defined

$$[\alpha \otimes X, \beta \otimes Y] = (\alpha \wedge \beta) \otimes [X, Y].$$

Recalling the matrix representation of $\mathcal{X}^{-1} \cdot d\mathcal{X}$, a moment of thought gives that

$$[\mathcal{X}^{-1} \cdot d\mathcal{X}, \mathcal{X}^{-1} \cdot d\mathcal{X}] = \mathcal{X}^{-1} \cdot (d\mathcal{X}) \cdot \mathcal{X}^{-1} \cdot (d\mathcal{X}) + \mathcal{X}^{-1} \cdot (d\mathcal{X}) \cdot \mathcal{X}^{-1} \cdot (d\mathcal{X}).$$

Hence,

$$d(\mathcal{X}^{-1} \cdot d\mathcal{X}) + \frac{1}{2}[\mathcal{X}^{-1} \cdot d\mathcal{X}, \mathcal{X}^{-1} \cdot d\mathcal{X}] = 0.$$

□

Theorem 3.1.6.3 (Ado's Theorem). *Every Lie algebra $\mathfrak{g}$ admits a faithful representation;*

$$\mathfrak{g} \hookrightarrow \mathfrak{gl}(n, \mathbb{R}).$$

Theorem 3.1.6.4. *If G is a Lie group with the Lie algebra $\mathfrak{g}$ and a sub Lie algebra $\mathfrak{h} \subseteq \mathfrak{g}$, then there exists a connected Lie subgroup H of G with the Lie algebra $\mathfrak{h}$.*

Proof. Let $\mathfrak{h} \subseteq \mathfrak{g} = \mathfrak{X}(G)$. Consider a plane field E on G spanned by $\mathfrak{h}$. So, we have

$$\Gamma(E) = \{\sum_i f_i X_i \mid X_i \in \mathfrak{h}, f_i \in C^\infty(G)\}.$$

Suppose that $A, B \in \Gamma(E)$. Letting $A = \sum a_i X_i$ and $B = \sum b_i X_i$ where $a_i, b_i \in C^\infty(G)$, we have

$$[A, B] = \sum_{i,j} a_i (X_i b_j) x_j + \sum_{i,j} a_i b_j [X_i, X_j] - \sum_{i,j} b_j (X_j a_i) X_i.$$

Since $\mathfrak{h}$ is a Lie subalgebra, we have $[X_i, X_j] \in \mathfrak{h}$. So, we conclude that $[A, B] \in \Gamma(E)$. hence, by the Frobenius theorem, there exists the maximal integral submanifold H containing $e \in G$. Note that by definition, a submanifold might not be an embedding and the maximality implies that H is connected. Now, we are going to show H is a Lie subgroup of G. Let $h \in H$. Clearly $e \in l_{h^{-1}}(H)$ and we deduce that E is a left-invariant plane field, i.e., $(l_{h^{-1}})_*(E) = dl_{h^{-1}}(E) = E$, since E is generated by $\mathfrak{h}$, which is also generated by the set of left-invariant vector fields. So, we conclude that $l_{h^{-1}}(H)$ is an integral submanifold of $(l_{h^{-1}})_*(E) = dl_{h^{-1}}(E) = E$. Without loss of generality, assuming $l_{h^{-1}}(H)$ to be maximal by extension, since $l_{h^{-1}}(H)$ also contains e, by the uniqueness of a maximal integrable submanifold through e, we conclude that

$$l_{h^{-1}}(H) = H.$$

So, we have that $l_{h^{-1}}(g) = h^{-1}g \in H$ for all $g, h \in H$. That is, H is stable under a C^∞ mapping $(g, h) \mapsto h^{-1}g$. Since $e \in H$, we conclude that H is a Lie subgroup of L. □

Example 3.1.6.5. *Let* $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ *be a torus. It is easy to see that*

$$T_{(0,0)}(T^2) = \mathfrak{g} = \{a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} \mid a, b \in \mathbb{R}\} = \mathbb{R}^2.$$

Notice that $\mathbb{R}^2$ *is an abelian Lie algebra. So, since any linear subspaces of* $\mathbb{R}^2$ *form Lie subalgebras of* $\mathbb{R}^2$*, there are the corresponding Lie subgroups of* T^2*. Suppose that* $\mathfrak{h} \subseteq \mathfrak{g}$ *is generated by*

$$a_0\frac{\partial}{\partial x} + b_0\frac{\partial}{\partial y}.$$

It is easy to see that the corresponding Lie subgroup H *is the skew-line of* T^2 *unless* $\frac{a_0}{b_0} \in \mathbb{Q}$*. So,* H *is not an embedding but a submanifold (immersion).*

We know that every homomorphism of Lie groups gives a Lie algebra homomorphism. Now, we investigate the converse.

Question 3.1.6.1. *Does every homomorphism of Lie algebras* $\mathfrak{g} \to \mathfrak{h}$ *arise in this way? That is, does every homomorphism of Lie algebras* $\mathfrak{g} \to \mathfrak{h}$ *come from a homomorphism of Lie groups* $G \to H$ *where* $\mathfrak{g}$ *and* $\mathfrak{h}$ *are the Lie algebras of* G *and* H*, respectively?*

In general, the answer is "No".

Example 3.1.6.6. *Let* $G = \mathbb{R}/\mathbb{Z} \cong S^1$ *and* $H = \mathbb{R}$*. Note that* G *is not simply connected. It is easy to see that the corresponding Lie algebras are* $\mathfrak{g} = \mathbb{R}$ *and* $\mathfrak{h} = \mathbb{R}$*. We shall show that every Lie group homomorphism from* G *to* H *induces a trivial Lie algebra homomorphism from* $\mathfrak{g}$ *to* $\mathfrak{h}$*. So, nontrivial Lie algebra homomorphisms from* $\mathfrak{g}$ *to* $\mathfrak{h}$*, which obviously exists, can not come from their Lie group homomorphism. Let* $G \xrightarrow{\varphi} H$ *be a Lie group homomorphism. Since*

$$\varphi \circ l_g(g_1) = \varphi(gg_1) = \varphi(g)\varphi(g_1) = l_{\varphi(g)} \circ \varphi(g_1),$$

we deduce that

$$(d\varphi)_g \circ dl_g = dl_{\varphi(g)} \circ (d\varphi)_e.$$

Since dl_g *and* $dl_{\varphi(g)}$ *are isomorphisms, we conclude that the linear map* $d\varphi$ *has a constant rank, i.e.,* $\textbf{rank}(d\varphi) =$ *constant. Note that it is well-known that any smooth map from a compact smooth manifold has a maximum. That is,* $d\varphi = 0$ *for some point in* G*. Hence, we have*

$$d\varphi \equiv 0.$$

However, if G is simply connected, we can give the affirmative answer to Question 3.1.6.1. Before we give a its proof, we need some prerequisites.

3.1.7 Covering spaces of manifolds

Suppose that $M \xrightarrow{f} N$ be a continuous map from a topological manifold M to a smooth manifold N. Assume that f is a local homeomorphism. That is, for all $m \in M$, there exists a neighborhood U_m of m such that $f|_{U_m}$ is a homeomorphism. Then by pulling back of the smooth structure of N by f, we have a locally unique differentiable structure on M. Since smoothness is a local notion, we infer that f become a smooth map. That is, there exists a smooth structure on M such that f is a local diffeomorphism.

Definition 3.1.7.1 (Covering spaces). *Let $M \xrightarrow{f} N$ be a continuous map. We say that f is a covering space if for all $y \in N$, there exists a neighborhood V_y of y in N such that for each component U of $f^{-1}(V_y)$, we have*

$$f|_U : U \to V_y \text{ is a homeomorphism.}$$

Note that the we say that M is evenly covered when the above property occurs.

