Some Aspects of Elementary Lie Theory

Oskar Henriksson

Bachelor's thesis
2018:K19

Lund University

Faculty of Science
Centre for Mathematical Sciences
Mathematics

Centrum Scientiarum Mathematicarum

Abstract

In this work, we present some of the basic concepts and constructions in the theory of matrix Lie groups. For each matrix Lie group, we use the matrix exponential to construct a Lie algebra, and we then use the matrix exponential to show how different properties of the Lie group affect the Lie algebra and vice versa. In particular, we use the Baker–Campbell–Hausdorff formula to prove a one-to-one correspondence between the representations of a path-connected, simply connected matrix Lie group and the representations of its Lie algebra. The physically motivated groups $\mathbf{SO}(3)$ and $\mathbf{SU}(2)$ are used as a case study.

Throughout this work it has been my firm intention to give reference to the stated results and credit to the work of others. All theorems, propositions, lemmas and examples left unmarked are assumed to be too well-known for a reference to be given.

Populärvetenskaplig sammanfattning

Det här arbetet ger en enkel introduktion till teorin för så kallade *Lie-grupper* (uppkallade efter den norske matematikern Sophus Lie), som är en slags matematiska objekt som befinner sig i gränslandet mellan två av matematikens huvudområden: geometri och algebra.

Å ena sidan kan Lie-grupper betraktas som *släta mångfalder*, vilket betyder att de är kurvor, ytor eller högre-dimensionella analoger som är släta i bemärkelsen att vi till varje punkt kan passa en tangentlinje, ett tangentplan eller ett högre-dimensionellt tangentrum som precis ”tangerar” Lie-gruppen. Samtidigt kan vi på Lie-gruppen definiera en algebraisk *gruppoperation*, ett slags ”räknesätt” i stil med vanlig multiplikation, sådant att det dels finns en analog till den vanliga 1:an och dels en analog till vanlig division. Det som utmärker Lie-grupper är att dessa två perspektiv – det geometriska och det algebraiska – är kompatibla med varandra, på ett sådant sätt att gruppoperationen är ”oändligt deriverbar” (i en viss specifik mening). Det är konsekvenserna av denna samverkan mellan geometri och algebra som denna uppsats är tänkt att ge en introduktion till.

Ett enkelt exempel på en Lie-grupp är den vanliga enhetscirkeln i planet. Dels är cirkeln en slät kurva, och dels kan varje punkt på cirkeln representeras av en vinkel (t.ex. mätt från x-axeln), vilket möjliggör en enkel gruppoperation definierad som addition av operandernas vinklar. Ett annat exempel, som vi fokuserar extra mycket på i denna uppsats, är $\mathbf{SO}(3)$, som är mängden av alla rotationer runt en fix punkt som kan utföras i tre dimensioner.

En av anledningarna till att matematiker intresserar sig för Lie-grupper är att de kan användas för att undersöka olika typer av symmetrier hos vektorer. Läran om hur grupper på detta vis interagerar med vektorer kallas för *representationsteori*, och är intressant, bland annat eftersom många fenomen i fysiken kan beskrivas med hjälp av vektorer, och dessutom ofta innefattar någon form av fysikalisk symmetri. Exempelvis kan $\mathbf{SO}(3)$ användas för att bättre förstå rotationssymmetrin hos lösningarna till Schrödinger-ekvationen för en väteatom.

Ett viktigt verktyg i studiet av Lie-grupper är de tidigare nämnda tangentrummen. Som en del av arbetet visar vi att det i vissa fall är möjligt att ”översätta” fram och tillbaka mellan frågor om en Lie-grupp och frågor om tangentrummet vid 1:an (detta tangentrum kallas för Lie-gruppens *Lie-algebra*). Detta förenklar det matematiska arbetet avsevärt, eftersom linjer, plan och högre-dimensionella rum har en enklare struktur än Lie-grupper. Tyvärr visar sig det här angreppssättet ha begränsningar när det gäller bland annat just $\mathbf{SO}(3)$, eftersom $\mathbf{SO}(3)$ inte är vad man kallar för *enkelt sammanhängande* utan har ”topologiska hålrum” i sig. I slutet av arbetet visar vi dock att detta problem går att komma runt, i vart fall när det gäller $\mathbf{SO}(3)$, genom att undersöka en besläktad Lie-grupp, som kallas för $\mathbf{SU}(2)$, som ”täcker över” $\mathbf{SO}(3)$ på ett sådant sätt att de problematiska ”hålrummen” försvinner.

