Chapter 1

Nilpotent Lie groups

In this chapter, we define the basic notions from Lie theory and explain the relations between Lie groups and Lie algebras.

1.1 Basics of Lie groups and Lie algebras

Definition 1.1.1. *A* **Lie group** G *is a group that is at the same time a* C^∞ *differentiable manifold, such that the group operations are* C^∞ *maps*[1].

The identity element of G is denoted by e. Write L_g and R_g respectively for the left and right translations by $g \in G$, i.e. $L_g h = gh$ and $R_g h = hg$ for all $g \in G$.

For example, for a finite dimensional vector space V,

$$\mathrm{GL}(V) = \{\text{linear automorphism } A \text{ of } V \text{ with } \det A \neq 0\}$$

is a Lie group. $\mathrm{GL}(V)$ is written as $\mathrm{GL}(d, \mathbb{R})$ if $V = \mathbb{R}^d$.

The Lie algebra $\mathfrak{g}$ of G is the tangent space T_eG of G at e. For example, the Lie algbera of $\mathrm{GL}(V)$ is the space of endomorphisms $\mathfrak{gl}(V) = \mathrm{End}(V) = \{\text{linear isomorphisms of } V\}$. The Lie algebra is a vector space. However, to be a Lie algebra, the vector spaces needs to be endowed with an additional structure called the Lie bracket, which we will define now.

Given an element $g \in G$, the map $\Psi_g : h \to ghg^{-1}$ is a continuous automorphism of G. In particular, it sends e to e and is differentiable.

[1] In fact, G being a topological manifold with continuous group operations would guarantee the existence of such a desired C^∞ structure, or even real analytic structure. This was Hilbert's Fifth Problem and was known by the works of Gleason, Montgomery and Zippin in the 1952.

Moreover, $g \to \Psi_g$ is a group morphism from G to $\mathrm{Aut}(G)$. Let $\mathrm{Ad}_g = D_e\Psi_g$ be the derivative at e, then it is a linear automorphism of $T_eG = \mathfrak{g}$. In particular, it can be regarded as an element of $\mathrm{GL}(\mathfrak{g}) \cong \mathrm{GL}(\dim \mathfrak{g}, \mathbb{R})$. If G is a linear group, i.e. a subgroup of $\mathrm{GL}(d, \mathbb{R})$, it is not hard to see that $\mathrm{Ad}_g X = gXg^{-1}$ for $g \in G$ and $X \in \mathfrak{g} \subset \mathfrak{gl}(d, \mathbb{R}) = \mathrm{Mat}_{d\times d}(\mathbb{R})$. $\mathrm{Ad} : g \to \mathrm{Ad}_g$ is called the **Adjoint** action, as it is a group action of G on $\mathfrak{g}$ by automorphisms. $\Phi : g \to \mathrm{Ad}_g$ is a group morphism from G to $\mathrm{GL}(\mathfrak{g})$.

Let us differentiate one more time at e: $D_e\Phi$ is now a map from $\mathfrak{g}$ to $\mathfrak{gl}(\mathfrak{g}) \cong \mathrm{Mat}_{\dim G\times\dim G}(\mathbb{R})$. For $X \in \mathfrak{g}$, we denote its image $D_e\Phi(X)$ by ad_X, called the **adjoint** map. For $X, Y \in \mathfrak{g}$, $\mathrm{ad}_X Y$ is also denoted by $[X, Y]$, called the **Lie bracket**.

Definition 1.1.2. *The* **Lie algebra** *of G is the tangent space $\mathfrak{g} = T_eG$, equipped with the Lie bracket $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. Some times $\mathfrak{g}$ is also denoted as* $\mathrm{Lie}(G)$*.*

Example 1.1.3. If $G \subset \mathrm{GL}(d, \mathbb{R})$ is a linear group, then $[X, Y] = XY - YX$.

Proof. To calculate the image $\mathrm{ad}_X \in \mathfrak{gl}(\mathfrak{g})$, we differentiate the smooth curve $\mathrm{Ad}_{g(t)} \in \mathrm{GL}(\mathfrak{g})$ at $t = 0$ for a smooth curve $g : \mathbb{R} \to G$ with $g(0) = e$ and $g'(0) = X$. Indeed:

$$\begin{aligned}\mathrm{ad}_X(Y) =&(\mathrm{Ad}_{g(t)} Y)'|_{t=0} = (g(t)Yg(t)^{-1})'|_{t=0}\\ =&(g'(t)Yg(t) + g(t)Y(g(t)^{-1}g'(t)g(t)^{-1}))|_{t=0}\\ =&g'(t)Y - Yg'(t) = XY - YX.\end{aligned}$$

□

In this example, it is clear that when G is a linear Lie group, $[X, Y] = \mathrm{ad}_X Y$ satisfies:

- (biliearity) $[X, aY + bZ] = a[X, Y] + b[X, Z]$ and $[aX + bZ, Y] = a[X, Y] + b[Z, Y]$;
- (skew-symmetricity) $[X, Y] = -[Y, X]$;
- (Jacobi identity) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

In fact, these also hold for general Lie groups.

Proposition 1.1.4. If G is a Lie group, then the Lie bracket on its Lie algebra $\mathfrak{g}$ is a bilinear form and satisfies skew-symmetricity and the Jacobi idenity.

Proof of skew-symmetricity. Bilinearity of $[\cdot,\cdot]$ is clear from the definition. We give the proof of skew-symmetricity, while that of the Jacobi identity is left as Exercise 1.1.1.

For $X, Y \in G$, let $g(t), h(t)$ be smooth curves such that $g(0) = h(0) = e$, $g'(0) = X$, $h'(0) = Y$.

The first claim is that

$$(g(t)h(t))'|_{t=0} = X + Y. \tag{1.1}$$

To see this, it suffices to note that

$$\begin{aligned}(g(t)h(t))'|_{t=0} =& \left(\frac{\partial}{\partial r}(g(r)h(s)) + \frac{\partial}{\partial s}(g(r)h(s))\right)\Big|_{r=s=0} \\ =& \left(\frac{\partial}{\partial r}(g(r)h(s)) + \frac{\partial}{\partial s}(g(r)h(s)g(r)^{-1}g(r))\right)\Big|_{r=s=0} \\ =& \big(R_{h(s)}g'(r)|_{r=0}\big)\Big|_{s=0} + \big((R_{g(r)})_* \operatorname{Ad}_{g(r)} h'(s)|_{s=0}\big)\Big|_{r=0} \\ =& X + Y.\end{aligned}$$

An immediate consequence is the special case $h(t) = g(t)^{-1}$, which yields that $g'(t) + (g(t)^{-1})' = 0$ for all smooth curve $g(t)$ with $g(0) = 0$.

Now we may consider $(*) = \frac{\partial}{\partial t}\frac{\partial}{\partial s}\big(g(t)h(s)g(t)^{-1}h(s)^{-1}\big)$. On the one hand, by earlier claims,

$$(*) = \frac{\partial}{\partial t}(\operatorname{Ad}_{g(t)} Y - Y) = \operatorname{ad}_X Y = [X, Y].$$

On the other hand, by the same argument,

$$\begin{aligned}(*) =& \frac{\partial}{\partial s}\frac{\partial}{\partial t}\big(g(t)h(s)g(t)^{-1}h(s)^{-1}\big) = -\frac{\partial}{\partial s}\frac{\partial}{\partial t}\big(h(s)g(t)h(s)^{-1}g(t)^{-1}\big) \\ =& -[Y, X].\end{aligned}$$

The skew-symmetricity is hence established. □

Lie algebras can be abstractly defined using these properties.

Definition 1.1.5. *A* **Lie algebra** *is a vector space equipped with a Lie bracket $[\cdot,\cdot]$, which is a bilinear form that satisfy both skew-symmetricity and the Jacobi identity.*

Definition 1.1.6. *A Lie algebra $\mathfrak{g}$ is* **abelian** *if its Lie bracket is trivial, i.e. $[X, Y] = 0$ for all X, Y.*

Observe that the skew-symmetricity guarantees that $[X, X] = 0$.

Definition 1.1.7. *A* **Lie group morphism** *is a smooth map between two Lie groups that is at the same time a group morphism.*

A **Lie algebra morphism** *is a linear transform ψ between two algebras such that $[\psi X, \psi Y] = \psi[X, Y]$ for all X, Y.*

Note that $\mathrm{ad} : X \to \mathrm{ad}_X$ is a map from a Lie algebra to $\mathfrak{gl}(\mathfrak{g})$, which is a Lie algebra itself.

Lemma 1.1.8. ad *is a Lie algebra morphism, i.e.* $[\mathrm{ad}_X, \mathrm{ad}_Y] = \mathrm{ad}_{[X,Y]}$ *for all* $X, Y \in \mathfrak{g}$.

Proof. It suffices to show that

$$\begin{aligned}[\mathrm{ad}_X, \mathrm{ad}_Y]Z =& \mathrm{ad}_X \mathrm{ad}_Y Z - \mathrm{ad}\, Y \,\mathrm{ad}_X Z = [X, [Y, Z]] - [Y, [X, Z]] \\ =& [X, [Y, Z]] + [Y, [Z, X]]\end{aligned}$$

and

$$\mathrm{ad}_{[X,Y]} Z = [[X, Y], Z] = -[Z, [X, Y]]$$

are equal, which follows from the Jacobi identity. □

The left translation L_g by an element $g \in G$ sends h to hg. It induces a pushforward $(L_g)_*$ from T_hG to $T_hg(G)$. In particular, for all $X \in \mathfrak{g} = T_eG$, $(L_g)_*X \in T_gG$ defines a smooth vector field on G as g varies in G.

Lemma 1.1.9. *The ordinary differential equation*

$$g'(t) = (L_{g(t)})_*X \tag{1.2}$$

with initial data $g(0) = e$ has a unique solution $g(t)$, which is itself a group morphism from $\mathbb{R}$ to G.

Proof. Since the right hand side $(L_{g(t)})_*X$ of the ODE depends Lipschitz continuously (which is because the group operations are C^∞), by the basic theory of differential equations, the equation, conditional to the initial condition $g(0) = e$, has a unique solution $g(t)$ at least locally near $t = 0$, i.e. in $t \in (-\tau, \tau)$ for some τ.

We now show that if $t, s, t + s \in (-\tau, \tau)$, then $g(t)g(s) = g(t + s)$. In fact,

$$\frac{\partial}{\partial s}(g(t)g(s)) = (L_{g(t)})_*g'(s) = (L_{g(t)})_*(L_{g(s)})_*X = (L_{g(t)g(s)})_*X, \tag{1.3}$$

and

$$\frac{\partial}{\partial s} g(t+s) = g'(t+s) = (L_{g(t+s)})_* X.$$

So $g(t)g(s)$ and $g(t+s)$ solve the same ODE $h'(s) = (L_{h(s)})_* X$ in s, with the same initial data $g(t)g(s)|_{s=0} = g(t+s)|_{s=0} = g(t)$, and therefore coincide by uniqueness. In particular, $g(t)$ commutes with $g(s)$ for all $t, s \in (-\frac{\tau}{2}, \frac{\tau}{2})$.

Given this property, one can easily extend the map $g : (-\tau, \tau) \to G$ to $(-\infty, \infty)$ by $g(nt) = g(t)^n$ for all $n \in \mathbb{N}$ and $s \in (-\frac{\tau}{2}, \frac{\tau}{2})$. One can check that the definition is self consistent, as if $nt = ms$ for $n, m \in \mathbb{N}$, then they both equal mnr where $r = \frac{t}{m} = \frac{s}{n}$. Then $g(t)^n = (g(r)^m)^n = (g(r)^n)^m = g(s)^m$.

Moreover, for all $t, s \in \mathbb{R}$, there exists n such that $t = nt_1$ and $s = ns_1$ for some $t_1, s_1 \in (-\frac{\tau}{2}, \frac{\tau}{2})$, and hence $g(t)g(s) = g(t_1)^n g(s_1)^n = g(t_1 + s_1)^n = g(n(t_1 + s_1)) = g(t+s)$. So $t \to g(t)$ is a group morphism from $\mathbb{R}$. At the same time, observe g is differentiable at every t. This is because for a fixed large value of n, $g(t_1)$ is differentiable at $t_1 = \frac{t}{n}$, and $g(t) = g(t_1)^n$ depends differentiably on $g(t_1)$.