Note that it is well-known that if $M \xrightarrow{f} N$ is a covering space, then the induced map on the fundamental groups $\pi_1(M) \xrightarrow{f_*} \pi_1(N)$ is injective. Also, from theorems in Algebraic topology, we also have that given N, there exists a covering space, which is called a universal covering space

$$\widetilde{N} \xrightarrow{f} N \text{ such that } \pi_1(\widetilde{N}) = 1.$$

In particular, this is unique up to homeomorphism and every other covering space of N is a quotient of $\widetilde{N}$. Reminding Exercise 1.2.7.4, we have the followings:

Theorem 3.1.7.1. *Every covering space of a Lie group is a Lie group.*

From the proof of Exercise 1.2.7.4, we know that the universal covering space of a connected Lie group H is given by

$$\widetilde{H} = \{[\gamma] \mid \gamma : [0,1] \to H \text{ where } \gamma \text{ is a path starting at } e \in H\}.$$

Moreover, the covering map $p : \widetilde{H} \to H$ is given by $p([\gamma]) = \gamma(1)$. So, it is easy to see that

$$p^{-1}(e) = \pi_1(H, e).$$

That is, since $\widetilde{H}$ is a group and $\widetilde{H} \xrightarrow{p} H$ is a homomorphism as we have seen, we conclude that

$$\mathbf{ker} p = \pi_1(H, e), \text{ i.e., an exact sequence } \pi_1(H, e) \to \widetilde{H} \to H.$$

Actually, since a connected Lie group is path connected, we have for any $p \in H$,

$$\pi_1(H, p) = \pi_1(H, e).$$

So, the sequence $\pi_1(H, e) \to \widetilde{H} \to H$ also tells us that $\widetilde{H}$ is a fiber bundle over H with fiber $\pi_1(H)$.

Example 3.1.7.1. *If $H = \mathbb{R}/\mathbb{Z} \cong S^1$, then we have $\widetilde{H} \cong \mathbb{R}$ and $\mathbb{R} \xrightarrow{p} \mathbb{R}/\mathbb{Z}$ with*

$$\mathbf{ker} p = \mathbb{Z} \cong \pi_1(S^1, e).$$

Note that if $H = \mathbb{R}^n/\mathbb{Z}^n \cong \underbrace{S^1 \times \cdots \times S^1}_{n} \overset{def}{=} T^n$, then we have $\widetilde{H} \cong \mathbb{R}^n$ and $\mathbb{R}^n \xrightarrow{p} \mathbb{R}^n/\mathbb{Z}^n$ with

$$\mathbf{ker} p = \mathbb{Z}^n \cong \pi_1(T^n, e).$$

Also, if $H = \mathbf{SO}(3) \cong \mathbb{RP}^3$, then we have $\widetilde{H} \cong \mathbf{SU}(2) \cong \mathbb{H}_1 \cong S^3$ and $\mathbf{SU}(2) \xrightarrow{p} \mathbf{SO}(3)$ with

$$\mathbf{ker} p = \mathbb{Z}_2 \cong \pi_1(\mathbf{SO}(3), e).$$

Note that $\pi_1(\mathbf{SO}(n)) = \mathbb{Z}_2$ and the universal covering space of $\mathbf{SO}(n)$ is denoted by

$$\widetilde{\mathbf{SO}(n)} = \mathbf{Spin}(n).$$

From this, it is good to ask the next question.

Question 3.1.7.1. *What group arises as* $\pi_1(G)$ *if* G *is a connected Lie group?*

The answer is that only **finitely generated abelian groups** can arise as $\pi_1(G)$ if G is a connected Lie group.

Theorem 3.1.7.2. *Let* G *be a connected Lie group. Then* $\pi_1(G)$ *is abelian.*

Proof. Let $\widetilde{G} \xrightarrow{\varphi} G$ be a universal covering space of a connected Lie group G. We know that φ is a homomorphism with

$$\mathbf{ker}(\varphi) = \pi_1(G).$$

We infer that $\pi_1(G) = \varphi^{-1}(e)$ is a closed set with no accumulation points by the fact that φ is evenly covered, i.e., $\varphi^{-1}(U)$ is a union of disjoint copies homeomorphic to U where U is a small neighborhood of e the identity. So, $\pi_1(G) = \varphi^{-1}(e)$ is discrete. Now, we shall show that any topological discrete normal subgroup of a connected topological group is abelian: If D is topological discrete normal subgroup of a connected topological group G, then for $g \in G$ and $\delta \in D$

$$\delta \mapsto g\delta g^{-1}$$

gives a well-defined continuous action of G on D by the fact D is normal. Since D is discrete, it must be a constant map by the continuity and discreteness. So, we conclude that

$$g\delta g^{-1} = e\delta e^{-1} = \delta, \text{ which means } D \in \mathbf{Center}(G).$$

So, D is central. Hence, D is abelian. Since $\pi_1(G)$ is a discrete normal Lie subgroup of a connected Lie group G, it is abelian. □

In order to show that $\pi_1(G)$ is finitely generated, we need one of deep theorems in Lie group theory. Without a proof, we state the following

Theorem 3.1.7.3. *Let* G *be a connected Lie group. Then there exists a compact Lie group* $K \subseteq G$ *such that* G *is diffeomorphic to* $K \times \mathbb{R}^d$ *for some* $d \in \mathbb{Z}$.

By Theorem 3.1.7.3 and the fact that $\pi_1(K)$ is finitely generated if K is compact, noting

$$\pi_1(A \times B) = \pi_1(A) \times \pi_1(B),$$

we conclude that $\pi_1(G)$ is finitely generated.

Example 3.1.7.2. *If* $G = \mathbf{PSL}(2, \mathbb{C})$*, then*

$$G \cong \mathbb{RP}^3 \times \mathbb{R}^3 \cong \mathbf{PU}(2) \times \mathbb{R}^3 \cong \mathbf{SO}(3) \times \mathbb{R}^3.$$

If $G = \mathbf{PSL}(2, \mathbb{R})$*, then*

$$G \cong S^1 \times \mathbb{R}^2 \cong \mathbf{PSO}(2) \times \mathbb{R}^2.$$

Suppose that $G \xrightarrow{\varphi} H$ be a smooth map between Lie groups such that

$$d\varphi : T_eG \to T_eH$$

is a Lie algebra isomorphism. So, by the inverse function theorem, there exists a neighborhood U of $e \in G$ such that $\varphi|_U : U \to \varphi(U)$ is a diffeomorphism. Using the left-invariant vector fields of G, we actually have $d\varphi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra isomorphism. So, by the same arguments, we deduce that $G \xrightarrow{\varphi} H$ is a local diffeomorphism. From this, note that

$$\dim G = \dim H.$$

Note that Definition1.1.8.1 say that $G \xrightarrow{f} H$ is called a **local Lie group isomorphism** if f is a homomorphism and f is a local diffeomorphism. From the above, we also have an equivalent definition of a local Lie group isomorphism. That is, a smooth map $G \xrightarrow{f} H$ is called a **local Lie group isomorphism** if f is a homomorphism and $d\varphi : T_eG \to T_eH$ is a Lie algebra isomorphism.

Theorem 3.1.7.4. *Let G and H be connected Lie groups with the Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Suppose that $\varphi : G \to H$ is a Lie group homomorphism. Then φ is a covering map if and only if $(d\varphi)_e : \mathfrak{g} \to \mathfrak{h}$ is an isomorphism.*

Proof. Suppose that φ is a covering space and $(d\varphi)_e$ is not injective. Let

$$\mathfrak{a} = \mathbf{ker}((d\varphi)_e).$$

Since $\mathfrak{a}$ is a nontrivial Lie subalgebra of $\mathfrak{g}$, by Theorem 3.1.6.4, we have a nontrivial connected Lie subgroup A containing e of G. However, since $\varphi : G \to H$ is a Lie group homomorphism, we must have $\varphi(A) = e$, which contradicts the assumption that φ is a covering space, i.e., $\varphi^{-1}(e)$ is discrete. Now, suppose that $(d\varphi)_e$ is not surjective. Let

$$\mathfrak{b} = \mathbf{Im}((d\varphi)_e).$$

Since $\mathfrak{b}$ is a proper Lie subalgebra of $\mathfrak{h}$, by Theorem 3.1.6.4, we have a proper connected Lie subgroup B containing e of H where $\varphi : G \to B \subset H$ is onto around neighborhoods of identities $e_G \in G$ and $e_B \in B$ by the fact that φ is a homomorphism. However, since B is a proper submanifold of H, this contradicts the assumption that φ is a covering space. Hence, $(d\varphi)_e$ is a Lie algebra isomorphism.