Acknowledgement

First and foremost, I wish to thank my advisor, Prof. Arne Meurman, for his support and guidance during my work on this thesis, and for his many useful comments on the manuscript. I would also like to thank my fellow students and friends at Lund University and UC San Diego for interesting conversations throughout my undergraduate studies, and for constantly reminding me of how fun, beautiful and exciting mathematics is.

Oskar Henriksson

Contents

Introduction

A *Lie group* (named after Sophus Lie, 1842–1899) is a mathematical object that possesses both a geometric *smooth manifold structure* and an algebraic *group structure*. These two structures are required to be compatible with each other, in the sense that the group multiplication and group inversion are both *smooth maps*. A familiar example is the unit circle S^1 in the complex plane, which on the one hand is a one-dimensional smooth manifold (i.e. a smooth curve) and on the other hand is a group under complex multiplication.

In this work, we present some of the most elementary aspects of Lie theory, at a level accessible for most undergraduate students with basic knowledge in abstract algebra and topology.

In Chapter 1, we first give a short introduction to manifold theory, leading up to the formal definition of a Lie group. The chapter concludes with a carefully worked-out example, showing that S^1 admits a Lie group structure.

In Chapter 2, we restrict our attention to so-called *matrix Lie groups*, which are Lie groups that can be realized as groups of complex square matrices. The self-contained theory for such Lie groups was first formulated in [vN29], and is useful since many of the Lie groups that show up in applications are of matrix type. Examples of this include $\mathbf{SO}(3)$, $\mathbf{SU}(2)$ and the Heisenberg group, which play prominent roles in quantum mechanics [Hal13].

Throughout Chapter 2, we develop concepts from the more general theory of Lie groups (see for example [Bt85] and [Kna02]) in the context of matrix Lie groups. One of the main themes will be the idea that questions about matrix Lie groups oftentimes can be translated into simpler questions about the tangent space at the identity element—the *Lie algebra* of the group—which is a vector space and thus lends itself to powerful tools from linear algebra. To prove this, we employ the matrix exponential and logarithm. As part of the chapter, we also prove that every matrix Lie group can be given the structure of a Lie group.

In Chapter 3, we use this idea of a Lie group–Lie algebra correspondence to study the *representations* of matrix Lie groups, i.e. the ways in which they act continuously and linearly on vector spaces. We end the chapter by giving a classic example of a non-matrix Lie group.

Chapter 4 concludes the thesis with a case study, where some of the concepts discussed in the thesis are applied to $\mathbf{SO}(3)$ and $\mathbf{SU}(2)$.

The material in this work is mostly based on [Hal15], [Sti08] and [Kna02].

Chapter 1

Preliminaries from differential geometry

Our main objective in this chapter will be to introduce the formal definition of a Lie group. To do that, we will need a couple of definitions and results from manifold theory, which we here present without proofs. For a more thorough introduction to manifolds, we recommend [Lee13] and [Gud18].

Definition 1.1. Let (M, τ) be a topological Hausdorff space with countable basis, and let $m \in \mathbb{Z}_0^+$. Then (M, τ) is said to be a ***topological manifold*** of dimension m, if there for each $p \in M$ exists an open set $U \subseteq M$ with $p \in U$, an open set $V \subseteq \mathbb{R}^m$ and a homeomorphism $x\colon U \to V$. We call a pair (U, x) of this form a ***chart*** (or ***local coordinates***) on (M, τ).

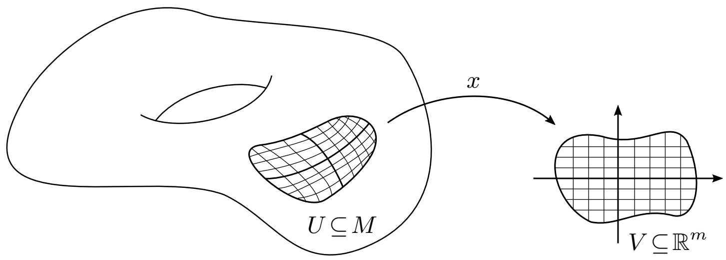

Figure 1.1: A topological manifold M with a chart (U, x).