It remains to prove that $g(t)$ is a global solution to (1.2) . Notice that as $g(s)$ is a solution to the equation (1.2) near $s = 0$, (1.3) still applies. Thus $g'(t) = \frac{\partial}{\partial s} g(t+s)\big|_{s=0} = \frac{\partial}{\partial s}(g(t)g(s))\big|_{s=0} = (L_{g(t)g(0)})_* X = (L_{g(t)})_* X$. Hence $g(t)$ is a solution for all $t \in \mathbb{R}$. In particular, it is unique by the fundamental theorem of ordinary differential equations. □

Definition 1.1.10. *Given* $X \in \mathfrak{g}$, *the element* $g(1)$ *is denoted by* $\exp X$, *where* $g(t)$ *is the solution in Lemma 1.1.9. The map* $\exp : \mathfrak{g} \to G$ *is called the* **exponential** *map.*

Obviously, $\exp 0 = e$, since for $X = e$, the equation (1.2) has constant solutions. Another easy corollary to Lemma 1.1.9 is that $\{\exp(tX)\}_{t \in \mathbb{R}}$ is a one-parameter subgroup, i.e. a continuous group morphism from $\mathbb{R}$.

Lemma 1.1.11. *For* $X \in \mathfrak{g}$, $\{\exp(tX)\}_{t \in \mathbb{R}}$ *is the unique one-parameter subgroup* $g(t)$ *satisifying* $g'(0) = X$.

Proof. If $g(t)$ is such a group, then $g(0) = e$ and

$$g'(t) = \frac{\partial}{\partial s}(g(t)g(s))|_{s=0} = (L_{g(t)})_* g'(s)|_{s=0} = (L_{g(t)})_* X.$$

That is, $g(t)$ satisfies (1.2). Hence by Lemma 1.1.9, $\exp(tX)$ is the unique solution. □

In particular,

$$\exp(X)^{-1} = \exp(-X).$$

We remark that exp is locally a diffeomorphism near $0 \in \mathfrak{g}$. This follows from the following

Lemma 1.1.12. *The map* exp *is differentiable at* 0*, and* $D_0 \exp = \mathrm{Id}$*.*

Proof. $(D_0 \exp)X = (\exp(tX))'|_{t=0} = (L_{\exp(tX)})_* X|_{t=0} = (L_e)_* X = X$. □

In general, exp does not have to be either injective or surjective, even if G is connected. Though we will see later in this chapter that, for the main object of these notes, namely connected nilpotent Lie groups, exp is a diffeomorphism between $\mathfrak{g}$ and G.

We now state two consequences of Lemma 1.1.12.

Corollary 1.1.13. *If $\Psi : G \to H$ is a Lie group morphism between two Lie groups G and H. Then $D_e\Psi$ is a Lie algebra morphism and the following diagram commutes:*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{D_e\Psi} & \mathfrak{h} \\ \downarrow{\scriptstyle \exp} & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\Psi} & H \end{array}$$

where $\mathfrak{g} = \mathrm{Lie}(G)$, $\mathfrak{h} = \mathrm{Lie}(H)$*.*

Proof. To show $D_e\Psi$ is a Lie algebra morphism, observe that for $X, Y \in \mathfrak{g}$,

$$\begin{aligned} D_e\Psi([X,Y]) =& D_e\Psi(\frac{\partial}{\partial t} \mathrm{Ad}_{\exp tX} Y) = \frac{\partial}{\partial t} D_e\Psi(\mathrm{Ad}_{\exp tX} Y) \\ =& \frac{\partial}{\partial t}\frac{\partial}{\partial s}\Psi(\exp tX \exp sY \exp(-tX)) \\ =& \frac{\partial}{\partial t}\frac{\partial}{\partial s}\Psi(\exp tX)\Psi(\exp sY)\Psi(\exp(-tX)) \\ =& \frac{\partial}{\partial t} \mathrm{Ad}_{\Psi(\exp tX)} Y = \mathrm{ad}_{\frac{\partial}{\partial t}\Psi(\exp tX)} D_e\Psi(Y) \\ =& \mathrm{ad}_{D_e\Psi(X)} D_e\Psi(Y) = [D_e\Psi(X), D_e\Psi(Y)]. \end{aligned} \tag{1.4}$$

Furthermore, $\exp(tX)$ is a one-parameter subgroup in G, and thus so is $\Psi \circ \exp(tX)$ in H. Thanks to Lemma 1.1.12, the derivative $(\Psi \circ \exp(tX))'|_{t=0}$ is given by $D_e\Psi \circ D_0 \exp(X) = D_e\Psi(X)$. By Lemma 1.1.11, $\exp(D_e\Psi(X))$ is the only one-parameter subgroup with this property. So $\Psi \circ \exp(tX) = \exp(D_e\Psi X)$, which is the statement of the corollary. □

Corollary 1.1.14. *If two connected Lie subgroups H_1, H_2 of G have the same Lie algebra $\mathfrak{h}$, then $H_1 = H_2$.*

Proof. Let U be a small neighborhood of $0 \in \mathfrak{h}$, then by Lemma 1.1.12, exp is a diffeomorphism between U and $\exp U$, which is a neighborhood of identity in both H_1 and H_1. Let H be the group generated by $\exp(U)$. Then $H \subseteq H_1 \cap H_2$. Moreover, H_1' is both open and closed in H_1, as both H_1' and its complement in H_1 are invariant under translations by elements from $\exp U$ and are thus open. Since H_1 is connected, $H = H_1$. Similarly $H = H_2$. $\square$

On the other hand, one may wonder whether there is a Lie subgroup H corresponding to each Lie subalgebra $\mathfrak{h}$. It turns out this is the case.

Proposition 1.1.15. Suppose G is a Lie group and $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g} = \mathrm{Lie}(G)$. Then there is a unique connected subgroup H (not necessarily closed) such that $\mathrm{Lie}(H) = \mathfrak{h}$.

While we will omit the proof of Proposition 1.1.15, let us remark that the idea of it. Every vector $X \in \mathfrak{h}$ can be pushed forward by left translations to get a left invariant vector field F_X. All such vector fields span a distribution $\mathcal{F}$ of subspaces in the vector space bundle TG. The key observation is the Lie bracket between X, Y coincides with the notion of Lie brackets in differential geometry between the two vector fields F_X, F_Y. And that $\mathfrak{h}$ being a Lie algebra makes the distribution $\mathcal{F}$ integrable; that is, there is a submanifold M_g through any given point g such that the tangent bundle TM_g coincides with $\mathcal{F}$ everywhere on M_g. It can then be proved that M_e is the desired subgroup H. For the detailed proof, see e.g. [Hel01, Theorem II.2.1].

Corollary 1.1.16. *Suppose G, H are Lie groups and $\mathfrak{g}$, $\mathfrak{h}$ are their respective Lie algebras. Suppose G is simply connected, then for any Lie algebra morphism $\psi : \mathfrak{g} \to \mathfrak{h}$, there is a unique Lie group morphism $\Psi : G \to H$ such that $\psi = D_e\Psi$.*

Proof. Since ψ interwines the Lie brackets on $\mathfrak{g}$ and $\mathfrak{h}$, the linear graph $\mathfrak{g}^\triangle = (\mathrm{Id} \times \psi)\mathfrak{g}^\triangle = \{(X, \psi X)\}$ is a Lie subalgebra of $\mathfrak{g} \times \mathfrak{h}$. By Proposition 1.1.15, there is a Lie subgroup $G^\triangle$ in $G \times H$ which has $\mathfrak{g}^\triangle$ as the Lie algebra. Note that the derivative $D_e\pi_1$ of the projection $\pi : G^\triangle \to G$ is the projection from $\mathfrak{g}^\triangle$ to $\mathfrak{g}$, which is an Lie algebra isomorphism. Hence π is locally a diffeomorphism. By Lemma 1.1.17 below, we know that π is a covering map. Because G is simply connected and $G^\triangle$ is connected, π must be bijective. So the composition $\pi' \circ \pi^{-1}$ is a group morphism from G to H, where π' stands fro the projection from $G^\triangle$ to H.

Observe that by Corollary 1.1.13, Ψ is completely determined on the image $\exp \mathfrak{g}$ of exp. Since Ψ is a group morphism and $\exp \mathfrak{g}$ generates G, Ψ is unique. □

Lemma 1.1.17. *Suppose a group morphism $\pi : G \to H$ between two topological groups is also a local homeomorphism near $e \in G$; i.e. there exists an open neighborhood U of e such that $\pi|_U$ is a homeomorphism between U and its image $\pi(U)$. Then π is a covering map.*

Proof. For every $g \in G$, gU is an open neighborhood of g, and for $a \in gU$, $\pi(a) = \pi(g)\pi(g^{-1}a)$. So $\pi|_{gU}$ is a homeomorphism between gU and $\pi(g) \cdot \pi(U)$. Hence π is a covering map. □

We finish this section by stating the following well-known theorem of Cartan.

Theorem 1.1.18. *(Cartan) All closed subgroups of a Lie group are Lie subgroups, i.e. submanifolds that are closed under group operations.*

For the proof, follow Exercise 1.1.4.

Notation 1.1.19. In the rest of this book, we shall denote the Lie algebras of Lie groups G, H, K, N, etc. by the corresponding calligraphic lower case letters $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{k}$, $\mathfrak{n}$, etc. without further explanation.

Exercises

Exercise 1.1.1. Prove the Jacobi identity for the Lie algebra $\mathfrak{g}$ of a Lie group G. (Hint: use Corollary 1.1.13 to show that for all $X, Y \in \mathfrak{g}$, $[\mathrm{ad}_X, \mathrm{ad}_Y]_{\mathfrak{gl}(\mathfrak{g})} = \mathrm{ad}_{[X,Y]}$, then apply Example 1.1.3.)

Exercise 1.1.2. Prove that: the solution $g(t)$ in Lemma 1.1.9 coincides with $\exp(tX)$.

Exercise 1.1.3. Prove that: if $G \subseteq \mathrm{GL}(d, \mathbb{R})$ is a linear Lie group, then $\exp(tX) = \sum_{n=0}^{\infty} \frac{(tX)^n}{n!}$ for all $t \in \mathbb{R}$ and $X \in \mathfrak{g} \subseteq \mathfrak{gl}(d, \mathbb{R})$.

Exercise 1.1.4. Assume that H is a connected closed subgroup of a Lie group G. Let $\mathfrak{h} = \{X \in \mathfrak{g} : \exp(tX) \in H, \forall t \in \mathbb{R}\}$.

(1) Show that $\mathfrak{h}$ is a vector subspace of $\mathfrak{g}$.

(2) Show that the projective space $\mathbb{P}\mathfrak{h}$ is closed in $\mathbb{P}\mathfrak{g}$.

(3) Show that exp is surjective from a small neighborhood of $0 \in \mathfrak{h}$ to a sufficiently small neighborhood of $e \in H$.

(5) Show that H is a submanifold and hence a Lie group.

1.2 Lie's Theorems and Ado's Theorem

In this part, we present three theorems due to Sophus Lie, which claim the correspondence between Lie groups and their Lie algebras are, in the weaker sense of local structures near identity, bijective.

Definition 1.2.1. *Two Lie groups G, G' are locally isomorphic if there are open neighborhoods of identity U, U' respectively in G, G' and a smooth diffeomorphism $\Psi : U \to U'$, such that for all $g, h \in U$ such that $gh \in U$, $\Psi(gh) = \Psi(g)\Psi(h)$.*

Theorem 1.2.2. *(Lie) If Lie groups G, G' are locally isomorphic, then $\mathfrak{g}$ and $\mathfrak{g}'$ are isomorphic.*

Theorem 1.2.3. *(Lie) For Lie groups G, G', if their Lie algebras $\mathfrak{g}$, $\mathfrak{g}'$ are isomorphic, then G and G' are locally isomorphic.*

Theorem 1.2.4. *(Lie) For every finite dimensional Lie algebra $\mathfrak{g}$, there exists, up to Lie group isomorphisms, a unique simply connected Lie group G such that $\mathrm{Lie}(G) \cong \mathfrak{g}$.*

Another important theorem is Ado's theorem:

Theorem 1.2.5. *(Ado) Every finite dimensional Lie algebra $\mathfrak{g}$ admits a faithful linear representation, i.e. an injective Lie algebra morphism $\psi : \mathfrak{g} \to \mathfrak{gl}(d, \mathbb{R})$ for some d.*

We first provide the proof of Theorem 1.2.2. The other theorems, though valid for general Lie groups, are easier to prove in the nilpotent case, which is our main focus. These proofs in the nilpotent case will appear in later sections.