Suppose that $(d\varphi)_e : \mathfrak{g} \to \mathfrak{h}$ is an isomorphism. By the previous argument, φ is a local diffeomorphism. By Exercise 3.1.7.1 and $\varphi(e_G) = e_H$ from the assumption that φ is a homomorphism, we deduce that H is generated by $\varphi(U)$ where U is a small neighborhood of $e \in G$. So, we conclude that φ is onto. Now, in order to show that φ is a covering space, the only remaining thing to show is that G is evenly covered. That is, we have to construct a neighborhood V_h in H for each $h \in H$ such that each component of $\varphi^{-1}(V_h)$ is homeomorphic to V_h. However, using left-multiplications in H, which are diffeomorphisms of H, it is sufficient to construct a neighborhood of e_H satisfying the required property. Since φ is a local diffeomorphism and $D = \varphi^{-1}(e_H)$ is discrete, we can find a neighborhood V of e_G in G such that $\varphi|_V$ is a diffeomorphism and $V \cap D = \{e_G\}$. From this, $\varphi|_V : V \to G$ is an isomorphism. Now, suppose that there exists $a \in VV^{-1} \cap D$. So, letting $a = v_1 v_2^{-1}$ for some $v_1, v_2 \in V$, we have

$$\varphi(v_2) = \varphi(av_2) = \varphi(v_1).$$

Since $\varphi|_V : V \to G$ is an isomorphism, we have $v_1 = v_2$. Hence, $a = e_G$. That is,

$$VV^{-1} \cap D = \{e_G\}.$$

So, the following claim will complete the proof. The claim is

$$\varphi^{-1}(\varphi(V)) = \coprod_{\delta \in D} \delta \cdot V.$$

Since $\varphi(\bigcup_{\delta \in D} \delta \cdot V) = \bigcup_{\delta \in D} \varphi(\delta) \cdot \varphi(V) = \varphi(V)$, we have

$$\bigcup_{\delta \in D} \delta \cdot V \subseteq \varphi^{-1}(\varphi(V)).$$

Suppose that $\sigma \in \varphi^{-1}(\varphi(V))$. So, $\varphi(\sigma) \in \varphi(V)$. So, there exists $g \in V$ such that $\varphi(\sigma) = \varphi(g)$. Since φ is a homomorphism, we deduce that

$$\sigma g^{-1} \in D = \varphi^{-1}(e_H) \Longleftrightarrow \sigma \in D \cdot g.$$

So, $\sigma \in \bigcup_{\delta \in D} \delta \cdot V$, which shows that

$$\bigcup_{\delta \in D} \delta \cdot V \supseteq \varphi^{-1}(\varphi(V)).$$

Now, we shall show that $\delta_1 \cdot V \cap \delta_2 \cdot V = \emptyset$ for $\delta_1, \delta_2 \in D$ with $\delta_1 \neq \delta_2$. Let $g \in \delta_1 \cdot V \cap \delta_2 \cdot V$. So,

$$g = \delta_1 \cdot g_1 = \delta_2 \cdot g_2 \text{ where } g_1, g_2 \in G.$$

Hence,

$$\delta_2^{-1} \cdot \delta_1 = g_2 \cdot g_1^{-1} \in VV^{-1}.$$

Since $D \cap VV^{-1} = \{e_G\}$, we have $\delta_2^{-1} \cdot \delta_1 = e_G$, which is a contradiction. So, we have $\delta_1 \cdot V$ is disjoint from $\delta_2 \cdot V$ for $\delta_1, \delta_2 \in D$ with $\delta_1 \neq \delta_2$.

$$\varphi^{-1}(\varphi(V)) = \bigcup_{\delta \in D} \delta \cdot V = \coprod_{\delta \in D} \delta \cdot V.$$

For an arbitrary $h \in H$, using the facts left-multiplications in H are diffeomorphism and φ is onto, letting $g \in \varphi^{-1}(h)$, we have

$$\varphi^{-1}(h \cdot \varphi(V)) = \bigcup_{\delta \in D} \delta \cdot V = \coprod_{\delta \in D} \delta \cdot (g \cdot V).$$

□

Exercise 3.1.7.1. *Let G be a connected topological group and U be an open set of e. Then U generates G. That is,*

$$G = \bigcup_{n=1}^{\infty} U^n.$$

Proof. Letting $V = U \cap U^{-1}$, it is easy to see that

$$\bigcup_{n=1}^{\infty} V^n \text{ is an open nonempty subgroup of } G.$$

Our claim is that this is also closed. Let $\{g_k\} \in \bigcup_{n=1}^{\infty} V^n$ be a sequence. So, it is easy to see that $g_{k+1}g_k^{-1} \in \bigcup_{n=1}^{\infty} V^n$. Since adding $g_0 = e$,

$$\lim_{k\to\infty} g_k = \prod_{k=0}^{\infty} g_{k+1}g_k^{-1} \in \bigcup_{n=1}^{\infty} V^n \text{ by definiton,}$$

we conclude that

$$\bigcup_{n=1}^{\infty} V^n \text{ is closed.}$$

So, since G is connected, we conclude that

$$G = \bigcup_{n=1}^{\infty} V^n \subseteq \bigcup_{n=1}^{\infty} U^n \subseteq G.$$

□

Now, we can give a proof of the affirmative answer to Question 3.1.6.1.

Theorem 3.1.7.5. *Let G and H be a Lie groups with the Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively.* ***Assume*** *$\pi_1(G) = 1$. Then every Lie algebra homomorphism $\mathfrak{g} \xrightarrow{\Phi} \mathfrak{h}$ arises from a unique Lie group homomorphism $G \xrightarrow{\varphi} H$.*

Proof. First, we note that

$$\mathfrak{g} \oplus \mathfrak{h} = \mathfrak{X}(G \times H).$$

Define

$$\mathsf{graph}(\Phi) = \Phi' : \mathfrak{g} \to \mathfrak{g} \oplus \mathfrak{h} \text{ by } \Phi'(X) = (X, \Phi(X)).$$

It is easy to see that $\mathbf{Im}(\Phi')$ is a Lie subalgebra of $\mathfrak{g} \oplus \mathfrak{h}$ and it is Lie algebra isomorphic to $\mathfrak{g}$. By Theorem 3.1.6.4, there exists a unique connected Lie subgroup A of $G \times H$ with the Lie algebra $\mathfrak{a} = \mathfrak{X}(A)$ isomorphic to $\mathbf{Im}(\Phi')$. Let π_G and π_H be the canonical projections from $G \times H$ to G and H, respectively. Clearly. $\pi_G|_A : A \to G$ induce a Lie algebra isomorphism $d\pi_G : \mathfrak{a} \to \mathfrak{g}$. Now, Theorem 3.1.7.4 tells us that $\pi_G|_A$ is a covering space. Since G is simply connected, $\pi_G|_A$ is in fact a diffeomorphism, so a Lie group isomorphism. Define

$$\varphi : G \to H \text{ by } \varphi = \pi_H \circ (\pi_G|_A)^{-1}.$$

Obviously, $d\pi_H \circ d(\pi_G|_A)^{-1} = \Phi$. Also, note that we already know it is uniqueness if it exists. □

Remark 3.1.7.1. *Combining Theorem 3.1.7.5 with Ado's theorem, which state that every Lie algebra $\mathfrak{g}$ admits a faithful representation of $\mathfrak{gl}(n, \mathbb{R})$, we deduce that every Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R}) = \mathbf{M}_n(\mathbb{R})$ gives a Lie group G. Note that G and the universal covering space $\widetilde{G}$ give the same Lie algebra.*

3.1.8 The Exponential map again

Note that in Exercise 2.1.8.4, we showed that every left-invariant vector field of a connected Lie group G is complete. Let G be a connected Lie group with the Lie algebra $\mathfrak{g}$. Suppose that $X \in \mathfrak{g}$. Then there exists the unique global flow ξ_t corresponding X. Define

$$\exp : \mathfrak{g} \to G \text{ by } \exp(tX) = \xi_t(e) \text{ for } t \in \mathbb{R}.$$

Note that $\exp(X) = \xi_1(e)$ and $\exp(0) = e$.