The definition tells us that for each point $p \in M$, we can introduce an m-dimensional coordinate system in some open neighbourhood U of p, where the coordinates for each point are given by the components of the map $x\colon U \to V \subseteq \mathbb{R}^m$. Note that when two such local coordinate systems, say (U_α, x_α) and (U_β, x_β), overlap, we can "translate" between them using a ***transition map*** (or a ***change of coordinates***) given by

$$x_\beta \circ x_\alpha^{-1}\Big|_{x_\alpha(U_\alpha \cap U_\beta)} \colon x_\alpha(U_\alpha \cap U_\beta) \subseteq \mathbb{R}^m \to x_\beta(U_\alpha \cap U_\beta) \subseteq \mathbb{R}^m \,.$$

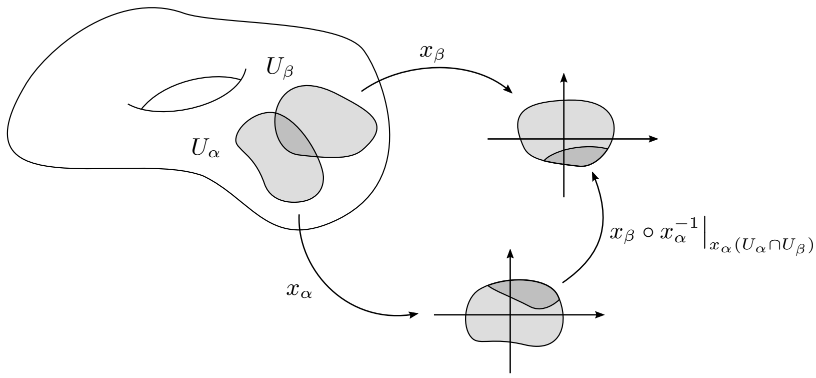

Figure 1.2: A transition map between two charts (U_α, x_α) and (U_β, x_β).

Since x_α^{-1} and x_β are both continuous, it is clear that such a change of coordinates is always continuous. For a *smooth* manifold, we will require that there exists a collection of charts that together cover the whole manifold, in such a way that all changes of coordinates in the overlaps are not only continuous, but smooth, in the usual sense of multivariable calculus.

Definition 1.2. Let (M, τ) be an m-dimensional topological manifold. A ***smooth atlas*** on (M, τ) is a collection $\mathcal{A} = \{(U_\alpha, x_\alpha)\}_{\alpha \in J}$ of charts on (M, τ) such that (i) it ***covers*** all of M, meaning that $\bigcup_{\alpha \in J} U_\alpha = M$, and (ii) the charts are ***smoothly compatible***, meaning that all transition maps are smooth. A smooth atlas $\widehat{\mathcal{A}}$ on (M, τ) is called a ***smooth structure*** if it is not properly contained in any other smooth atlas. A triple $(M, \tau, \widehat{\mathcal{A}})$, where (M, τ) is a topological manifold and $\widehat{\mathcal{A}}$ is a smooth structure on (M, τ), is called a ***smooth manifold***.

We will generally not need (or want) to write down the full smooth structure of a smooth manifold; usually it will be enough to consider an atlas. This is allowed by the useful fact that for any smooth atlas $\mathcal{A}$ on a manifold M, there exists a unique smooth structure $\widehat{\mathcal{A}}$ such that $\mathcal{A} \subseteq \widehat{\mathcal{A}}$.

Example 1.3. We can turn $\mathbb{R}^n$ into a smooth manifold by equipping it with its standard topology and the smooth structure determined by the atlas $\mathcal{A} = \{(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})\}$. Unless stated otherwise, we will always assume that $\mathbb{R}^n$ is equipped with this smooth structure.

We now go on to define what it means for a map between smooth manifolds to be *smooth*. The idea will be to use the charts from the smooth structure to "translate" maps between the manifolds into maps between the usual Euclidean spaces, and use the already well-established notion of differentiability that we have there.

Definition 1.4. A continuous map $\Phi\colon M \to N$ between smooth manifolds $(M, \tau_M, \widehat{\mathcal{A}}_M)$ and $(N, \tau_N, \widehat{\mathcal{A}}_N)$ is said to be ***smooth*** if for each $(U, x) \in \widehat{\mathcal{A}}_M$ and each $(V, y) \in \widehat{\mathcal{A}}_N$, the ***coordinate representation***

$$y \circ \Phi \circ x^{-1}\Big|_{x(U\cap\Phi^{-1}(V))} : x(U \cap \Phi^{-1}(V)) \subseteq \mathbb{R}^m \to \mathbb{R}^n$$

of Φ is smooth. If furthermore Φ is invertible with smooth inverse, then Φ is said to be a ***diffeomorphism*** between $(M, \tau_M, \widehat{\mathcal{A}}_M)$ and $(N, \tau_N, \widehat{\mathcal{A}}_N)$.