Proof of Theorem 1.2.2. Suppose U, V, Ψ are as in Definition 1.2.1. Then $D_e\Psi$ is a Lie algebra morphism. The proof of this fact is the same as in (1.4), while one should note in addition that for all $X, Y \in \mathfrak{g}$, there exists ϵ such that for $t, s \in (-\epsilon, \epsilon)$, all the group multiplications involved in (1.4) take place in either U or U', depending on the context, in which case Ψ behaves like a group morphism. Since Ψ is a diffeomorphism, $D_e\Psi$ is bijective and hence a Lie algebra isomorphism. □

Lemma 1.2.6. *If G is a connected topological group, then the universal cover $\widetilde{G}$ is also a topological group and the projection $\pi : \widetilde{G} \to G$ is a group morphism.*

Proof. Consider the multiplication map $\tau : (a,b) \to ab$ from $G \times G \to G$. Because $\widetilde{G} \times \widetilde{G}$ is the universal cover of G, τ' lifts to a smooth map $\tau : \widetilde{G} \times \widetilde{G} \to \widetilde{G}$ such that $\tau(e,e) = e$. The two elements $\tau(a, \tau(b,c))$ and $\tau(\tau(a,b),c)$ of $\widetilde{G}$ both project to $a'b'c' \in G$, where a', b', c' respectively denote the projections of a, b, c to G. Hence $\tau(a, \tau(b,c))$ is the translate of $\tau(\tau(a,b),c)$ by some $\gamma(a,b,c) \in \Gamma$, where $\Gamma = \pi_1(G)$ is the deck tranformation group associated to the projection $\widetilde{G} \to G$. Since Γ is discrete and γ is continuous in a, b, c, we know that γ must be a constant, and in fact be the identity element by checking $(a,b,c) = (e,e,e)$. So $\tau(a, \tau(b,c)) = \tau(\tau(a,b),c)$. Similarly, one can find a smooth map $\iota : \widetilde{G} \to \widetilde{G}$ that lifts the inverse map $\iota' : a \to a^{-1}$ on G. such that $\iota(e) = e$. Using similar arguments one can show $\tau(a, \iota(a)) = e$. Thus $\widetilde{G}$ is a simply connected group that covers G. □

Proof of Theorem 1.2.3. Let $\widetilde{G}$, $\widetilde{G}'$ respectively be the universal covers of G and G', then by Lemma 1.2.6 they are both simply connected Lie groups whose Lie algebras are both isomorphic to $\mathfrak{g}$. Let $\psi : \mathrm{Lie}(\widetilde{G}) \to \mathrm{Lie}(\widetilde{G}')$ be a Lie algebra isomorphism. Then by Corollary 1.1.16, there exist Lie group morphisms $\Psi : \widetilde{G} \to \widetilde{G}'$ and $\Phi : \widetilde{G}' \to \widetilde{G}$ such that $D_e\Psi = \psi$ and $D_e\Phi = \psi^{-1}$. Then $D_e(\Psi \circ \Phi) = \mathrm{Id}$. By uniqueness in 1.1.16, $\Psi \circ \Phi : \widetilde{G} \to \widetilde{G}$ must be the identity map on G. Similarly, so $\Psi \circ \Phi$ is the identity map on $\widetilde{G}'$. So $\widetilde{G} \cong \widetilde{G}'$ as Lie groups. Therefore, G and G' are both covered by the same Lie group $\widetilde{G}$ and thus locally isomorphic. □

Proof of Theorem 1.2.4 assuming Theorem 1.2.5. By Theorem 1.2.5, there is a Lie algebra $\mathfrak{g}' \in \mathfrak{gl}(d, \mathbb{R})$ for some d such that $\mathfrak{g} \cong \mathfrak{g}'$. By Proposition 1.1.15, one can find a Lie group $G' \subset \mathrm{GL}(d, \mathbb{R})$ such that the Lie algebra of G is $\mathfrak{g}'$. (It should be remarked here that the topology of G' is not inherited from $\mathrm{GL}(d, \mathbb{R})$. Instead, the topology is the one determined by path connectivity if G is not closed.) One then take the universal cover G of G', which is by Lemma 1.2.6 has a group structure that lifts that of G'. In particular, since G' is a Lie group in the current setting, G is a simply connected Lie group with the same Lie algebra $\mathfrak{g}$. □

The proof of Ado's theorem (Theorem 1.2.5) is not a simple one for general Lie algebras, however considerably simpler for nilpotent Lie algebras. The proof in this special case will be presented in the appendix. While the general proof is omitted, the special case is going to be sufficient for our purposes.

Exercises

Exercise 1.2.1. Prove that all: connected abelian Lie groups are isomorphic to $\mathbb{R}^k \times \mathbb{T}^l$ for some $k, l \geq 0$.

1.3 Baker-Campbell-Hausdorff formula

For real numbers x, y, the equality $\exp x \exp y = \exp(x+y)$ always holds. Nevertheless, this property fails for the exponential map on Lie algebras, as a consequence of the non-commutativity represented by $[\cdot, \cdot]$. The Baker-Campbell-Hausdorff formula characterizes how the Lie brackets determines the multiplicative operations of the Lie group G.

Theorem 1.3.1 (Baker-Campbell-Hausdorff formula)**.** *Suppose X, Y are elements of the Lie algebra $\mathfrak{g}$ of a Lie group G, then*

$$\exp X \exp Y = \exp Z \tag{1.5}$$

formally holds where $Z = Z(X, Y)$ is an infinite series whose all terms are of the form

$$b[X_1, [X_2, [\cdots, [X_{m-1}, X_m]] \cdots],$$

where $b \in \mathbb{Q}$ and $X_i \in \{X, Y\}$.

When X, Y are in a sufficiently neighborhood of 0, the equality actually holds.

The first a few terms of $Z(X, Y)$ reads:

$$Z = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + \cdots. \tag{1.6}$$

The omitted terms all have at least three layers of Lie brackets.

Here, the meaning of "formally holds" is that $\exp X \exp Y$ and $\exp Z$ are both real analytic functions of $(X, Y) \in \mathfrak{g}^2$ and their power series agree. In general, the power series does not always converge.

Before going further, we first make the observation that if $[X, Y] = 0$, then $Z = X + Y$. In this case, $\exp X \exp Y = \exp Y \exp X = \exp(X + Y)$. In other words, the Lie bracket is responsible for the non-commutativity in G.

We will discuss the proof of Theorem 1.3.1 only in the case of linear Lie groups. Indeed, this special case implies the theorem for general Lie groups thanks to Ado's theorem (Exercise 1.3.3). In the proof, we will omit certain details (Exercise 1.3.1).

Lemma 1.3.2. *For any Lie algebra* $\mathfrak{g}$ *and* $X, Y \in \mathfrak{g}$,

$$\mathrm{Ad}_{\exp X} Y = \exp(\mathrm{ad}_X) Y = \sum_{n=0}^{\infty} \frac{\mathrm{ad}_X^n Y}{n!} \tag{1.7}$$

Proof. Applying Corollary 1.1.13 to $\Psi = \mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$, we see that $\mathrm{Ad}_{\exp X} = \exp(\mathrm{ad}_X)$, where ad_X is a matrix in $\mathfrak{gl}(\mathfrak{g}) \cong \mathrm{Mat}_{\dim G \times \dim G} \mathbb{R}$. Thus by Exercise 1.1.3,

$$\mathrm{Ad}_{\exp X} Y = \exp(\mathrm{ad}_X) Y = \left(\sum_{n=0}^{\infty} \frac{\mathrm{ad}_X^n}{n!} \right) Y = \sum_{n=0}^{\infty} \frac{\mathrm{ad}_X^n Y}{n!}$$

for all $X, Y \in \mathfrak{g}$. □

When $\mathfrak{g} = \mathfrak{gl}(d, \mathbb{R}) = \mathrm{Mat}_{d \times d}(\mathbb{R})$, both sides in (1.7) can be explicitly written as infinite series, using Exercise 1.1.3 and the relations

$$\mathrm{Ad}_X Y = XYX^{-1}, \mathrm{ad}_X Y = XY - YX. \tag{1.8}$$

Since X, Y are arbitrarily chosen matrices in arbitrary dimensions, it is not hard to believe that (1.7) holds formally for abstract symbols X, Y. In particular, we may express this relation in the setting where $X = X(t)$ is a smooth curve in $\mathfrak{gl}(d, \mathbb{R})$, and Y represents the differential operator $\frac{\mathrm{d}}{\mathrm{d}t}$. While this discussion is not a rigorous proof, we obtain the statement below:

Lemma 1.3.3. *For a smooth curve* $X(t) \in \mathfrak{gl}(d, \mathbb{R})$*, adopting the conventions* (1.8)*, we have the equality*

$$\mathrm{Ad}_{\exp X(t)} \big(\frac{\mathrm{d}}{\mathrm{d}t}\big) = \exp \mathrm{ad}_{X(t)} \big(\frac{\mathrm{d}}{\mathrm{d}t}\big) = \sum_{n=0}^{\infty} \frac{\mathrm{ad}_{X(t)}^n (\frac{\mathrm{d}}{\mathrm{d}t})}{n!} \tag{1.9}$$

between linear differential operators acting on smooth curves $\gamma : \mathbb{R} \to \mathbb{R}^d$.

The rigorous proof of Lemma 1.3.3 is left as Exercise 1.3.1.

One can derive from this lemma the differentiation formula of exp at any base point, which generalizes Lemma 1.1.12.

Corollary 1.3.4. *For* $X, Y \in \mathfrak{gl}(d, \mathbb{R})$,

$$(D_X \exp) Y = (\exp X) \sum_{n=1}^{\infty} \frac{(-1)^{n-1} \mathrm{ad}_X^{n-1}}{n!} Y.$$

Proof. First of all, observe the fact that $\mathrm{ad}_{X(t)} \frac{\mathrm{d}}{\mathrm{d}t} = -X'(t)$ as an operator. To see this, observe that for all smooth curves $\nu(t)$ in $\mathbb{R}^d$,

$$\begin{aligned}\mathrm{ad}_{X(t)} \frac{\mathrm{d}}{\mathrm{d}t}(\nu(t)) =& (X(t) \circ \frac{\mathrm{d}}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t} \circ X(t))\nu(t) = X(t)\nu'(t) - (X \cdot \nu)'(t) \\ =& - X' \cdot \nu(t).\end{aligned}$$

Suppose $\gamma(t)$ is a constant function in t. Applying both sides of (1.9) to $-X(t)$ and $\gamma = \gamma(t)$, we respectively get:

$$\begin{aligned}\mathrm{Ad}_{\exp(-X(t))} \big(\frac{\mathrm{d}}{\mathrm{d}t}\big)(\gamma) =& \exp(-X(t)) \frac{\mathrm{d}}{\mathrm{d}t}(\exp X(t)\gamma) \\ =& \exp(-X(t))(\exp X(t))'\gamma,\end{aligned}$$

and

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{\mathrm{ad}^n_{-X(t)}(\frac{\mathrm{d}}{\mathrm{d}t})}{n!}(\gamma) =& \sum_{n=1}^{\infty} \frac{\mathrm{ad}^{n-1}_{-X(t)}(X'(t))}{n!} \cdot \gamma \\ =& \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\,\mathrm{ad}^{n-1}_{X(t)}(X'(t))}{n!} \cdot \gamma.\end{aligned}$$

Here we used the fact above.

Comparing both equalities, we can conclude the proof. □

Proof of Theorem 1.3.1 for linear Lie groups. Assume that $G \subseteq \mathrm{GL}(d, \mathbb{R})$ and hence $\mathfrak{g} \subseteq \mathfrak{gl}(d, \mathbb{R})$. We will prove the Baker-Campbell-Hausdorff formula holds as an actual equality for matrices X, Y from a sufficiently small neighborhood of 0 in $\mathfrak{g}$. Since $\exp tX \exp tY$ is real analytic in t, this would prove the formula as a formal equality.

When X, Y have small norms, $\exp X \exp tY$ is also close to 0. For $t \in (-2, 2)$, there is a unique $Z(t)$ chosen from a small neighborhood of 0 such that $\exp Z(t) = \exp X \exp tY$. This is because exp is a local diffeomorphism at 0. Moreover, Z depends differentiably on t.