Theorem 3.1.8.1. *Let $\varphi : G \to H$ be a homomorphism of Lie groups. Then* exp *is equivariant: That is, the following diagram commutes*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\exp} & G \\ {\scriptstyle d\varphi}\downarrow & & \downarrow{\scriptstyle \varphi} \\ \mathfrak{h} & \xrightarrow[\exp]{} & H \end{array}$$

Proof. This comes from the uniqueness of flows and the fact that $d\varphi$ induces an action on a flow by φ. □

It is easy to see that exp satisfies the following properties:
1. exp is well-defined and smooth.
2. By identifying $\mathfrak{g}$ with T_eG, we have

$$\exp(\mathbf{Ad}(g)X_e) = g\exp(X_e)g^{-1}.$$

That is, it is equivariant with respect to $G \xrightarrow{\mathbf{Ad}} \mathbf{Aut}(\mathfrak{g})$ given by $\mathbf{Ad}(g)X_e = (\iota_g)_*(X_e)$ where $G \xrightarrow{\iota} \mathbf{Aut}(G)$ by $\iota_g(h) = ghg^{-1}$:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\exp} & G \\ {\scriptstyle (\iota_g)_* = \mathbf{Ad}(g)}\downarrow & & \downarrow{\scriptstyle \iota_g} \\ \mathfrak{g} & \xrightarrow[\exp]{} & G \end{array}$$

Proof. This comes from Theorem 3.1.8.1 and the fact that $\mathbf{Ad}(g)$ acts on a flow by conjugation from the definition of $\mathbf{Ad}$. □

3. $(d\exp)_0 : \mathfrak{g} = T_0\mathfrak{g} \to T_eG = \mathfrak{g}$ is the identity map where 0 is the identity of $\mathfrak{g}$.

Proof.

$$\frac{d}{dt}|_{t=0}\exp(tX) = X \cdot \exp(tX)|_{t=0} = X.$$

□

4. The map exp gives a local diffeomorphism from a neighborhood of 0 in $\mathfrak{g}$ to a neighborhood of e in G.

Proof. By 3 and the inverse function theorem, there exists a neighborhood U_0 of $0 \in \mathfrak{g}$ such that

$$\exp : U \to \exp(U) \subseteq G, \text{ which is a neighborhood of } e$$

is a diffeomorphism. Note that $\exp(U)$ is called a local Lie group or group chunk. Note that since every Lie algebra is contractible by the fact that they are vector spaces, we conclude that if $\exp$ is a global diffeomorphism, then the Lie group must be contractible. So, non contractibility of a Lie group is one of obstructions for $\exp$ to be a global diffeomorphism. ∎

5. $\exp((t_1 + t_2)X) = \exp(t_1 X)\exp(t_2 X)$ for $t_1, t_2 \in \mathbb{R}$.

Proof. Note that in Exercise 2.1.8.5, we showed that the flow of a left-invariant vector field on a Lie group G is a right multiplication. Keeping this in your mind, we have

$$\exp((t_1 + t_2)X) = \xi_{t_1+t_2}(e) = \xi_{t_2}(\xi_{t_1}(e)) = \xi_{t_1}(e) \cdot \xi_{t_2}(e) = \exp(t_1 X)\exp(t_2 X).$$

Note that this proof also show that

$$\exp(t_1 X)\exp(t_2 X) = \exp(t_2 X)\exp(t_1 X).$$

∎

6. $\exp(-X) = (\exp(X))^{-1}$.

Proof. By the above, we have

$$e = \exp(0) = \exp(X - X) = \exp(X)\exp(-X).$$

So, $\exp(-X) = (\exp(X))^{-1}$. ∎

7. For $X, Y \in \mathfrak{g}$, if $[X, Y] = 0$, then we have

$$\exp(X + Y) = \exp(X)\exp(Y).$$

Proof. Let ξ_t and η_t be the flows corresponding to X and Y, respectively. By Theorem 2.1.6.2, we know that $[X, Y] = 0$ implies that ξ_t and η_t commute. Moreover, Exercise 2.1.8.5 showed that the flow of a left-invariant vector field on a Lie group G is in fact given by a right multiplication. So, the upshot is that $\xi_t(e) \cdot \eta_t(e)$ becomes a flow if $[X, Y] = 0$ and G is a **Lie group**: Let $\Phi_t(e) = \xi_t(e) \cdot \eta_t(e) = \eta_t(\xi_t(e))$. We have to show that $\Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2}$.

$$\begin{aligned}
\Phi_{t_1+t_2}(e) &= \xi_{t_1+t_2}(e) \cdot \eta_{t_1+t_2}(e) = \xi_{t_2}(\xi_{t_1}(e)) \cdot \eta_{t_2}(\eta_{t_1}(e)) = \xi_{t_1}(e) \cdot \xi_{t_2}(e) \cdot \eta_{t_1}(e) \cdot \eta_{t_2}(e) \\
&= \xi_{t_1}(e) \cdot \eta_{t_1}(e) \cdot \xi_{t_2}(e) \cdot \eta_{t_2}(e) = \eta_{t_2}(\xi_{t_1}(e) \cdot \eta_{t_1}(e) \cdot \xi_{t_2}(e)) = \eta_{t_2}(\xi_{t_2}(\xi_{t_1}(e) \cdot \eta_{t_1}(e))) \\
&= \Phi_{t_2}(\Phi_{t_1}(e))
\end{aligned}$$

To illustrate the content of Exercise 2.1.8.5, we give another proof:

$$\begin{aligned}
\Phi_{t_1+t_2}(e) &= \eta_{t_1+t_2}(\xi_{t_1+t_2}(e)) = \eta_{t_1}(\eta_{t_2}(\xi_{t_1+t_2}(e))) = \eta_{t_1}(\eta_{t_2}(\xi_{t_2}(\xi_{t_1}(e)))) \\
&= \eta_{t_1}(\xi_{t_1}(e) \cdot \eta_{t_2}(\xi_{t_2}(e))) = \eta_{t_1}(\eta_{t_2}(\xi_{t_2}(e)) \cdot \xi_{t_1}(e)) = \eta_{t_1}(\xi_{t_1}(\eta_{t_2}(\xi_{t_2}(e)))) \\
&= \Phi_{t_1}(\Phi_{t_2}(e))
\end{aligned}$$

Now, by Exercise 3.1.6.4, we know that the multiplications of a Lie group give the additions on $T_eG = \mathfrak{g}$. So, $\Phi_t(e) = \xi_t(e) \cdot \eta_t(e)$ gives $X_e + Y_e$. That is, the flow Φ_t gives a vector field $X + Y$. Hence,

$$\exp(X + Y) = \Phi_1(e) = \xi_1(e) \cdot \eta_1(e) = \exp(X)\exp(Y).$$

Actually, in this proof we showed that if $[X, Y] = 0$, then

$$\exp(X)\exp(Y) = \exp(Y)\exp(X).$$

□

Remark 3.1.8.1. *In a Lie group, since a flow satisfies $\xi_t(g) = g\xi_t(e)$, we deduce that*

$$g\exp(tX) \text{ is nothing but the flow of } X.$$

Theorem 3.1.8.2. *If $G = \mathbf{GL}(n, \mathbb{R})$, then we have*

$$\exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k \text{ for } X \in \mathfrak{gl}(n, \mathbb{R}).$$

Note that it is still true that if we replace $\mathbb{R}$ to $\mathbb{C}$.