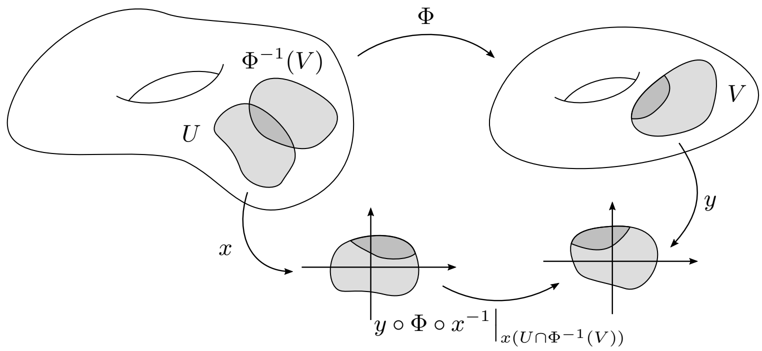

Figure 1.3: A coordinate representation of a map Φ via two charts (U, x) and (V, y).

A useful fact is that it is sufficient to check differentiability for the coordinate representations associated to the charts from some smooth atlases $\mathcal{A}_M \subseteq \widehat{\mathcal{A}}_M$ and $\mathcal{A}_N \subseteq \widehat{\mathcal{A}}_N$, rather than all possible combinations of charts from the full smooth structures. This makes it easy to see that for a map $\Phi\colon \mathbb{R}^m \to \mathbb{R}^n$, the notion of smoothness from Definition 1.4 coincides with the notion of smoothness from multivariable calculus.

Next we define a simple notion of a submanifold.

Definition 1.5. Let $(M, \tau_M, \widehat{\mathcal{A}}_M)$ be an m-dimensional smooth manifold. A subset $N \subseteq M$ is said to be an n-dimensional ***submanifold*** of M (where $n \leqslant m$), if for each $p \in N$ there exists a ***slice chart*** of N in M, i.e. a chart $(U_p, x_p) \in \widehat{\mathcal{A}}_M$ with $p \in U_p$ such that

$$x_p(U_p \cap N) = x_p(U_p) \cap (\mathbb{R}^n \times \{0\}) \,.$$

It can be shown that a submanifold N satisfying the criterion above is an n-dimensional smooth manifold in and of itself, if equipped with the subspace topology and the smooth structure associated with the atlas

$$\mathcal{A}_N = \Big\{ \Big(U_p \cap N, (\pi \circ x_p)|_{U_p \cap N} \Big) : p \in N \Big\} \,,$$

where $\pi\colon \mathbb{R}^n \times \mathbb{R}^{m-n} \to \mathbb{R}^n$ denotes the natural projection onto the first factor. The so-obtained smooth structure on N turns out to be independent of the choice of the slice charts (U_p, x_p) in the construction above, and is called the ***induced structure*** on N in M.

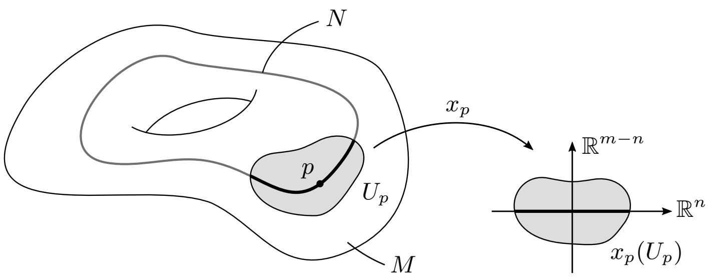

Figure 1.4: A slice chart (U_p, x_p) of a submanifold N in M.

Theorem 1.6. *Let $(M_1, \tau_1, \widehat{\mathcal{A}}_1)$ and $(M_2, \tau_2, \widehat{\mathcal{A}}_2)$ be smooth manifolds, and let $N_1 \subseteq M_1$ and $N_2 \subseteq M_2$ be submanifolds. If $\Phi\colon M_1 \to M_2$ is a smooth map, such that $\Phi(N_1) \subseteq N_2$, then $\Phi|_{N_1}\colon N_1 \to N_2$ is smooth as well.*

Next, we define what we mean by the product of two manifolds.