Applying Corollary 1.3.4 to $\exp X \exp tY = \exp Z(t)$ yields

$$(\exp X \exp tY)' = \exp X \exp tY \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\,\mathrm{ad}^{n-1}_{Z(t)}}{n!} Z'(t).$$

As the derivative on the left hand side equals $(\exp X \exp tY)Y$. So

$$Y = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\,\mathrm{ad}^{n-1}_{Z(t)}}{n!} Z'(t).$$

Note that the first term in the power series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}\,\mathrm{ad}_{Z(t)}^{n-1}}{n!}$ is the identity transform $\mathrm{Id} \in \mathrm{GL}(\mathfrak{g})$. Therefore, for small values of $Z(t)$, which is being assumed, the linear transform $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}\,\mathrm{ad}_{Z(t)}^{n-1}}{n!}$ has an inverse transform $\phi(\mathrm{ad}_{Z(t)}) \in \mathrm{GL}(\mathfrak{g})$, where ϕ is the MacLaurin series determined by

$$\phi(z) = \left(\sum_{n=1}^{\infty} \frac{(-1)^{n-1} z^{n-1}}{n!}\right)^{-1} = \frac{z}{1-e^{-z}} = 1 + \frac{z}{2} + \frac{z^2}{12} + \cdots .$$

Then,

$$Z'(t) = \phi(\mathrm{ad}_{Z(t)})Y.$$

Furthermore, remember that for $t \in (-2, 2)$, $X, Y, Z(t)$ are in a small neighborhood U of 0 in $\mathfrak{g}$, and thus so are $\mathrm{ad}_X, \mathrm{ad}\,Y, \mathrm{ad}_{Z(t)}$ in $\mathrm{End}_{\mathfrak{g}}$. When U is small, $\exp|_U$ is a diffeomorphism, and by Lemma 1.3.2,

$$\begin{aligned}\mathrm{ad}_{Z(t)} &= \exp^{-1}(\mathrm{Ad}_{\exp Z(t)}) = \exp^{-1}(\mathrm{Ad}_{\exp X \exp tY}) = \exp^{-1}(\mathrm{Ad}_x \mathrm{Ad}_{tY}) \\ &= \exp^{-1}(\exp(\mathrm{ad}_X)\exp(t\,\mathrm{ad}_Y)).\end{aligned}$$

Since exp is, in this context, from $\mathrm{End}(\mathfrak{g})$ to $\mathrm{GL}(\mathfrak{g})$, it is given by the power series $\exp A = \mathrm{Id} + A + \frac{A^2}{2!} + \cdots$. The inverse $\exp^{-1}$ takes place in a neighborhood of $\mathrm{Id} \in \mathrm{GL}(g)$, and is given by the series $\exp^{-1}(\mathrm{Id} + A) = A - \frac{A^2}{2} + \frac{A^3}{3} - \cdots$. It follows that $\mathrm{ad}_{Z(t)}$ can be written as an infinite power series $\psi(\mathrm{ad}_X, t\,\mathrm{ad}_Y)$ in ad_X and $t\,\mathrm{ad}_Y$, whose terms have the form $c\,\mathrm{ad}_X^{i_1} \circ (t\,\mathrm{ad}_Y)^{j_1} \circ \cdots \circ \mathrm{ad}_X^{i_k}$ or $c\,\mathrm{ad}_X^{i_1} \circ (t\,\mathrm{ad}_Y)^{j_1} \circ \cdots \circ \mathrm{ad}_X^{i_k} \circ (t\,\mathrm{ad}_Y)^{j_k}$, where c is always rational. Hence $\phi \circ \psi$ is a power series of similar form, and $Z'(t) = \phi \circ \psi(\mathrm{ad}_X, \mathrm{ad}_Y)Y$.

Integrating on $t \in [0, 1]$, we know that

$$Z(1) = Z(0) + \int_0^1 \phi \circ \psi(\mathrm{ad}_X, \mathrm{ad}_Y)Y\,\mathrm{d}t = X + \int_0^1 \phi \circ \psi(\mathrm{ad}_X, \mathrm{ad}_Y)Y\,\mathrm{d}t.$$

Remark that this is a series of the form in the statement of Theorem 1.3.1. Because $\exp Z(1) = \exp X \exp Y$, the theorem is proved. □

The constant and first order terms in the series of $\exp(\mathrm{ad}_X)\exp(t\,\mathrm{ad}_Y)$ are $\mathrm{Id} + \mathrm{ad}_X + t\,\mathrm{ad}_Y$, thus $\psi(\mathrm{ad}_X, \mathrm{ad}_Y) = \mathrm{ad}_X + t\,\mathrm{ad}_Y +$(higher order terms). And $\phi \circ \psi(\mathrm{ad}_X, \mathrm{ad}_Y) = \mathrm{Id} + \frac{1}{2}(\mathrm{ad}_X + t\,\mathrm{ad}_Y) +$ (higher order terms) as well. So

$$Z(1) = X + \int_0^1 (Y + \frac{1}{2}([X, Y] + t[Y, Y] \cdots)\mathrm{d}t = X + Y + \frac{1}{2}[X, Y] + \cdots ,$$

where "$\cdots$" represents higher order terms involving two or more Lie brackets.

Exercises

Exercise 1.3.1. Prove Lemma 1.3.3. (Hint: Compute all the partial derivatives in s of the real analytic expression $e^{-sX(t)}\frac{\partial}{\partial t}(e^{sX(t)}\gamma(t))$ at $s = 0$.)

Exercise 1.3.2. Verify that the third order terms at the beginning of the Baker-Campbell-Hausdorff formula are those in (1.6).

Exercise 1.3.3. Assuming Ado's theorem, prove that: if the Baker-Campbell-Hausdorff formula holds for all linear Lie groups, then it holds for all Lie groups. (Hint: use Ado's theorem and Proposition 1.1.15 to prove that given a Lie group G, there is a linear Lie group G' locally isomorphic to G.)

1.4 Nilpotent Lie algebras

The Baker-Campbell-Hausdorff formula tells us that how the local commutator relations determine the structure of a Lie groups. However, in practice, it is difficult to do so, because of the following reasons. First, the coefficients in the formula, though can be made explicit with some efforts, are highly complicated. Second, there can be infinitely many terms, because the repeated Lie brackets can be infinitely long. The group relations are analytic, however not polynomial in general.

Nilpotent Lie groups are Lie groups whose Baker-Campbell-Hausdorff formula has only finitely many terms, or equivalently in whose Lie algebra repeated brackets of sufficiently high orders all vanish. In this case, the group structure is characterized by polynomial equations. We will give more precise definitions of such Lie algebras in this section.

If $\mathfrak{h}, \mathfrak{h}' \subset \mathfrak{g}$ is a Lie subalgebra. Denote $[\mathfrak{h}, \mathfrak{h}'] = \mathbb{R}\text{-span}\{[X, Y] : X \in \mathfrak{h}, Y \in \mathfrak{h}'\}$. It can be easily verified to be a vector subspace of $\mathfrak{g}$. The subalgebra $\mathfrak{h}$ is called an **ideal** of $\mathfrak{g}$ if $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$.

Lemma 1.4.1. *If $\mathfrak{h}$, $\mathfrak{h}'$ are ideals of $\mathfrak{g}$, then $[\mathfrak{h}, \mathfrak{h}']$ is an ideal of each of $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{h}'$.*

Proof. Let X, Y, Z be respectively vectors from $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{h}'$, then

$$\begin{aligned}[X, [Y, Z]] =& -[Y, [Z, X]] - [Z, [X, Y]] \in [\mathfrak{h}, [\mathfrak{h}', \mathfrak{g}]] + [\mathfrak{h}', [\mathfrak{g}, \mathfrak{h}]] \\ =& [\mathfrak{h}, \mathfrak{h}'] + [\mathfrak{h}', \mathfrak{h}] = [\mathfrak{h}, \mathfrak{h}'].\end{aligned}$$

Hence $[\mathfrak{g}, [\mathfrak{h}, \mathfrak{h}']] \subseteq [\mathfrak{h}, \mathfrak{h}']$, or in other words $[\mathfrak{h}, \mathfrak{h}']$ is an ideal in $\mathfrak{g}$.

Furthermore, because $\mathfrak{h}$ and $\mathfrak{h}'$ are ideals, $[\mathfrak{h}, \mathfrak{h}'] \subseteq \mathfrak{h} \cap \mathfrak{h}'$. It then follows that $[\mathfrak{h}, \mathfrak{h}']$ is also an ideal in both $\mathfrak{h}$ and $\mathfrak{h}'$.

$[\mathfrak{h}, [\mathfrak{g}, \mathfrak{h}]] \subseteq [\mathfrak{g}, [\mathfrak{g}, \mathfrak{h}]] \subseteq [\mathfrak{g}, \mathfrak{h}]$. In particular, $[[\mathfrak{g}, \mathfrak{h}], [\mathfrak{g}, \mathfrak{h}]] \subseteq [\mathfrak{g}, \mathfrak{h}]$, so $\mathfrak{h}$ is a subalgebra, and furthermore and ideal. □

If $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then the quotient vector space $\mathfrak{g}/\mathfrak{h}$ can be equipped with a Lie bracket descending from that of $\mathfrak{g}$. Namely, if $X_1, Y_1 \in \mathfrak{h}$ are respectively represented by $X, Y \in \mathfrak{g}$, then $[X_1, Y_1] \in \mathfrak{g}/\mathfrak{h}$ is the equivalence class represented by $[X, Y]$. If X', Y' also represent X_1, Y_1 respectively, then $X' = X + A$, $Y' = Y + B$ where $A, B \in \mathfrak{h}$. So $[X', Y'] = [X, Y] + [X, B] + [A, Y'] \in [X, Y] + \mathfrak{h}$. Thus the Lie bracket on $\mathfrak{g}/\mathfrak{h}$ is independent of the choice of representatives X, Y, and hence well-defined.

Definition 1.4.2. *Given a Lie algebra* $\mathfrak{g}$*, the* **lower central series** *(or* **derived series***) is the sequence of Lie subalgebras* $\mathfrak{g} = \mathfrak{g}_{(1)} \rhd \mathfrak{g}_{(2)} \rhd \mathfrak{g}_{(3)} \rhd \cdots$ *inductively defined by* $\mathfrak{g}_{(i+1)} = [\mathfrak{g}, \mathfrak{g}_{(i)}]$.

Lemma 1.4.1 guarantees that the $\mathfrak{g}_{(i)}$'s are ideals of G and $\mathfrak{g}_{(i)} \lhd \mathfrak{g}_{(j)}$ if $i \geq j$.

The word "central" in the name comes from the fact that $\mathfrak{g}_{(i)}/\mathfrak{g}_{(i+1)}$ is in the **center** of $\mathfrak{g}/\mathfrak{g}_{(i+1)}$, i.e. $[X, Y] = 0$ for all $X \in \mathfrak{g}_{(i)}/\mathfrak{g}_{(i+1)}$ and $Y \in \mathfrak{g}/\mathfrak{g}_{(i+1)}$.

Lemma 1.4.3. $[\mathfrak{g}_{(i)}, \mathfrak{g}_{(j)}] \subseteq \mathfrak{g}_{(i+j)}$.

Proof. We prove the statement inductively as i increases. For $i = 1$ and all j, this is the definition of lower central series. Suppose $i \geq 2$ and $[\mathfrak{g}_{(i-1)}, \mathfrak{g}_{(j)}] \subseteq \mathfrak{g}_{(i+j-1)}$, then it suffices to show $[[X, Y], Z] \in \mathfrak{g}_{i+j-1}$ for all $X \in \mathfrak{g}$, $Y \in \mathfrak{g}_{(i-1)}$ and $Z \in \mathfrak{g}_{(j)}$ because $[\mathfrak{g}_{(i)}, \mathfrak{g}_{(j)}]$ is linearly spanned by vectors of this form. By Jacobi identity, it suffices to show $[X, [Y, Z]]$ and $[Y, [X, Z]]$ are both in $\mathfrak{g}_{(i+j)}$. By inductive assumption, $[X, [Y, Z]] \in [\mathfrak{g}, \mathfrak{g}_{(i+j-1)}] \subseteq \mathfrak{g}_{(i+j)}$, and $[Y, [X, Z]] \in [\mathfrak{g}_{(i-1)}, \mathfrak{g}_{(j+1)}] \subseteq \mathfrak{g}_{(i+j)}$. The proof is completed. □

Definition 1.4.4. *A Lie algebra* $\mathfrak{g}$ *is* **nilpotent** *if* $\mathfrak{g}_{(i+1)} = \{0\}$ *for some* $i \in \mathbb{N}$*. The smallest such* i *is called the* **step of nilpotency***, and* $\mathfrak{n}$ *is called* s**-step nilpotent**.

If the step of nilpotency is s, then $\mathfrak{g}_{(n)} = 0$ for all $n \geq s + 1$.

Lemma 1.4.5. $\mathfrak{g}$ *is nilpotent with step of nilpotency at most* s*, if and only if* $[X_1, [X_2, [\cdots, [X_s, X_{s+1}] \cdots]] = 0$ *for all* $X_1, \cdots, X_{s+1} \in \mathfrak{g}$.

Proof. All such iterated Lie brackets belong to $\mathfrak{g}_{(s+1)}$, and conversely $\mathfrak{g}_{(s+1)}$ is generated by these special vectors. □

If the step of nilpotency is s, then $\mathfrak{g}_{(n)} = 0$ for all $n \geq s + 1$.