Proof. Note that in Subsection 1.2.1, we showed that $\sum_{k=0}^{\infty} \frac{1}{k!} X^k$ converges absolutely. Since we know that

$$\exp(tX) \in \mathbf{GL}(n, \mathbb{R}),$$

let

$$f(t) = \exp(tX) = \sum_{k=0}^{\infty} A_k t^k \text{ where } A_k \in \mathbf{GL}(n, \mathbb{R}) \text{ and } A_0 = I \text{ the identity.}$$

Note that we have $f(0) = I$ and $f'(t) = Xf(t)$. So,

$$Xf(t) = \sum_{k=0}^{\infty} XA_k t^k = \sum_{k=0}^{\infty} (k+1)A_{k+1} t^k = f'(t).$$

Hence, we have

$$A_{k+1} = \frac{1}{k+1} XA_k \text{ with } A_0 = I.$$

Therefore,

$$A_k = \frac{1}{k!} X^n.$$

□

Remark 3.1.8.2. *We know that* exp *is a local diffeomorphism of a neighborhood U_0 of $0 \in \mathfrak{gl}(n, \mathbb{R})$. So,* exp *has the inverse locally, which will be denoted by* log*. In the case of a Lie group $G \subseteq \mathbf{GL}(n, \mathbb{R})$, we have an explicit formula, of course, locally:*

$$\log(A) = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} A^k.$$

$$\begin{array}{ccc} U_0 & \xleftarrow[\log]{} & \exp(U_0) \\ \downarrow & & \downarrow \\ \mathfrak{g} & \xrightarrow{\exp} & G \end{array}$$

Example 3.1.8.1. *Note that in Example 1.1.4.1 and Exercise 1.1.4.1, we showed the following: Let*

$$\mathfrak{sl}(2,\mathbb{R}) \cong \mathfrak{su}(2) \cong \mathbb{H}_0 = \{q \mid q = ai + bj + ck \text{ where } a,b,c \in \mathbb{R}\}$$

be the set of traceless quaternions, which is the Lie algebra of a Lie group $\mathbb{H}_1 \cong \mathbf{SU}(2) \cong S^3$*. We have*

$$\exp : \mathbb{H}_0 \to \mathbb{H}_1.$$

We showed that $\exp(ti) = \sum_{k=0}^{\infty} \frac{t^k}{k!} i^k = \cos(t) + \sin(t)i$ *and for* $q \in \mathbb{H}_0$ *and* $t \in \mathbb{R}$*,*

$$\exp(tq) = \cos(t) + \sin(t)q.$$

In general, since $\mathbb{H}$ *is the Lie algebra of a Lie group* $\mathbb{H}^\times$*,* $\exp : \mathbb{H} \to \mathbb{H}^\times$ *is given by*

$$\exp(tq) = \cos(t\|q\|) + \frac{\sin(t\|q\|)}{\|q\|} q.$$

So, we have for $q = \cos(\theta)j + \sin(\theta)k \in \mathbb{H}$*,*

$$\exp(2\pi q) = 1 \text{ and } \exp(\pi q) = -1.$$

Hence, even if θ *varies, we still have* $\exp(\pi q) = -1$*. That is, identifying* $T_{\pi q}\mathbb{H} = \mathbb{H}$ *and* $T_{\pi q}\mathbb{H}^\times = \mathbb{H}$*, we conclude that*

$$d(\exp)_{\pi q} : \mathbb{H} \to \mathbb{H}$$

is singular, i.e., actually $d(\exp)_{\pi q} = 0$*. We infer that the one-parameter subgroup* $\exp(\pi q) \in \mathbb{H}_1 \subseteq \mathbb{H}^\times$ *is geodesic.*

Example 3.1.8.2. exp *is not necessarily surjective.*

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbf{SL}(2,\mathbb{R}) \text{ but } \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \notin \exp(\mathfrak{sl}(2,\mathbb{R})).$$

Theorem 3.1.8.3 (Baker-Campbell-Hausdorff Formula). *Let* $X, Y \in \mathfrak{g}$*. Suppose that* $\exp(sX)$ *and* $\exp(tY)$ *are sufficiently close to* $e \in G$*. Then we have*

$$\log(\exp(sX)\exp(tY)) = C(sX, tY) \text{ where}$$

$C(sX,tY) = X + Y + \frac{1}{2}[sX, tY] + \frac{1}{12}([sX,[sX,tY]] + [tY,[tY,sX]]) + \cdots$.

Proof. See "Lie groups, Lie algebras, and their representations" by V.S. Varadarajan. □

Note that the exact formula of $C(sX, tY)$, which you can find in the above book is given by a power series in $sX, tY, [sX, tY]$. If $\mathfrak{g}$ in nilpotent, obviously $C(sX, tY)$ is given by a polynomial in $sX, tY, [sX, tY]$. Now, we prove the followings:

Theorem 3.1.8.4. *Let H be a connected Lie subgroup of a connected Lie group G. Then H is a normal subgroup if and only if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.*

In order to prove this, we need a lemma. In Remark 3.1.1.1, we stated the followings without proof:

$$\begin{array}{ccc} G & \xrightarrow{\mathbf{Ad}} & \mathbf{Aut}(\mathfrak{g}) \\ {\scriptstyle \exp}\Big\uparrow & & {\scriptstyle \exp}\Big\uparrow \\ \mathfrak{g} & \xrightarrow{(\mathbf{Ad})_*=\mathbf{ad}} & \mathbf{End}(\mathfrak{g}). \end{array}$$

Since $\mathbf{Aut}(\mathfrak{g}) = \mathbf{Aut}(T_eG) \cong \mathbf{GL}(n,\mathbb{R})$, it is a Lie group. Moreover, $\mathbf{End}(\mathfrak{g}) \cong T_{id}\mathbf{GL}(n,\mathbb{R}) = \mathfrak{gl}(n,\mathbb{R})$ is the corresponding Lie algebra. So, the commutativity of the diagram shall come from Theorem 3.1.8.1 and the fact that $(\mathbf{Ad})_* = \mathbf{ad}$, which we shall prove in the next lemma.

Lemma 3.1.8.1.

$$\mathbf{Ad}(\exp(tX)) = \exp(\mathbf{ad}(tX)).$$

Proof. We have to show that $((\mathbf{Ad})_*(X))(Y) = (\mathbf{ad}(X))(Y) \overset{def}{=} [X,Y]$ for all $X, Y \in \mathfrak{g}$. Let ξ_t and η_s be the corresponding flow of X and Y. Reminding that the flow of a left-invariant vector field on a Lie group is a right multiplication,

$$\begin{aligned} ((\mathbf{Ad})_*(X))(Y) &= \frac{d}{dt}|_{t=0}(\mathbf{Ad}(\xi_t(e))(Y)) = \frac{d}{dt}|_{t=0}(r_{\xi_{-t}(e)})_*(Y) \\ &= \frac{d}{dt}|_{t=0}(dr_{\xi_{-t}(e)})(Y) = \frac{d}{dt}|_{t=0}\frac{d}{ds}|_{s=0}(\eta_s(\xi_t(e))\xi_{-t}(e)) = \frac{d}{dt}|_{t=0}\frac{d}{ds}|_{s=0}(\xi_{-t}\circ\eta_s(\xi_t(e)) \\ &= \frac{d}{dt}|_{t=0}(d\xi_{-t})(Y_{\xi_t(e)}) = \frac{d}{dt}|_{t=0}(\xi_{-t}\cdot Y) = \mathcal{L}_X(Y) = [X,Y] = (\mathbf{ad}(X))(Y). \end{aligned}$$

□

Actually, since $(\mathbf{Ad})_* = \mathbf{ad}$ implies that $\mathbf{ad}$ is linear, Lemma 3.1.8.1 can be written

$$\mathbf{Ad}(\exp(tX)) = \exp(t\mathbf{ad}(X)).$$

Also, note that in the proof since $\exp(tX) = \xi_t$ and $\exp(sY) = \eta_s$, we infer that

$$(\mathbf{Ad}(\exp(tX_e)))(sY_e) = \frac{d}{ds}|_{s=0}\exp(tX_e)\exp(sY_s)\exp(-tX_e).$$

From this, if we go one further step, we can deduce a simplest form:

$$\exp\big(\mathbf{Ad}(\exp(tX))(sY)\big) = \exp(tX)\exp(sY)\exp(-tX).$$

Definition 3.1.8.1 (Ideals of a Lie algebra). *A Lie subalgebra $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is called an (left) ideal if $\mathbf{ad}(Y)(X) = [Y,X] \in \mathfrak{h}$ for all $Y \in \mathfrak{g}$ and $X \in \mathfrak{h}$. Equivalently, $\mathfrak{h}$ is stable under $\mathbf{ad}(\mathfrak{g})$.*