Proposition 1.7. *Let $(M, \tau_M, \widehat{\mathcal{A}}_M)$ be an m-dimensional smooth manifold, with smooth structure $\widehat{\mathcal{A}}_M = \{(U_\alpha, x_\alpha)\}_{\alpha\in I}$, and let $(N, \tau_N, \widehat{\mathcal{A}}_N)$ be an n-dimensional smooth manifold with smooth structure $\widehat{\mathcal{A}}_N = \{(V_\gamma, y_\gamma)\}_{\gamma\in J}$. Then the product space $(M \times N, \tau_{M\times N})$ (where $\tau_{M\times N}$ is the product topology) is a topological manifold of dimension $m + n$, and the collection*

$$\mathcal{A} = \{(U_\alpha \times V_\gamma, x_\alpha \times y_\gamma)\}_{(\alpha,\gamma)\in I\times J}$$

is a smooth atlas on $(M \times N, \tau_{M\times N})$, so that $(M \times N, \tau_{M\times N}, \widehat{\mathcal{A}})$ is a smooth manifold (called the **product manifold** *of M and N).*

We are now finally ready for the formal definition of a Lie group.

Definition 1.8. A ***Lie group*** is a tuple $(G, \cdot, \tau, \widehat{\mathcal{A}})$, where $(G, \cdot)$ is a group and $(G, \tau, \widehat{\mathcal{A}})$ is a smooth manifold, such that both multiplication $\mu\colon G \times G \to G$, given by $\mu(p, q) = p \cdot q$, and inversion $i\colon G \to G$, given by $i(p) = p^{-1}$, are smooth maps. Here, $G \times G$ denotes the product manifold.

Example 1.9. One of the simplest examples of a Lie group is the usual Euclidean space $\mathbb{R}^n$, equipped with the usual component-wise addition, the standard topology and the standard smooth structure from Example 1.3. It is then easily verified that $(x, y) \mapsto x + y$ and $x \mapsto -x$ become smooth maps.

Example 1.10. The unit circle can also be endowed with a Lie group structure. We start with the set $S^1 = \{e^{it} : t \in \mathbb{R}\} \subseteq \mathbb{C}$, and equip it with the usual complex multiplication $\cdot$, so that $e^{it} \cdot e^{is} = e^{i(t+s)}$. This clearly gives a group operation on S^1. We furthermore equip S^1 with the subspace topology in $\mathbb{C}$, and to obtain a smooth structure, we consider the set

$$\mathcal{A} = \{\, (U, x)\, ,\ (V, y)\, \}\, ,$$

where $U = S^1 \setminus \{1\}$, $V = S^1 \setminus \{-1\}$ and the chart maps are

$$\begin{aligned} x\colon S^1 \setminus \{1\} &\to (0, 2\pi)\,, \qquad & e^{it} &\mapsto t\,,\ \text{ for } t \in (0, 2\pi) \\ y\colon S^1 \setminus \{-1\} &\to (0, 2\pi)\,, \qquad & e^{\pi+t} &\mapsto t\,,\ \text{ for } t \in (0, 2\pi)\,. \end{aligned}$$

We claim that $\mathcal{A}$ is a smooth atlas on S^1. It is clear that both U and V are open and cover all of S^1, and that both x and y are homeomorphisms. It is furthermore easy to verify that the transition maps are smooth. For example, $x(U \cap V) = (0, \pi) \cup (\pi, 2\pi)$ and we see geometrically that

$$(y \circ x^{-1})(t) = \begin{cases} t + \pi & \text{for } t \in (0, \pi) \\ t - \pi & \text{for } t \in (\pi, 2\pi)\,, \end{cases}$$

from which we conclude that the transition map $y \circ x^{-1}|_{x(U \cap V)}$ is smooth. We conclude that $\mathcal{A}$ gives rise to a smooth structure on S^1. Finally we check that this smooth structure is compatible with the group structure. We first show that inversion

$$i\colon S^1 \to S^1\,, \qquad e^{it} \mapsto e^{-it}$$

is smooth, by showing that all of the four possible coordinate representations of i via $\widehat{\mathcal{A}}$ are smooth. As an example, pick (U, x) as a chart for the domain and (V, y) as a chart for the codomain. The domain of the corresponding coordinate representation is $x(U \cap i^{-1}(V)) = x(U \cap V) = (0, \pi) \cup (\pi, 2\pi)$, and we have

$$(y \circ i \circ x^{-1})(t) = \begin{cases} \pi - t & \text{for } t \in (0, \pi) \\ 3\pi - t & \text{for } t \in (\pi, 2\pi)\,, \end{cases}$$

so $(y \circ i \circ x^{-1})|_{x(U \cap i^{-1}(V))}$ is clearly smooth. The other coordinate representations of i can be shown to be smooth in a similar fashion. To show that the multiplication

$$\mu\colon S^1 \times S^1 \to S^1\,, \qquad (e^{is}, e^{it}) \mapsto e^{i(s+t)}$$

is smooth, we first note that the smooth structure on the product manifold $S^1 \times S^1$ contains the smooth atlas

$$\mathcal{A}_{S^1 \times S^1} = \{(U \times U, x \times x), (U \times V, x \times y), (V \times U, y \times x), (V \times V, y \times y)\}\,.$$