Lemma 1.4.6. *If $\mathfrak{g}$ is nilpotent, with step s, then all of its Lie subalgebras and quotient Lie algebras are also nilpotent and their step of nilpotencies are at most s.*

Proof. Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a subalgebra and $\mathfrak{f} \subseteq \mathfrak{g}$ be an ideal, then one can easily prove by induction that

$$\mathfrak{h}_{(i)} \subseteq \mathfrak{g}_{(i)}$$

and

$$(\mathfrak{g}/\mathfrak{f})_{(i)} = \mathfrak{g}_{(i)}/(\mathfrak{f} \cap \mathfrak{g}_{(i)}).$$

Both of these become trivial when $i = s + 1$. □

Definition 1.4.7. *A matrix $A \in \mathrm{Mat}_{d\times d}(\mathbb{R})$ is* **nilpotent** *(resp.* **unipotent***) if all of its eigenvalues are equal to 0 (resp. 1).*

Lemma 1.4.8. *If $\mathfrak{g}$ is nilpotent, then for every $X \in \mathfrak{g}$, $\mathrm{ad}_X \in \mathrm{End}(\mathfrak{g})$ is a nilpotent matrix and $\mathrm{ad}_X^s = 0$.*

Proof. Using Jordan canonical form, one can see that a matrix A is nilpotent if and only if $A^n = 0$ for sufficiently large n. This is true for ad_X with $n = s$. □

Example 1.4.9. By definition, a Lie algebras is 1-step nilpotent if and only if it is abelian.

Example 1.4.10. Let $\mathfrak{n} \subseteq \mathfrak{gl}(d, \mathbb{R})$ denote the vector subspace of consisting of strictly upper triangular matrices

$$\begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1,d-1} & a_{1,d} \\ & 0 & a_{23} & \cdots & a_{2,d-1} & a_{2,d} \\ & & \ddots & & \vdots & \vdots \\ & & & \ddots & \vdots & \vdots \\ & & & & 0 & a_{d-1,d} \\ & & & & & 0 \end{pmatrix}$$

Then $\mathfrak{n}$ is closed under $[\cdot,\cdot]$; that is, for all $X, Y \in \mathfrak{n}$, $[X, Y] = XY - YX$ is still in $\mathfrak{n}$. This is because $\mathfrak{n}$ is closed under matrix multiplication and both XY and YX are also in $\mathfrak{n}$.

Furthermore, $\mathfrak{n}$ can be written as $\{(a_{ij})_{i,j=1}^d : a_{ij} = 0 \text{ unless } j \geq i + 1\}$. For $X = (x_{ij}), Y = (y_{ij}) \in \mathfrak{n}$, we now calculate the entries of $[X, Y]$.

The entry of $[X, Y]$ with index ij is given by $\sum_{k=1}^d (x_{ik}y_{kj} - y_{ik}x_{kj})$. In order for any term of the form $x_{ik}y_{kj}$ or $y_{ik}x_{kj}$ to be non-zero, one needs

$k \geq i+1$ and $j \geq k+1$, so $j \geq i+2$. Thus $\mathfrak{n}_{(2)} = [\mathfrak{n}, \mathfrak{n}]$ is inside the subspace $\{(a_{ij})_{i,j=1}^{d} : a_{ij} = 0 \text{ unless } j \geq i+2\}$. Indeed, $\mathfrak{n}_{(2)}$ coincides with this space (Exercise 1.4.3).

Inductively, we can show $\mathfrak{n}_{(k)} \subseteq [\mathfrak{n}, \mathfrak{n}]$ is contained in (in fact, equal to) the vector subspace $\{(a_{ij})_{i,j=1}^{d} : a_{ij} = 0 \text{ unless } j \geq i+k\}$. Starting at $k = d$, this space is trivial. Hence $\mathfrak{n}$ is $(d-1)$-step nilpotent.

Example 1.4.11. The $(2n+1)$-dimensional Heisenberg Lie algebra $\mathfrak{h}_{2n+1}$ is the vectors space spanned by $2n+1$ linearly independent vectors $X_1, \cdots, X_n$, $Y_1, \cdots, Y_n$, Z. All the Lie brackets among these base vectors are trivial, except: $[X_i, Y_i] = -[Y_i, X_i] = Z$.

Then all three-fold iterated Lie brackets of the form $[*, [*, *]]$ vanish, hence $(\mathfrak{h}_{2n+1})_{(3)} = \{0\}$. But $(\mathfrak{h}_{2n+1})_{(2)}$ is non-trivial because the Lie bracket is non-trivial. Hence $\mathfrak{g}_{2n+1}$ is two-step nilpotent.

We finish this chapter by a useful fact. To state it, first let $\mathfrak{g}_{(2)} = [\mathfrak{g}, \mathfrak{g}]$ be the commutator subalgebra of a nilpotent Lie algebra $\mathfrak{g}$. Then $\mathfrak{g}_{(2)}$ is an ideal. Suppose in addition that $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$.

Lemma 1.4.12. *For a Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, if $\mathfrak{h} + \mathfrak{g}_{(2)} = \mathfrak{g}$ then $\mathfrak{h} = \mathfrak{g}$.*

Proof. We shall inductively prove that $\mathfrak{h} + \mathfrak{g}_{(i)} = \mathfrak{g}$ for every $\mathfrak{g}_{(i)}$ in the lower central series with $i \geq 2$. As the lower central series eventually becomes trivial, this shows $\mathfrak{h} = \mathfrak{g}$.

Suppose the inductive step is valid for index i. Then every $X \in \mathfrak{g}$ can be written as $Y + Z$ where $Y \in \mathfrak{h}$ and $Z \in \mathfrak{g}_{(i)}$. The claim for $i+1$ would follow if we could further show that for every $Z \in \mathfrak{g}_{(i)}$, $Z \in \mathfrak{h} + \mathfrak{g}_{(i+1)}$. Without loss of generality we may assume $Z = [U, V]$ where $U \in \mathfrak{g}$ and $V \in \mathfrak{g}_{(i-1)}$ as $\mathfrak{g}_{(i)}$ is spanned by such vectors. By inductive hypothesis, $U = Y_U + Z_U$ and $V = Y_V + Z_V$, where $Y_U, Y_V \in \mathfrak{h}$ and $Z_U, Z_V \in \mathfrak{g}_{(i)}$. Then

$$\begin{aligned} [U, V] =& [Y_U, Y_V] + ([Y_U, Z_V] + [Z_U, V]) \\ \in& \mathfrak{h} + ([\mathfrak{g}, \mathfrak{g}_{(i)}] + [\mathfrak{g}_{(i)}, \mathfrak{g}]) \subseteq \mathfrak{h} + \mathfrak{g}_{(i+1)}. \end{aligned}$$

The proof is complete. □

Exercises

Exercise 1.4.1. Show that in the lower central series, $[\mathfrak{g}_{(i)}, \mathfrak{g}_{(j)}] \lhd \mathfrak{g}_{(i+j)}$.

Exercise 1.4.2. The **upper central series** $\{0\} = \mathfrak{g}^{(0)} \subseteq \mathfrak{g}^{(1)} \subseteq \mathfrak{g}^{(2)} \subseteq \cdots$ of a Lie algebra $\mathfrak{g}$ is the sequnce of subalgebras inductively defined by

$$\mathfrak{g}^{(i+1)} = \{X \in \mathfrak{g} : [X, Y] \in \mathfrak{g}^{(i)}, \forall Y \in \mathfrak{g}\}.$$

Show that:

(1) $\mathfrak{g}$ is nilpotent if and only if $\mathfrak{g}^{(n+1)} = \mathfrak{g}$ for some $n \in \mathbb{N}$;

(2) The smallest such n equals the step of nilpotency s of $\mathfrak{g}$.

(3) $\mathfrak{g}_{(n)} \subseteq \mathfrak{g}^{(s-n+1)}$ for all $n \leq s$.

Exercise 1.4.3. In Example 1.4.10, show that

$$\mathfrak{g}_{(k)} = \{(a_{ij})_{i,j=1}^{d} : a_{ij} = 0 \text{ unless } j \geq i + k\}.$$

Exercise 1.4.4. Show that the Lie algebra $\mathfrak{n}$ of strictly upper triangular 3×3 matrices is isomorphic to the 3-dimensional Heisenberg Lie algebra.

Exercise 1.4.5. If $A \in \mathrm{Mat}_{d\times d}(\mathbb{R})$ is a nilpotent matrix, then

$$A = \sum_{n=1}^{d-1} \frac{(-1)^{n-1}}{n} (\exp A - \mathrm{Id})^n.$$

1.5 Nilpotent Lie groups

Nilpotent Lie groups are defined in a similar way as in the Lie algebra category.

For subgroups H, H' of a G, let the commutator group $[H, H']$ be the subgroup generated by $\{aba^{-1}b^{-1} : a \in H, b \in H'\}$. Note that H is a normal subgroup of G (denoted as $H \lhd G$) if and only if $[G, H] \subseteq H$.

It is easy to check that for $H \subseteq G$, $[G, H]$ is a normal subgroup of G. Similary to Lemma 1.4.1, if in addition $H \lhd G$ is a normal subgroup, then $[G, H] \subseteq H$, and hence is also normal in H. If we start with $H = G$, this defines a sequence of decreasing normal subgroups:

Definition 1.5.1. *Given a group G, its* **lower central series** *(or* **derived series***) is the sequence of normal subgroups $G = G_{(1)} \rhd G_{(2)} \rhd G_{(3)} \rhd \cdots$ inductively defined by $G_{(i+1)} = [G, G_{(i)}]$.*

Definition 1.5.2. *A Lie group G is* **nilpotent** *if $G_{(i+1)} = \{0\}$ for some $i \in \mathbb{N}$. The smallest such i is called the* **step of nilpotency***, and G is called* **s-step nilpotent***.*

A group is abelian if and only if it is 1-step nilpotent.

Similar to Lemma 1.4.5, we have

Lemma 1.5.3. *If G is nilpotent, with step s, then all of its subgroups and quotient groups are also nilpotent and their step of nilpotencies are at most s.*

Proof. Let $H \subseteq G$ be a subgroup and $F \subseteq G$ be a normal subgroup, then induction shows that $H_{(i)} \subseteq G_{(i)}$ and $(G/F)_{(i)} = G_{(i)}/(F \cap G_{(i)})$. Both series become trivial when $i = s + 1$. □

Lemma 1.5.4. *If G is a connected Lie group, H is a connected Lie subgroup and $\mathfrak{g}$, $\mathfrak{h}$ are the corresponding Lie algebras, then $H \lhd G$ if and only if $\mathfrak{h} \lhd \mathfrak{g}$.*

Proof. Suppose $H \lhd G$, then for all differentiable curves $g(t) \in G$ and $h(s) \in H$ with $g(0) = h(0) = e$, $g(t)h(s)g(t)^{-1} \subseteq H$. Differentiating in s, we see that $\mathrm{Ad}_{g(t)} h'(0) \in \mathfrak{h}$. Differentiating again in t yields $\mathrm{ad}_{g'(0)} h'(0) \in \mathfrak{h}$. As $g'(0) \in \mathfrak{g}$ and $h'(0) \in \mathfrak{h}$ can be arbitrarily chosen, it follows that $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$, or equivalently $\mathfrak{h} \lhd \mathfrak{g}$.

Conversely, if $\mathfrak{h} \lhd \mathfrak{g}$, we wish to show that $ghg^{-1} \in H$ for all $g \in G$, $h \in \mathfrak{h}$. Let U, V be neighborhoods of 0 respectively in $\mathfrak{g}$ and $\mathfrak{h}$ such that $\exp|_U : U \to \exp U$ and $\exp|_V : V \to \exp V$ are diffeomorphisms.

We first show that $\mathrm{Ad}_g \mathfrak{h} \subseteq \mathfrak{h}$. First, suppose that $g = \exp X$ for some $X \in U$. Because ad_X preserves the Lie subalgebra $\mathfrak{h}$, so does $\mathrm{Ad}_g = \exp(\mathrm{ad}_X)$. The claim follows for general g, because $\mathrm{Ad} : g \to \mathrm{Ad}_g$ is a group morphism and G, as a connected topological group, is generated by the neighborhood $\exp U$ of identity.

We then prove $ghg^{-1} \in H$ for all $h \in H$. If $h = \exp Y$ for some $Y \in V$, then $h = h(1)$ where $h = h(1)$ where $h(t) = \exp tY$ is the one parameter subgroup with initial velocity $h'(0) = Y$. Hence $gh(t)g^{-1}$ is the one-parameter subgroup with initial veclocity $\mathrm{Ad}_g Y \in \mathfrak{h}$. As such a subgroup exists in H and is unique in G, we obtain that $gh(t)g^{-1} \in H$. So $ghg^{-1} \in H$. To generalize this to all $h \in H$, it suffices to note that H is connected and thus generated by $\exp V$. This completes the proof. □

Definition 1.5.5. *We define* $G_{(i)}$**-admissible curves** $\gamma : \mathbb{R} \to G_{(i)}$ *recursively as follows:*

For $i = 1$, a $G_{(1)}$ admissible curve is a one-parameter subgroup $\{\exp tY\}$ in G.