Proof of Theorem 3.1.8.4. Suppose that H is normal in G. Let $X \in \mathfrak{h}$ and $Y \in \mathfrak{g}$. Since we know that $\exp(sX)$ and $\exp(tY)$ are one-parameter subgroups of X and Y, respectively, using the assumption that H is normal, we have

$$\exp\big(\mathbf{Ad}(\exp(tY))(sX)\big) = \exp(tY)\exp(sX)\exp(-tY) \in H.$$

So, we deduce that $\mathbf{Ad}(\exp(tY))(sX) \in \mathfrak{h}$. By Lemma 3.1.8.1 we know that

$$\begin{aligned}\mathbf{Ad}(\exp(tY))(sX) = \exp(\mathbf{ad}(tY))(sX) = \exp(t\mathbf{ad}(Y))(sX) &= (\sum_{n=0}^{\infty}\frac{t^n}{n!}\mathbf{ad}(Y)^n)(sX)\\ &= sX + ts[Y,X] + s\frac{t^2}{2}[Y,[Y,X]] + \cdots .\end{aligned}$$

Since $\mathfrak{h} \cong T_eH$ is a vector space, $\mathfrak{h}$ is a Lie group with the Lie algebra $\mathfrak{h} \cong T_e\mathfrak{h}$. So, if we think $\mathbf{Ad}(\exp(tY))(sX)$ is a smooth path in $\mathfrak{h}$ with respect to t, the tangent vector at the identity e on $\mathfrak{h}$ is nothing but

$$s[Y,X] \in T_e\mathfrak{h} = \mathfrak{h} \text{ by the power series expansion.}$$

So, we conclude that $[Y,X] \in \mathfrak{h}$, which shows that $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. Conversely, suppose that $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. It is easy to see that H is generated by the set of flows $\exp(\mathfrak{h})$. Now we have to show that

$$g(\exp(\mathfrak{h}))g^{-1} \subseteq \exp(\mathfrak{h}).$$

Actually, it suffices to show that this case when $g \in \exp(\mathfrak{g})$. That is, for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$ and $Y \in \mathfrak{g}$, we have to show that

$$\exp\big(\mathbf{Ad}(\exp(tX))(sY)\big) = \exp(tX)\exp(sY)\exp(-tX) \in H.$$

By the above we know that

$$\exp(tX)\exp(sY)\exp(-tX) = \exp(sY + ts[X,Y] + s\frac{t^2}{2}[X,[X,Y]] + \cdots).$$

Since $\mathfrak{h}$ is an ideal, we deduce that $sY + ts[X,Y] + s\frac{t^2}{2}[X,[X,Y]] + \cdots \in \mathfrak{h}$. Hence,

$$\exp(tX)\exp(sY)\exp(-tX) \in H.$$

□

Question 3.1.8.1. *When does* $\mathfrak{g}$ *admit a nondegenerate symmetric bilinear form invariant under* $\mathbf{Ad}$*?*

From Subsection 1.2.2, we know that

$$\{\text{(nondegenerate) positive definite bilinear forms on } \mathfrak{g} = T_1(G)\}$$
$$\Updownarrow$$
$$\{\text{left-invariant (psedo-) Reimannian metrics on } G\}$$

Since $\mathbf{Ad} : G \to \mathbf{Aut}(\mathfrak{g})$ acts on $\mathfrak{g}$ by $(r_{g^{-1}})_* = dr_{g^{-1}}$, it is obvious that

$$\{\text{(nondegenerate) positive definite bilinear forms on } \mathfrak{g} = T_1(G) \text{ which are invariant under } \mathbf{Ad}\}$$
$$\Updownarrow$$
$$\{\text{bi-invariant (psedo-) Reimannian metrics on } G\}$$

Example 3.1.8.3. *Suppose that*

$$\rho : G \to \mathbf{GL}(n, \mathbb{R}) \text{ by } \rho(g_1 \cdot g_2) = \rho(g_1) \cdot \rho(g_2)$$

be a representation of a Lie group G. We know that ρ induces

$$d\rho : \mathfrak{g} \to \mathfrak{gl}(n, \mathbb{R}) = \mathbf{M}_n(\mathbb{R}), \text{ which is also a representation.}$$

Note that

$$d\rho(X + Y) = d\rho(X) + d\rho(Y).$$

Define a bilinear form $\mathcal{B}$ on $\mathfrak{g}$ by

$$\mathcal{B}(X, Y) = \mathbf{tr}((d\rho(X)) \cdot (d\rho(Y)))$$

By Exercise 3.1.8.1, $\mathcal{B}$ is nondegenerate. Now let $g \in G$. Using the fact that $\mathbf{Ad}(g) = (r_{g^{-1}})_* = (\iota_g)_*$ on a left-invariant vector field, we have

$$\begin{aligned}
\mathcal{B}(\mathbf{Ad}(g)X, \mathbf{Ad}(g)Y) &= \mathbf{tr}((d\rho(\mathbf{Ad}(g)X)) \cdot (d\rho(\mathbf{Ad}(g)Y))) \\
&= \mathbf{tr}((d\rho(d\iota_g X)) \cdot (d\rho(d\iota_g Y))) \\
&= \mathbf{tr}((d(\rho \circ \iota_g)X) \cdot (d(\rho \circ \iota_g)Y)) \\
&= \mathbf{tr}(\frac{d}{dt}|_{t=0}\rho(g \exp(tX) g^{-1}) \cdot \frac{d}{dt}|_{t=0}\rho(g \exp(tY) g^{-1})) \\
&= \mathbf{tr}(\rho(g)(\frac{d}{dt}|_{t=0}\rho(\exp(tX))) \cdot (\frac{d}{dt}|_{t=0}\rho(\exp(tY)))\rho(g)^{-1}) \\
&= \mathbf{tr}(\frac{d}{dt}|_{t=0}\rho(\exp(tX)) \cdot \frac{d}{dt}|_{t=0}\rho(\exp(tY))) \\
&= \mathbf{tr}(d\rho(X) \cdot d\rho(Y)) = \mathcal{B}(X, Y).
\end{aligned}$$

That is, $\mathfrak{B}$ is a bi-invariant (pseudo-) Riemannian metric on G.

Exercise 3.1.8.1. *Show that $\mathcal{B}$ is nondegenerate on $\mathfrak{g}$.*

Proof. It suffices to show that if $\mathbf{tr}(A \cdot B) = 0$ for all $A \in \mathbf{M}_n(\mathbb{R})$, then $B = 0$. Equivalently, it suffices to show that if $B \neq 0$, the there exists $A \in \mathbf{M}_n(\mathbb{R})$ such that $\mathbf{tr}(A \cdot B) \neq 0$. Let $B = (b_{ij})_{(n \times n)} \neq 0$. So, there is some $b_{\alpha\beta} \neq 0$. Let

$$A = (a_{ij})_{(n \times n)} \text{ such that } a_{\beta\alpha} = \frac{1}{b_{\alpha\beta}} \text{ and the other entries are all zeros.}$$

It is easy to see that $A \cdot B$ is a matrix with one diagonal entry is 1 and the other entries are all zeros. So, $\mathbf{tr}(A \cdot B) = 1$. ☐

Since $\mathbf{Ad} : G \to \mathbf{GL}(n, \mathbb{R})$ is naturally a representation of a Lie group G, which is called the **adjoint representation**, we deduce that

$$(\mathbf{Ad})_* = \mathbf{ad} : \mathfrak{g} \to \mathbf{End}(\mathfrak{g}) \cong \mathbf{Der}(\mathfrak{g})$$

is also a representation. Now, we define an inner product on $\mathfrak{g}$ by the following way:

$$\mathbf{B}(X, Y) = \mathbf{tr}((\mathbf{Ad})_*(X) \cdot (\mathbf{Ad})_*(Y)) = \mathbf{tr}((\mathbf{ad})(X) \cdot (\mathbf{ad})(Y)).$$

Definition 3.1.8.2 (The Killing form).

$$\mathbf{B}(X,Y) = \mathbf{tr}((\mathbf{ad})(X)\cdot(\mathbf{ad})(Y))$$

is called the killing from on G.