For $i \geq 1$, a $G_{(i)}$-admissible curve is a curve of the form

$$\begin{aligned}\exp(tY_1)\gamma_1(t)\exp(-tY_1)\gamma_1(t)^{-1}\exp(tY_2)\gamma_2(t)\exp(-tY_2)\gamma_2(t)^{-1}\\ \cdots\exp(tY_n)\gamma_n(t)\exp(-tY_n)\gamma_n(t)^{-1},\end{aligned} \tag{1.10}$$

where $1 \le n \le \dim \mathfrak{g}_{(i)}$*,* $Y_j \in \mathfrak{g}$*, and the* γ_j*'s are* $G_{(1)}$ *admissible curves.*

Lemma 1.5.6. *Suppose* G *is a connected Lie groups, then for all* $i \ge 1$ *and* $Z \in \mathfrak{g}_{(i)}$*, there exists a* C^1 *differentiable curve* $\gamma : \mathbb{R} \to G_{(i)}$ *such that* $g(0) = e$*,* $g'(0) = Z$*. Moreover,* γ *is* $G_{(i)}$*-admissible.*

Proof. We prove by induction. For $i = 1$, $G_{(i)} = G$, it suffices to take $g(t) = \exp tZ$. Supppse $i \ge 2$ and the statement holds for level $i - 1$.

Suppose $X \in \mathfrak{g}$, $Y_0 \in \mathfrak{g}_{(i)}$. By inductive hypothesis, there is a $G_{(i-1)}$-admissible curve $h : \mathbb{R} \to G_{(i-1)}$ such that $h(0) = e$, $h'(0) = Y_0$. Because h is differentiable at 0, one can modify its values on $(-\infty, 0)$ by setting $h(-t) = h(t)^{-1}$ for all $t > 0$. Then the resulting new curve is still continuous at 0. Moreover, by (1.1),

$$\frac{\mathrm{d}}{\mathrm{d}t} h(t)|_{t=0-} = -\frac{\mathrm{d}}{\mathrm{d}t} h(-t)|_{t=0+} = -\frac{\mathrm{d}}{\mathrm{d}t} h(t)^{-1}|_{t=0+} = \frac{\mathrm{d}}{\mathrm{d}t} h(t)|_{t=0+} = Y_0.$$

So the new curve is also C^1 and $h'(0) = Y_0$. Thus we may assume $h(-t) = h(t)^{-1}$ for all $t \in \mathbb{R}$.

Define a $G_{(i)}$-admissible curve

$$g(t) = (\exp X) h(t) \exp(-X) h(-t),$$

which lies in $[G, G_{(i-1)}] = G_{(i)}$. By applying (1.1) and Lemma 1.3.2,

$$\begin{aligned} g'(0) &= \left(\frac{\mathrm{d}}{\mathrm{d}t} (\exp X h(t) \exp(-X)) + \frac{\mathrm{d}}{\mathrm{d}t} h(-t) \right) \Big|_{t=0} \\ &= \mathrm{Ad}_{\exp X} Y_0 - Y_0 = (\exp(\mathrm{ad}_X) - \mathrm{Id}) Y_0 \\ &= \sum_{n=1}^{\infty} \frac{\mathrm{ad}_X^n Y_0}{n!} = \left[X, \big(\mathrm{Id} + \sum_{n=1}^{\infty} \frac{\mathrm{ad}_X^n}{(n+1)!}\big) Y_0 \right] \end{aligned}$$

By Lemma 1.4.1, $(\mathrm{Id} + \sum_{n=1}^{\infty} \frac{\mathrm{ad}_X^n}{(n+1)!}) Y_0 \in \mathfrak{g}_{(i-1)} + [\mathfrak{g}, \mathfrak{g}_{(i-1)}] \subseteq \mathfrak{g}_{(i-1)}$. Furthermore, when X is in a small neighborhood of 0, $\mathrm{ad}_X \in \mathfrak{gl}(\mathfrak{g})$ has small matrix norm, and hence so does the linear transform $\sum_{n=1}^{\infty} \frac{\mathrm{ad}_X^n}{(n+1)!}$. In consequence, $\mathrm{Id} + \sum_{n=1}^{\infty} \frac{\mathrm{ad}_X^n}{(n+1)!}$ is an invertible linear transform of $\mathfrak{g}_{(i-1)}$. So we can conclude that: for vectors $X \in \mathfrak{g}$ that are sufficiently close to 0, and all vectors $Y \in \mathfrak{g}_{(i-1)}$, there exists $Y_0 \in \mathfrak{g}_{(i-1)}$ such that the curve $g(t)$ inside the subgroup $G_{(i)}$ has value $g(0) = e$ and derivative $g'(0) = [X, Y]$ at 0. By replacing the pair (X, Y) by $(RX, R^{-1}Y)$ where $R > 1$, one can drop the assumption that X is close to 0 and assert that for all $X \in \mathfrak{g}$, $Y \in \mathfrak{g}_{(i-1)}$, there exists a $G_{(i)}$-admissible curve g in $G_{(i)}$ with $g(0) = e$, $g'(0) = [X, Y]$.

By applying (1.1) again, we can then, by taking an n-fold product where $n \leq \dim \mathfrak{g}_{(i)}$, construct an admissible curve g such that $g(0) = e$ and $g'(0)$ is any linear combination Z of vectors of the form $[X, Y]$, $X \in \mathfrak{g}$, $Y \in \mathfrak{g}_{(i-1)}$. This is what we need, since $\mathfrak{g}_{(i)}$ is by definition the collection of all such Z's. ☐

Theorem 1.5.7. *For a connected Lie group G, G is nilpotent if and only if $\mathfrak{g} = \mathrm{Lie}(G)$ is a nilpotent Lie algebra. Moreover, their steps of nilpotency are equal, and $\exp : \mathfrak{g} \to G$ is a covering map.*

Proof of the "only if" part. Suppose G is s-step nilpotent, then $G_{(s+1)}$ is trivial. Therefore, according to Lemma 1.5.6, $\mathfrak{g}_{(s+1)}$ cannot contain any non-zero vectors. So $\mathfrak{g}$ is nilpotent, and its step of nilpotency is at most s. ☐

We start the proof of the "if" part by declaring the following fact, which does not necessarily hold for general Lie groups.

Proposition 1.5.8. If the Lie algebra $\mathfrak{g}$ of a simply connected Lie group G is nilpotent, then:

(1) $\exp : \mathfrak{g} \to G$ is a diffeomorphism between $\mathfrak{g}$ and G.

(2) The Baker-Campbell-Hausdorff formula (1.5) holds for all $X, Y \in \mathfrak{g}$.

Proof. Suppose $\mathfrak{g}$ is s-step nilpotent. Then for sufficiently short vectors $X, Y \in \mathfrak{g}$, $\exp X \exp Y = \exp Z$ where $Z = Z(X, Y)$ is, by the Baker-Campbell-Hausdorff formula, a sum

$$Z(X, Y) = X + Y + \frac{1}{2}[X, Y] + \cdots$$

in which each term is a repeated Lie bracket between X and Y of at most s layers. There are only finitely many such combinations of Lie brackets, and $(X, Y) \to [X, Y]$ is a bilinear map. Therefore, $Z(X, Y)$ is a finite sum, more precisely a polynomial map from $\mathfrak{g} \times \mathfrak{g}$ to $\mathfrak{g}$ whose linear part is $X + Y$. In particular, $Z(X, Y)$ is defined for all $X, Y \in \mathfrak{g}$, not only those from a neighborhood of 0.

We claim that $\mathfrak{g}$ has a group structure, where the identity map is 0, the multiplication is given by

$$X \odot Y = Z(X, Y) \tag{1.11}$$

and inversion is given by $X \to -X$. For this, one needs to verify

$$Z(X, Z(Y, W)) = Z(Z(X, Y), W) \tag{1.12}$$

and

$$Z(0,X)=Z(X,0)=X,\;\; Z(X,-X)=0 \tag{1.13}$$

Since (1.13) follows immediately from the fact that $[X,-X]=[X,0]=[0,X]=0$, we focus on (1.12).

Notice that for X, Y, W of small norms, all intermediate expressions in (1.12) takes place on a small neighborhood U of $0\in\mathfrak{g}$ on which exp is a diffeomorphism. In this case

$$\begin{aligned}Z(X,Z(Y,W)) &= \exp^{-1}(\exp X\cdot\exp Z(Y,W)) = \exp^{-1}(\exp X\cdot\exp Y\exp W)\\ &= \exp^{-1}(\exp Z(X,Y)\cdot\exp W) = Z(Z(X,Y),W).\end{aligned}$$

So $Z(X,Z(Y,W))-Z(Z(X,Y),W)$ vanishes when (X,Y,W) lies in a certain neighborhood of the origin in $\mathfrak{g}^3$. On the other hand, $Z(X,Z(Y,W))-Z(Z(X,Y),W)$ is a polynomial map as Z is. This forces the polynomial to vanish on the entire $\mathfrak{g}^3$. So (1.12) holds for all X, Y, Z and we obtain the desired group structure on $\mathfrak{g}$.

Moreover, the Lie bracket $\mathfrak{g}$ determined by the group multiplication $\odot$ coincides with the original bracket on $\mathfrak{g}$, because $\exp:\mathfrak{g}\to G$ behaves like a group morphism on U and $D_e\exp=\mathrm{Id}$.

Notice that in the Lie group $(\mathfrak{g},\cdot)$, $\{tX\}$ is a one-parameter subgroup and $(tX)'|_{t=0}=X$. By the uniqueness in Lemma 1.1.11, the exponential map $\exp^{\odot}$ from $\mathfrak{g}$ to the Lie group $(\mathfrak{g},\odot)$ sends tX to tX for all $t\in\mathbb{R}$ and $X\in\mathfrak{g}$. In other words, $\exp^{\odot}=\mathrm{Id}$.

By Corollary 1.1.16 there is a unique Lie group morphism Ψ from $\mathfrak{g}$ to G, such that $D_e\Psi=\mathrm{Id}$. By Corollary 1.1.13, $\Psi\circ\exp^{\odot}=\exp\circ\mathrm{Id}$, or equivalently $\Psi=\exp$. Thus we can regard exp as a Lie group morphism from $(\mathfrak{g},\odot)$ and $(G,\cdot)$, with $D_e\exp=\mathrm{Id}$. So, exp is a covering map by Lemma 1.1.17, and then must be bijective because $\mathfrak{g}$ and G are both simply connected. In conclusion, $\exp:\mathfrak{g}\to G$ is a diffeomorphism. This shows part (1).

To prove part (2), notice that by the definition of (1.11)and the fact that exp is a group morphism, $\exp X\cdot\exp Y=\exp(X\odot Y)=\exp Z(X,Y)$ for all X, Y, which is the Baker-Campbell-Hausdorff formula. □

In fact, for the Baker-Campbell-Hausdorff formula to hold for all Lie algebra elements, we do not need the simply connectedness assumption on G. (Excercise 1.5.1)

Corollary 1.5.9. *If G is a simply connected Lie group whose Lie algebra $\mathfrak{g}$ is nilpotent, then for all ideals $\mathfrak{h}\subseteq\mathfrak{g}$, $H=\exp\mathfrak{h}$ is a simply connected Lie subgroup of G and* exp *is a diffeomorphism between $\mathfrak{h}$ and H.*

In particular, all connected closed subgroups of G are simply connected.

Proof. Note that $\mathfrak{h}$ is invariant under the group multiplication $\odot$ by Baker-Campbell-Hausdorff formula and the inversion $X \to -X$. Hence $\mathfrak{h}$ is a subgroup in $(\mathfrak{g}, \odot)$.

The corollary is a direct sequence from the fact that $\exp : (\mathfrak{g}, \odot) \to (G, \cdot)$ is a Lie group isomorphism. □

In order to show the "if" direction in Theorem 1.5.7, we further need the following lemma:

Lemma 1.5.10. *Assume that G is a simply connected Lie group whose Lie algebra $\mathfrak{g}$ is nilpotent. Assume $\mathfrak{h} \lhd \mathfrak{g}$ is an ideal, then $[G, \exp \mathfrak{h}] \subseteq \exp[\mathfrak{g}, \mathfrak{h}]$.*

Proof. We first show that $[G, \exp \mathfrak{h}] \subseteq \exp[\mathfrak{g}, \mathfrak{h}]$. For this, it suffices to prove that $\exp X \exp Y \exp(-X) \exp(-Y) \in \exp[\mathfrak{g}, \mathfrak{h}]$ for all $X \in \mathfrak{g}$, $Y \in \mathfrak{h}$.

By the Baker-Campbell-Hausdorff formula, this expression is given by $\exp(Z(Z(X, Y), Z(-X, -Y))$. However,

$$
\begin{aligned}
& Z(Z(Z(X,Y), Z(-X,-Y)) \\
=& Z(X + Y + \frac{1}{2}[X,Y] + \cdots, -X - Y + \frac{1}{2}[X,Y] + \cdots) \\
=& [X,Y] + [X + Y + \frac{1}{2}[X,Y] + \cdots, -X - Y + \frac{1}{2}[X,Y] + \cdots] \\
=& [X,Y] + [X + Y, -X - Y] + \cdots = [X,Y] + \cdots,
\end{aligned}
$$

where all the omitted terms are repeated Lie brackets that have at least three layers, and has at least one component that is equal to Y. Each of these terms belong to $[\mathfrak{g}, \mathfrak{h}]$ as well as $[X, Y]$. Therefore, we obtain that $\exp X \exp Y \exp(-X) \exp(-Y) \in \exp[\mathfrak{g}, \mathfrak{h}]$ □

Proof of the "if" part of Theorem 1.5.7. Suppose G is a connected Lie group whose Lie algebra $\mathfrak{g}$ is s-step nilpotent. Then, since its universal cover $\widetilde{G}$ is also a connected Lie group endowed with the same Lie algebra, by Proposition 1.5.8, $\exp : \mathfrak{g} \to \widetilde{G}$ is a diffeomorphism and indeed, by the proof of Proposition 1.5.8, an isomorphism between the Lie group $(\mathfrak{g}, \odot)$ and $\widetilde{G}$. By Lemma 1.5.10, $[\widetilde{G}, \exp \mathfrak{g}_{(i)}] \subseteq \exp \mathfrak{g}_{(i+1)}$ for all i. Because $\exp \mathfrak{g}_{(1)} = \exp \mathfrak{g} = \widetilde{G}$, it follows by induction that $\widetilde{G}_{(i)} \subseteq \exp \mathfrak{g}_{(i)}$. As $\mathfrak{g}_{(s+1)} = \{0\}$, $G_{(s+1)}$ is a trivial group. So $\widetilde{G}$ is nilpotent and its step of nilpotency is s. Lemma 1.5.3 implies that G is nilpotent with step bounded by s. This completes the proof of Theorem 1.5.7. □

The two lower central series, in Lie algebra and in Lie group, correspond to each other for simply connected nilpotent Lie groups.

Corollary 1.5.11. *For a simply connected nilpotent Lie group G, for all indices i, $G_{(i)} = \exp \mathfrak{g}_{(i)}$.*

Proof. By the proof of the "if" part above, $G_{(i)} \subseteq \exp \mathfrak{g}_{(i)}$. It suffices to show that $\exp \mathfrak{g}_{(i)} \subseteq G_{(i)}$. Because $\exp \mathfrak{g}_{(i)}$ is a connected Lie group and is generated by any open neighborhood of its identity, it suffices to show that for a sufficiently small neighborhood $U \subseteq G_{(i)}$ of e, $U \subseteq G_{(i)}$.

Note that in the recurvsive Definition 1.5.5 we can make $n = \dim \mathfrak{g}_{(i)}$ in (1.10) by adding trivial components when $n < \dim \mathfrak{g}_{(i)}$. So Lemma 1.5.6 tells us that for every $1 \leq i \leq s$, there exist constants $N = N(i)$, $M = M(i)$ and an expression $\gamma(t, Y_1, \cdots, Y_N) \in G_{(i)}$ representing a $G_{(i)}$-admissive curve that can be decomposed into a product $\exp_{tY_{h_1}} \cdots \exp tY_{h_M}$ of $\exp(tY_1), \cdots, \exp(tY_N)$, where $h_1, \cdots, h_M \in \{1, \cdots, N\}$ such that the map $L(Y_1, \cdots, Y_N) = \frac{\partial}{\partial t}\gamma(t, Y_1, \cdots, Y_N)\big|_{t=0}$ from $\mathfrak{g}^N$ to $\mathfrak{g}_{(i)}$ is surjective.

The product $P(X_1, \cdots, X_N) = \exp_{X_{h_1}} \cdots \exp X_{h_M}$ defines a differentiable map from $\mathfrak{g}^N$ to $G_{(i)}$. Observe that $L = (\nabla P)(Y_1, \cdots, Y_N)$, and thus ∇P is a surjective linear map from $\mathfrak{g}^N$ to $\mathfrak{g}_{(i)}$. By implicit function theorem, the image of P contains a small neighborhood U of the identity in $G_{(i)}$. This completes the proof. □

Example 1.5.12. The family of upper triangular unipotent $d \times d$ matrices, i.e. upper triangular matrices whose diagonal entries are all 1, form a subgroup N of $\mathrm{SL}(d, \mathbb{R})$. This group is $d-1$ step nilpotent as its Lie algebra is the nilpotent Lie algebra $\mathfrak{n}$ in Example 1.4.10.

Example 1.5.13. The $(2n+1)$ dimensional Heisenberg Lie group is $H_{2n+1} = \mathbb{R}^{2n+1}$ equipped with the group multiplication

$$(\mathbf{x}, \mathbf{y}, z)(\mathbf{x}', \mathbf{y}', z') = (\mathbf{x} + \mathbf{x}', \mathbf{y} + \mathbf{y}', z + z' + \mathbf{x} \cdot \mathbf{y}')$$

and group inversion $(\mathbf{x}, \mathbf{y}, z)^{-1} = (-\mathbf{x}, -\mathbf{y}, -z + \mathbf{x} \cdot \mathbf{y})$, where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $z \in \mathbb{R}$. Its Lie algebra is the Lie algebra $\mathfrak{h}_{2n+1}$ from Example 1.4.11. Therefore H_{2n+1} is 2-step nilpotent. The center of H_{2n+1} is the one-parameter subgroup $\{(\mathbf{0}, \mathbf{0}, z)\}$.

We proved in Corollary 1.5.9 that every connected closed subgroup H of a simply connected nilpotent Lie group G is simply connected. In fact, so is the quotient G/H if H is a normal subgroup.

Corollary 1.5.14. *If H is a connected closed normal subgroup of a simply connected nilpotent Lie group G, then G/H is a simply connected nilpotent group which is diffeomorphic to its Lie algebra $\mathfrak{g}/\mathfrak{h}$ through the exponential map.*

Proof. Because H is closed, G/H has a manifold structure and hence is a Lie group. Its Lie algebra can be naturally identified with $\mathfrak{g}/\mathfrak{h}$.

By Corollary 1.1.13 we have the commutative diagram

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\pi} & \mathfrak{g}/\mathfrak{h} \\ \downarrow{\scriptstyle \exp} & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\pi} & G/H \end{array} \tag{1.14}$$

We know that $\exp : \mathfrak{g}/\mathfrak{h} \to G/H$ is a local diffeomorphism at the identity. By this diagram, it is also surjective because $\exp : \mathfrak{g} \to G$ is a diffeomorphism. It suffices to show its injectivity.

Assume for the sake of contradiction that $Z, X \in \mathfrak{g}$ satisfies $\exp Z \in \exp X \cdot H$. We want to show that $Z - X \in \mathfrak{h}$. Because $H = \exp \mathfrak{h}$, $\exp Z = \exp X \exp Y$ for some $Y \in \mathfrak{h}$. Hence $Z = Z(X, Y)$ is given by the Baker-Campbell-Hausdorff formula. More precisely, $Z = X + Y + \cdots$, where all the omitted terms are repeated Lie bracktet involving Y at least once. Since $Y \in \mathfrak{h}$ and $\mathfrak{h}$ is an ideal (due to the normality of H), these omitted terms all belong to $\mathfrak{h}$. Thus $Z - X \in \mathfrak{h}$. □

From Lemma 1.4.8 and Theorem 1.5.7 we deduce:

Corollary 1.5.15. *If G is a connected Lie group of step s, then for every $g \in G$, $\mathrm{Ad}_g \in \mathrm{GL}(\mathfrak{g})$ is a unipotent matrix and $(\mathrm{Ad}_g - \mathrm{Id})^s = 0$.*

Proof. By Theorem 1.5.7, $g = \exp X$ for some X. By Corollary 1.1.13, $\mathrm{Ad}_g = \exp(\mathrm{ad}_X) = \sum_{n=0}^{\infty} \frac{1}{n!} \mathrm{ad}_X^n$. Because ad_X is a nilpotent matrix by Lemma 1.4.8, Ad_g is unipotent as a matrix on $\mathfrak{g}$. In fact, $\mathrm{Ad}_g - \mathrm{Id} = \mathrm{ad}_X \cdot B$ where $B = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \mathrm{ad}_X^n$ is a matrix commuting with ad_X. So $(\mathrm{Ad}_g - \mathrm{Id})^s = \mathrm{ad}_X^s B^s = 0$ as $\mathrm{ad}_X^s = 0$. Thus $\mathrm{Ad}_g - \mathrm{Id}$ is nilpotent and Ad_g is unipotent. □

Finally, we state another fact about connected closed subgroups in simply connected nilpotent Lie groups.

Lemma 1.5.16. *Let H is a closed subgroup of a simply connected nilpotent Lie group G. If H has finitely many connected components then it is connected and simply connected.*

Proof. Simply connectedness follows from connectedness so it suffices to show that there is only one connected component. Let H^0 be the connected component of H containing identity, then it is a connected closed subgroup and hence a Lie subgroup. Assume there is an element $h \in H$ that is not in H^0, we claim that $h^n \notin H^0$ for all $n \in \mathbb{Z} \backslash \{0\}$.

In fact, $h = \exp X$ for some $X \in \mathfrak{g}$ as $\exp : \mathfrak{g} \to G$ is a diffeomorphism, and $h^n = \exp(nX)$. If $h^n \in H^0$, then nX belongs to the Lie algebra $\mathfrak{h}$ of H^0 and thus so does X. This would contradict the fact that $h \notin H^0$.

It follows that $h^n H^0$ are different connected components of H for all $n \in \mathbb{Z}$ and H^0 has infinite index in H. This contradicts the assumption. The proof is completed. □

Exercises

Exercise 1.5.1. Show that for all connected nilpotent Lie group G, the Baker-Campbell-Hausdorff formula is true for all pairs of elements from $\mathfrak{g}$.

Exercise 1.5.2. Show that $\mathfrak{h}_{2n+1}$ from Example 1.4.11 is isomorphic to the Lie algebra of H_{2n+1} from Example 1.5.13. Moreover, the identity map

$$\sum_{i=1}^{n} x_i X_i + \sum_{i=1}^{n} y_i Y_i + zZ \to (x_1, \cdots, x_n, y_1, \cdots, y_n, z)$$

from $\mathfrak{h}_{2n+1}$ to H_{2n+1} is the exponential map.

Exercise 1.5.3. Let $u_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$.

(1) Prove that $t \to u_t$ is a group morphism. In particular, $u : (t, v) \to u_t v$, where $t \in \mathbb{R}$, $v \in \mathbb{R}^2$, is a left $\mathbb{R}$-action on $\mathbb{R}^2$.

(2) Prove that the semi-product group $\mathbb{R} \ltimes_u \mathbb{R}^2$ is a nilpotent Lie group.

(3) Prove that $\mathbb{R} \ltimes_u \mathbb{R}^2$ is isomorphic to the 3-dimensional Heisenberg group.

Exercise 1.5.4. Show that for all $a \in \mathrm{SL}(d, \mathbb{R})$, the group

$$\{a \in \mathrm{SL}(d, \mathbb{R}) : \lim_{n \to \infty} g^n a g^{-n} = \mathrm{Id}\}$$

is a nilpotent Lie group and describe its Lie algebra.

1.6 Mal'cev basis

From previous discussions, we know that simply connected Lie groups are diffeomorphic to their Lie algebras, and that in this case the Lie group structure and the Lie algebra structure determine each other. Which data do we need in order to completely describe these structures?

Definition 1.6.1. *A* **filtration** *of a nilpotent Lie algebra* $\mathfrak{g}$ *is a sequence of Lie subalgebras* $\mathfrak{g} = \mathfrak{g}_1 \rhd \mathfrak{g}_2 \rhd \cdots \rhd \mathfrak{g}_r = \{0\}$ *such that* $[\mathfrak{g}, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i+1}$.

In particular, each $\mathfrak{g}_i$ is an ideal of $\mathfrak{g}$. Moreover as in Lemma 1.4.3, $[\mathfrak{g}_i, \mathfrak{g}_j] \subseteq \mathfrak{g}_{i+j}$.

By definition, the lower central series $\mathfrak{g}_{(i)}$ of a nilpotent Lie algebra $\mathfrak{g}$ is a filtration. For a filtration, $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is an abelian Lie algebra; in other words, it is isomorphic to $\mathbb{R}^{d_i}$ for some d_i. Denote $m_i = \dim \mathfrak{g}_i$, then $m_i = \sum_{j=i}^r d_j$ and is decreasing in i. Write $m = m_1 = \dim \mathfrak{g}$.