It is important to note that if G is abelian, then since

$$\mathbf{Ad}(g)(X) = \frac{d}{dt}|_{t=0}(g\exp(tX)g^{-1}) = \frac{d}{dt}|_{t=0}(\exp(tX)) = X,$$

we deduce that

$$(\mathbf{Ad})_* = \mathbf{ad} = 0.$$

Hence, $\mathbf{B}(X,Y) = \mathbf{tr}((\mathbf{ad})(X)\cdot(\mathbf{ad})(Y)) = 0$. That is, $\mathbf{B}$ is degenerate. On the other hand, E. Cartan shows that

Theorem 3.1.8.5 (E. Cartan). $\mathfrak{g}$ *has nontrivial solvable ideals (, i.e., semisimple) if and only if the killing form* $\mathbf{B}$ *is nondegenerate.*

Proof. We will prove this later. □

Theorem 3.1.8.6. *If G is a compact connected Lie group, then G admits a bi-invariant Riemannian metric.*

Proof. In Example 3.1.1.3, we show that G is unimodular if and only if there exists a bi-invariant volume form on G. Suppose that G is compact and connected. Since the unimodular function $\det\circ\mathbf{Ad}$;

$$G \xrightarrow{\mathbf{Ad}} \mathbf{Aut}(\mathfrak{g}) \xrightarrow{\det} \mathbb{R}^{\times},$$

is a Lie group homomorphism, it is easy to see that $\mathbf{Im}(\det\circ\mathbf{Ad}(G))$ must be a compact and connected subgroup of $\mathbb{R}^{\times}$. However, the only compact and connected subgroup of $\mathbb{R}^{\times}$ is $\{+1\}$. Hence, we conclude that

$$\det\circ\mathbf{Ad} \equiv 1.$$

So, G has a bi-invariant volume from. Now, using exterior algebra of differentiable forms, it is not hard to see that there exists a bi-invariant Riemannian metric, which shall induce the volume form. □

Without a proof, we give the converse.

Theorem 3.1.8.7. *If $\mathbf{Ad}(G)$ preserves a positive definite inner product on $\mathfrak{g}$, i.e.,*

$$\mathbf{Ad} : G \to \mathbf{O}(n),$$

then

$$G \cong \text{ (abelian Lie group)}\times\text{(compact Lie group)}.$$

Exercise 3.1.8.2. *Show that*

$$\mathbf{Isom}_+(\mathbb{R}^3) = \mathbf{SO}(3) \ltimes \mathbb{R}^3 \text{ and } \mathbf{Isom}_+(\mathbb{R}^{2,1}) = \mathbf{SO}(2,1) \ltimes \mathbb{R}^3$$

have bi-invariant pseudo-Riemannian metrics. Note that $\mathbb{R}^3$ *is a normal subgroup of* $\mathbf{Isom}_+(\mathbb{R}^3)$ *and* $\mathbf{Isom}_+(\mathbb{R}^{2,1})$.

Also, show that the Heisenberg group $\mathbf{H}_3$ *does not have a bi-invariant pseudo-Riemannian metric.*

Theorem 3.1.8.8 (Warner p. 97). *Every closed subgroup* H *of a Lie group* G *is a Lie subgroup.*

Theorem 3.1.8.9 (Warner p.109). *A continuous homomorphism of Lie groups is a Lie group homomorphism.*

Chapter 4

General Theory of Lie Algebras

4.1 Overviews

4.1.1 Recalls and Preliminaries

Consider a finite-dimensional vector space $\mathfrak{g}$ over k. Define k-bilinear map

$$[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \text{ satisfying}$$

(1) Skew-symmetric: $[X, Y] + [Y, X] = 0$ for all $X, Y \in \mathfrak{g}$.
(2) Jacobi Identity: $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ for all $X, Y, Z \in \mathfrak{g}$.

Remark 4.1.1.1. *Recall that we say that $\mathfrak{g}$ is an abelian Lie algebras if $[X, Y] = 0$ for $X, Y \in \mathfrak{g}$. From the skew-symmetry, it is easy to see that we can make every vector space an abelian Lie algebra, i.e., by defining a trivial bracket. Note that any associative algebra becomes a Lie algebra by defining*

$$[X, Y] = X \cdot Y - Y \cdot X.$$

Of course, we denote $\cdot$ as the multiplication structure of the given associative algebra. So, the matrix algebra

$$\mathfrak{gl}(n) \overset{def}{=} \mathbf{M}_n(k)$$

becomes a Lie algebra. More generally, every finite-dimensional Lie algebra is a Lie sub algebra of the matrix algebra.

Recall that from Exercise 2.1.2.1 for an algebra A

$$\mathbf{Der}(A) = \{D : A \to A \mid D(ab) = D(a)b + aD(b) \text{ and } D \text{ is linear}\}$$

is a Lie algebra.

Remark 4.1.1.2. *By defining for $X, Y \in \mathfrak{g}$*

$$\mathbf{ad}(X) : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) = \mathbf{End}(\mathfrak{g}) \text{ by } \mathbf{ad}(X)(Y) = [X, Y],$$

from the Jacobi identity and skew-symmetry of $[,]$ in $\mathfrak{g}$, it is easy to see that

$$\mathbf{ad}(X)([Y, Z]) = [\mathbf{ad}(X)(Y), Z] + [Y, \mathbf{ad}(X)Z].$$

That is, $\mathbf{ad}(X)$ *is a derivation of* $\mathfrak{g}$*:*

$$\mathfrak{g} \xrightarrow{\mathbf{ad}} \mathbf{Der}(\mathfrak{g}) \hookrightarrow \mathfrak{gl}(\mathfrak{g}).$$

Exercise 4.1.1.1. *Show that* **ad** *is a Lie algebra homomorphism, i.e.,*

$$\mathbf{ad}([X,Y]) = [\mathbf{ad}(X), \mathbf{ad}(Y)].$$

Proof. Since $(\mathbf{Ad})_* = \mathbf{ad}$, it is a Lie algebra homomorphism. See the below of Remark 3.1.2.2. □

Suppose that we have $\mathfrak{a} \hookrightarrow \mathfrak{g}$ of Lie algebras. That is, $\mathfrak{a}$ is a Lie subalgebra of $\mathfrak{g}$, i.e., $[X,Y] \in \mathfrak{a}$ for $X, Y \in \mathfrak{a}$.

Exercise 4.1.1.2. *Every one dimensional linear subspace is an abelian Lie subalgebra.*

Proof. By skew-symmetry, it is necessarily an abelian Lie subalgebra. □

Note that by Theorem 3.1.6.4, there exists a connected Lie subgroup H corresponding a 1-dimensional abelian Lie subalgebra.

Notation 4.1.1.1. *By definition,* $\mathfrak{a}$ *is an ideal of* $\mathfrak{g}$ *if* $[A,X] \in \mathfrak{a}$ *for all* $A \in \mathfrak{a}$ *and* $X \in \mathfrak{g}$*. By notational convention, we denote*

$$\mathfrak{a} \lhd \mathfrak{g}.$$

Of course, since $\mathfrak{a}$ *is always a normal subgroup of* $\mathfrak{g}$*, this notation also makes sense.*

Let $\mathfrak{a} \xrightarrow{\varphi} \mathfrak{g}$ be a Lie algebra homomorphism. Clearly,

$$\mathbf{Ker}(\varphi) = \{a \in \mathfrak{a} \mid \varphi(a) = 0\}$$

is an ideal of $\mathfrak{a}$. Conversely, since $\mathfrak{j} \lhd \mathfrak{g}$ implies that $\mathfrak{g}/\mathfrak{j}$ inherits the Lie algebra structure of $\mathfrak{g}$, we have a surjective Lie algebra homomorphism

$$\mathfrak{g} \to \mathfrak{g}/\mathfrak{j} \text{ with } \mathfrak{j} \text{ as the kernel.}$$