Definition 1.6.2. *A* **Mal'cev basis** $\mathcal{X}$ *adapted to the filtration* $\{\mathfrak{g}_i\}$ *is a basis* $X_1, \cdots, X_m$ *of* $\mathfrak{g}$ *such that* $X_{m-m_i+1}, \cdots, X_m$ *spans* $\mathfrak{g}_i$ *for each* i.

Notice that given the filtration, a Mal'cev basis always exists.

Example 1.6.3. The basis $\{X_1, \cdots, X_n, Y_1, \cdots, Y_n, Z\}$ in Example 1.4.11, in that order, form a Mal'cev basis adapted to the lower central series filtration of the Heisenberg Lie algebra.

Suppose X_j, X_k are from the Mal'cev basis $\mathcal{X}$ and $j \leq k$, then there exists $1 \leq i \leq r$ such that $m - m_i + 1 \leq k \leq m - m_{i+1}$. Hence $X_k \in \mathfrak{g}_i$ and $[X_j, X_k] \in \mathfrak{g}_{i+1}$ because $\{\mathfrak{g}_i\}$ is a filtration. So there are constants c_{jk}^l for all $m - m_{i+1} + 1 \leq l \leq m$, such that

$$[X_j, X_k] = \sum_{l=m-m_{i+1}+1}^{m} c_{jk}^l X_l. \tag{1.15}$$

The constants c_{jk}^l, called **structural constants**, completely determines the Lie bracket on $\mathfrak{g}$, and hence, via Baker-Campbell-Hausdorff formula, the multiplication on the simply connected Lie group G associated to $\mathfrak{g}$ as well.

Once the Mal'cev basis $\mathcal{X}$ is chosen, the map $\psi : \mathbb{R}^d \to G$ given by

$$\psi(u_1, \cdots, u_m) = \exp(u_1 X_1 + \cdots + u_m X_m) \tag{1.16}$$

is a diffeomorphism by Corollary 1.5.9. The following fact is evident:

Lemma 1.6.4. *In the coordinate system (1.16), the multiplication on G is given by the polynomial map $\odot$ in (1.11). And a group morphism $\Psi : G \to G$ of the Lie group G is simply given by the linear transform $D_e\Psi$ thanks to the diagram (1.14).*

In this coordinate system we will write $\mathbf{u} = (u_1, \cdots, u_m)$ and, for $j = 1, \cdots, s$, $\mathbf{u}_j = (u_{m-m_j+1}, \cdots, u_{m-m_{j+1}})$. Then $\mathbf{u} = (\mathbf{u}_1, \cdots, \mathbf{u}_r)$.

Lemma 1.6.5. *There exist polynomial functions $\boldsymbol{\theta}_i : (\mathbb{R}^{m-m_i})^2 \to \mathbb{R}^{d_i}$ for all $1 \le i \le r$, such that $\psi(\mathbf{u})\psi(\mathbf{v}) = \psi(\mathbf{w})$ if and only if*

$$\mathbf{w}_i = \mathbf{u}_i + \mathbf{v}_i + \boldsymbol{\theta}_i(\mathbf{u}_1, \cdots, \mathbf{u}_{i-1}, \mathbf{v}_1, \cdots, \mathbf{v}_{i-1}). \tag{1.17}$$

Moreover, $\boldsymbol{\theta}_1 = 0$.

Proof. Write $U_i = \sum_{j=m-m_i+1}^{m-m_{i+1}} u_j X_j$ and similarly define V_i, W_i. By construction of ψ,

$$(\sum_{i=r}^{m} U_j) \odot (\sum_{i=1}^{r} V_i) = \sum_{i=1}^{r} W_i. \tag{1.18}$$

By Baker-Campbell-Hausdorff formula, (1.18) is equal to

$$\sum_{i=1}^{r}(U_i + V_i) + \sum(\text{bracket terms}),$$

where each bracket term has the form $b[Y_1, [Y_2, \cdots [Y_t, Y_{t+1}] \cdots]]$ with $b \in \mathbb{Q}$ and $Y_1, \cdots, Y_{t+1} \in \{U_1, \cdots, U_r, V_1, \cdots, V_r\}$. If at least one of these is from $\{U_i, \cdots, U_r, V_i, \cdots, V_r\} \subset \mathfrak{g}_i$, then the bracket term will be in $\mathfrak{g}_{i+1}$ as $[\mathfrak{g}, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i+1}$. In this case, that bracket term belongs to the span of $X_{m-m_{i+1}}, \cdots, X_m$, and therefore does not contribute to the component W_i in (1.18). Therefore, $W_i - U_i - V_i$ is a fintie sum consisting of the $\bigoplus_{j=m-m_i+1}^{m-m_{i+1}} \mathbb{R}X_j$-components of the bracket terms who only involve $U_1, \cdots, U_{r-1}, V_1, \cdots, V_{r-1}$. Because this set of vectors are linearly parametrized by $\mathbf{u}_1, \cdots, \mathbf{u}_{i-1}, \mathbf{v}_1, \cdots, \mathbf{v}_{i-1}$ and the Lie bracket is bilinear, $W_i - U_i - V_i$ is a polynomial in $\mathbf{u}_1, \cdots, \mathbf{u}_{i-1}, \mathbf{v}_1, \cdots, \mathbf{v}_{i-1}$. This is equivalent to (1.17). When $i = 1$, there are no such bracket terms as $i - 1 = 0$, and thus $\boldsymbol{\theta}_1 = 0$. □

Instead of the coordinate change ψ, we can also define another map $\phi : \mathbb{R}^d \to G$ by

$$\phi(v_1, \cdots, v_m) = \exp(v_1 X_1) \cdots \exp(v_m X_m). \tag{1.19}$$

The map is differentiable and a local diffeomorphism near 0, because it is easy to check that $D_0\phi(u_1, \cdots, u_m) = \sum_{j=1}^m u_m X_m$. Indeed, we will see that it is a global diffeomorphism. This follows from the following

Proposition 1.6.6. There is a bijective polynomial map $f : \mathbb{R}^m \to \mathbb{R}^m$ such that $\phi = \psi \circ f$. Moreover, there are polynomials $\boldsymbol{\zeta}_i, \widehat{\boldsymbol{\zeta}}_i : (\mathbb{R}^{m-m_i})^2 \to \mathbb{R}^{d_i}$, such that if $f(\mathbf{u}) = \mathbf{v}$, then

$$\mathbf{v}_i = \mathbf{u}_i + \boldsymbol{\zeta}_i(\mathbf{u}_1, \cdots, \mathbf{u}_{i-1}), \tag{1.20}$$

$$\mathbf{u}_i = \mathbf{v}_i + \widehat{\boldsymbol{\zeta}}_i(\mathbf{v}_1, \cdots, \mathbf{v}_{i-1}). \tag{1.21}$$

Moreover, $\boldsymbol{\zeta}_1 = \hat{\boldsymbol{\zeta}}_1 = 0$.

Proof. The proof is by induction in r, the step of the filtration with respect to which $\mathcal{X}$ is defined.

When $r = 1$, $\mathfrak{g}$ is abelian, and therefore $\phi = \psi$ by Baker-Campbell-Hausdorff formula. So it suffices to take $f = \mathrm{id}$ and $\boldsymbol{\zeta}_1 = \widehat{\boldsymbol{\zeta}}_1 = 0$.

Assume the $r-1$ case is known. Note that $\mathcal{X}' = \{X_{d_1+1}, \cdots, X_m\}$ is a Mal'cev basis of $\mathfrak{g}_2$. Let ψ', ϕ' be the corresponding maps defined for this basis. Then there is a bijective polynomial $f' : \mathbb{R}^{m_2} \to \mathbb{R}^{m_2}$ satisfying $\phi' = \psi' \circ f'$.

Suppose $\phi(\mathbf{u}) = \psi(\mathbf{v})$. Then

$$\psi(\mathbf{v}) = \phi(\mathbf{u}_1, 0)\phi'(\mathbf{u}'), \tag{1.22}$$

where $\mathbf{u}'$ stands for $(\mathbf{u}_2, \cdots, \mathbf{u}_r)$.

Then $\phi'(\mathbf{u}') = \psi'(\mathbf{v}') = \psi(0, \mathbf{v}')$ where $\mathbf{v}' = f'(\mathbf{u}')$. More precisely, according to the inductive hypothesis, for all $2 \le i \le r$,

$$\mathbf{v}'_i = \mathbf{u}_i + \boldsymbol{\zeta}'_i(\mathbf{u}_2, \cdots, \mathbf{u}_{i-1}), \forall 2 \le i \le r \tag{1.23}$$

$$\mathbf{u}_i = \mathbf{v}'_i + \widehat{\boldsymbol{\zeta}}'_i(\mathbf{v}'_2, \cdots, \mathbf{v}'_{i-1}), \forall 2 \le i \le r, \tag{1.24}$$

where $\boldsymbol{\zeta}'_i$ and $\widehat{\boldsymbol{\zeta}}'_i$ are polynomial maps.

Furthermore, by Baker-Campbell-Hausdroff formula,

$$\phi(\mathbf{u}_1, 0) = \exp(u_1 X_1) \cdots \exp(u_{d_1} X_{d_1}) = \exp\big(\sum_{j=1}^{d_1} u_j X_j + \boldsymbol{\eta}(\mathbf{u}_1)\big).$$

Here $\boldsymbol{\eta}$ is a finite sum of iterated Lie brackets among $u_1X_1, \cdots, u_{d_1}X_{d_1}$, and therefore belongs to $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}_2$ and is polynomial in $\mathbf{u}_1$. In other words,

$$\phi(\mathbf{u}_1, 0) = \psi(\mathbf{u}_1, \boldsymbol{\eta}(\mathbf{u}_1)). \tag{1.25}$$

Combining (1.22)-(1.25) with Lemma 1.6.5, we obtain that

$$\mathbf{v}_1 = \mathbf{u}_1 + 0 = \mathbf{u}_1 \tag{1.26}$$

and for $2 \leq i \leq r$,

$$\begin{aligned} \mathbf{v}_i =& (\boldsymbol{\eta}(\mathbf{u}_1))_i + \mathbf{v}_i' \\ & \qquad + \boldsymbol{\theta}_i\big(\mathbf{u}_1, (\boldsymbol{\eta}(\mathbf{u}_1))_2, \cdots, (\boldsymbol{\eta}(\mathbf{u}_1))_{i-1}, 0, \mathbf{v}_2', \cdots, \mathbf{v}_{i-1}'\big) \\ =& \mathbf{u}_i + \Big[(\boldsymbol{\eta}(\mathbf{u}_1))_i + \boldsymbol{\zeta}_i'(\mathbf{u}_2, \cdots, \mathbf{u}_{i-1}) \\ & \qquad + \boldsymbol{\theta}_i\big(\mathbf{u}_1, (\boldsymbol{\eta}(\mathbf{u}_1))_2, \cdots, (\boldsymbol{\eta}(\mathbf{u}_1))_{i-1}, 0, \mathbf{v}_2', \cdots, \mathbf{v}_{i-1}'\big)\Big] \end{aligned} \tag{1.27}$$

Note that the sum in square brackets is a polynomial of $\mathbf{u}_1, \cdots, \mathbf{u}_{i-1}$ because $\boldsymbol{\eta}$, $\boldsymbol{\zeta}_i'$ and $\mathbf{v}_2', \cdots \mathbf{v}_{i-1}'$ are all polynomial. This proves (1.20) for index i.

The array $(\mathbf{u}_1, \cdots, \mathbf{u}_{i-1})$ is a polynomial function of $(\mathbf{v}_1, \cdots, \mathbf{v}_{i-1})$, because (1.21) is assumed for all indices up to $i-1$. Thus the function $\boldsymbol{\zeta}_i(\mathbf{u}_1, \cdots, \mathbf{u}_{i-1})$ is a polynomial, which we denote by $-\widehat{\boldsymbol{\zeta}}_i$, in $(\mathbf{v}_1, \cdots, \mathbf{v}_{i-1})$. The equality (1.21) for i then follows from (1.20). □

Exercises

Exercise 1.6.1. Using the Mal'cev basis $\{X_1, \cdots, X_m, Y_1, \cdots, Y_m, Z\}$ of the $(2n+1)$-dimensional Heisenberg Lie algebra $\mathfrak{h}_{2n+1}$, express the multiplication rule on the corresponding Lie group H_{2n+1} in the coordinate system (1.19).

Exercise 1.6.2. Show that all simply connected 3-dimensional nilpotent Lie groups are isomorphic to either H_3 or $\mathbb{R}^3$.