That is, every ideal $\mathfrak{j}$ is the kernel of a Lie algebra homomorphism. Also, note that for a given Lie algebra homomorphism $\mathfrak{a} \xrightarrow{\varphi} \mathfrak{g}$, we have a Lie subalgebra

$$\mathfrak{a}/\mathbf{Ker}(\varphi) \cong \mathbf{Image}(\varphi) = \varphi(\mathfrak{a}).$$

Clearly, for all ideals $\mathfrak{j} \lhd \mathfrak{g}$,

$$\varphi^{-1}(\mathfrak{j}) \lhd \mathfrak{a}.$$

Note that for all ideals $\mathfrak{j} \lhd \mathfrak{g}$ and a given Lie subalgebra $\mathfrak{a} \subseteq \mathfrak{g}$, we have $\mathfrak{a} + \mathfrak{j}$ is a Lie subalgebra and $\mathfrak{a} \cap \mathfrak{j}$ is an ideal of $\mathfrak{a}$. So, the second Noether's isomorphism theorem says that

$$\frac{\mathfrak{a} + \mathfrak{j}}{\mathfrak{j}} \cong \frac{\mathfrak{a}}{\mathfrak{a} \cap \mathfrak{j}}.$$

Definition 4.1.1.1 (Direct sum of Lie algebras). *Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be a Lie algebra. For $X_1, Y_1 \in \mathfrak{g}_1$ and $X_2, Y_2 \in \mathfrak{g}_2$ by defining a Lie bracket as*

$$[X_1 \oplus X_2, Y_1 \oplus Y_2] = [X_1, Y_1] \oplus [X_2, Y_2]$$

we have a Lie algebra

$$\mathfrak{g}_1 \oplus \mathfrak{g}_2.$$

It is easy to see that

$$\mathfrak{g}_i \lhd \mathfrak{g}_1 \oplus \mathfrak{g}_2 \text{ and each projection } \mathfrak{g}_1 \oplus \mathfrak{g}_2 \to \mathfrak{g}_i \text{ is a Lie algebra homomorphism.}$$

Remark 4.1.1.3. *Since an abelian Lie algebra is nothing but a vector space, we deduce that an abelian Lie algebra is a direct sum of* 1*-dimensional Lie algebras, which are* 1*-dimensional vector spaces.*

Definition 4.1.1.2 (Semi-direct product of Lie algebras). *Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be Lie algebras and suppose that $\mathfrak{g}_1$ acts on $\mathfrak{g}_2$ by derivations, i.e., there is a Lie algebra homomorphism φ*

$$\mathfrak{g}_1 \xrightarrow{\varphi} \mathbf{Der}(\mathfrak{g}_2).$$

*Define $\mathfrak{g} = \mathfrak{g}_2 \rtimes_\varphi \mathfrak{g}_1$ to be $(\mathfrak{g}_2, \mathfrak{g}_1)$ as a **underlying vector space** and to have a **Lie bracket** as follows: For $X_1, Y_1 \in \mathfrak{g}_1$ and $X_2, Y_2 \in \mathfrak{g}_2$,*

$$[(X_2, X_1), (Y_2, Y_1)] = ([X_2, Y_2] + \varphi(X_1)Y_2 - \varphi(Y_1)X_2, [X_1, Y_1]).$$

Identifying $X \in \mathfrak{g}_2$ with $(X, 0) \in \mathfrak{g}_2 \rtimes_\varphi \mathfrak{g}_1$, for any $(X_2, X_1) \in \mathfrak{g}_2 \rtimes_\varphi \mathfrak{g}_1$ we deduce that

$$[(X, 0), (X_2, X_1)] = ([X, X_2] - \varphi(X_1)X, [0, X_1]) \in \mathfrak{g}_2 \times \{0\}.$$

That is, $\mathfrak{g}_2 \times \{0\} \lhd \mathfrak{g}_2 \rtimes_\varphi \mathfrak{g}_1 = \mathfrak{g}$. Clearly we have a canonical Lie algebra isomorphism:

$$\pi : \frac{\mathfrak{g}}{\mathfrak{g}_2 \times \{0\}} \to \{0\} \times \mathfrak{g}_1.$$

On the other hand, note that identifying $X \in \mathfrak{g}_1$ with $(0, X) \in \mathfrak{g}_2 \rtimes_\varphi \mathfrak{g}_1$, for any $(X_2, X_1) \in \mathfrak{g}_2 \rtimes_\varphi \mathfrak{g}_1$ we have

$$[(0, X), (X_2, X_1)] = (\varphi(X)X_2, [X, X_1]).$$

That is, $\{0\} \times \mathfrak{g}_1$ is not necessarily an ideal of $\mathfrak{g}_2 \rtimes_\varphi \mathfrak{g}_1 = \mathfrak{g}$. However, it is easy to see that $\{0\} \times \mathfrak{g}_1$ is a Lie subalgebra of $\mathfrak{g}_2 \rtimes_\varphi \mathfrak{g}_1 = \mathfrak{g}$. From this, now clearly we have a Lie algebra homomorphism

$$\mathbf{ad}_\mathfrak{g}|_{\{0\} \times \mathfrak{g}_1} : \{0\} \times \mathfrak{g}_1 \to \mathbf{Der}(\mathfrak{g}) \text{ from } \mathbf{ad}_\mathfrak{g} : \mathfrak{g} \to \mathbf{Der}(\mathfrak{g}).$$

Exercise 4.1.1.3. *Let $\Phi : \mathfrak{g}_1 \to \mathbf{Der}(\mathfrak{g}_2 \times \{0\})$ by $\Phi(X) = (\varphi(X), 0)$. Show that*

$$\mathbf{ad}_\mathfrak{g}(0, X) = \Phi(X).$$

Proof. We have to show that for all $(X_2, 0) \in \mathfrak{g}_2 \times \{0\}$ we have

$$(\mathbf{ad}_{\mathfrak{g}}(0, X))(X_2, 0) = \Phi(X)(X_2, 0) \stackrel{def}{=} (\varphi(X)X_2, 0).$$

It is a triviality, since

$$(\mathbf{ad}_{\mathfrak{g}}(0, X))(X_2, 0) = [(0, X), (X_2, 0)] = (\varphi(X)X_2, [X, 0]).$$

□

Definition 4.1.1.3. *We say that* $\mathfrak{g}$ *is solvable if there is a sequence*

$$\{0\} = \mathfrak{g}_k \lhd \cdots \lhd \mathfrak{g}_2 \lhd \mathfrak{g}_1 \lhd \mathfrak{g}_0 = \mathfrak{g}$$

such that $\frac{g_i}{\mathfrak{g}_{i-1}}$ *is an abelian Lie algebra.*

Theorem 4.1.1.1 (Lie). *If* $\mathfrak{g}$ *is solvable, then it is represented by upper triangular matrices.*

Theorem 4.1.1.2. *Every Lie algebra has a unique maximal solvable ideal, which is called a* ***radical****.*

Notation 4.1.1.2. *The unique maximal solvable ideal of a given Lie algebra* $\mathfrak{g}$ *is denoted by*

$$\mathbf{rad}(\mathfrak{g}) = \sqrt{\mathfrak{g}}.$$

Definition 4.1.1.4 (Semi-simple Lie algebra). *We say that* $\mathfrak{g}$ *is semi-simple if*

$$\sqrt{\mathfrak{g}} = 0.$$

That is, $\mathfrak{g}$ *is semi-simple if it has no solvable ideals.*

Definition 4.1.1.5 (Simple Lie algebra). *We say that* $\mathfrak{g}$ *is simple if it has no nonzero proper ideals.*

Theorem 4.1.1.3 (Levi decomposition). *Every Lie algebra is a semidirect product of its radical* $\sqrt{\mathfrak{g}}$ *and a semisimple Lie subalgebra* $\mathfrak{s}$*, which is called a Levi decomposition:*

$$\mathfrak{g} = \sqrt{\mathfrak{g}} \rtimes \mathfrak{s}.$$

Moreover,

Theorem 4.1.1.4. *Every semi-simple Lie algebra is a direct sum of simple Lie algebras* $\mathfrak{g}_1, \ldots, \mathfrak{g}_n$*:*

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n.$$