Thus, we only need to check eight coordinate representations of μ. One of them is obtained by choosing $(U \times U, x \times x)$ as a chart for $S^1 \times S^1$, and (V, y) as a chart for S^1. It is easy to see that

$$(x \times x)\Big((U \times U) \cap \mu^{-1}(V)\Big) = \Big\{(s, t) \in (0, 2\pi) \times (0, 2\pi) : s + t \notin \{\pi, 3\pi\}\Big\}\,,$$

and that for $(s,t) \in (x \times x)\left((U \times U) \cap \mu^{-1}(V)\right)$, it holds that

$$\left(y \circ \mu \circ (x \times x)^{-1}\right)(s,t) = \begin{cases} s+t+\pi & \text{if } s+t \in (0,\pi) \\ s+t-\pi & \text{if } s+t \in (\pi,3\pi) \\ s+t-3\pi & \text{if } s+t \in (3\pi,4\pi)\,, \end{cases}$$

i.e. $y \circ \mu \circ (x \times x)^{-1}|_{(x \times x)((U \times U) \cap \mu^{-1}(V))}$ is smooth. That the other coordinate representations of μ are smooth is shown similarly. We conclude that S^1, equipped with the group operation, the topology and the smooth structure described above, is indeed a Lie group.

Remark 1.11. In Section 2.4 we will describe a more general approach, by which S^1 (as well as a wider class of groups which we will call *matrix Lie groups*) can be equipped with a smooth structure compatible with the group operation. Just as in Example 1.10, the idea in Section 2.4 will be to use the exponential map (which in the next chapter will be extended to complex square matrices) to obtain local homeomorphisms between the group and Euclidean space.

Chapter 2

Matrix Lie groups and their Lie algebras

2.1 Definitions and examples

Our main focus in this work will be Lie groups that can be realized as groups of complex invertible matrices. Our first order of business will be to introduce an inner product, a norm and a topology on the complex vector space of complex $n \times n$ matrices, denoted $\mathbb{C}^{n\times n}$. We will use the ***Frobenius inner product*** $\langle \cdot, \cdot \rangle \colon \mathbb{C}^{n\times n} \to \mathbb{C}$ defined by

$$\langle Z, W \rangle = \operatorname{trace}(Z^* W) = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{z}_{ij} w_{ij} \,,$$

for complex matrices $Z = (z_{ij})$ and $W = (w_{ij})$. The corresponding induced norm $\|\cdot\| \colon \mathbb{C}^{n\times n} \to \mathbb{R}$ is called the ***Hilbert–Schmidt norm***, and is given by

$$\|Z\| = \sqrt{\langle Z, Z \rangle} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |z_{ij}|^2} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \left(a_{ij}^2 + b_{ij}^2 \right)} \,,$$

where $Z = (z_{ij}) = (a_{ij} + b_{ij} i)$ is a complex matrix. The induced topology will be considered the standard topology on $\mathbb{C}^{n\times n}$ throughout this thesis. Note that the same topology would have been obtained by any other choice of norm on $\mathbb{C}^{n\times n}$, since all norms on a finite-dimensional complex vector space are equivalent.

An alternative way to think about the Hilbert–Schmidt norm, is to identify $\mathbb{C}^{n\times n}$ with $\mathbb{C}^{n^2}$ (e.g. by stacking the columns of a matrix on top of each other) and then identify $\mathbb{C}^{n^2}$ with $\mathbb{R}^{2n^2}$ by replacing each complex entry with a 2×1 block consisting of its real and imaginary part. Then the Hilbert–Schmidt norm is just the usual Euclidean norm on $\mathbb{R}^{2n^2}$. In the 2×2 case, this corresponds to the following identifications:

$$\begin{pmatrix} a+bi & e+fi \\ c+di & g+hi \end{pmatrix} \rightsquigarrow \begin{pmatrix} a+bi \\ c+di \\ e+fi \\ g+hi \end{pmatrix} \rightsquigarrow \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{pmatrix}$$

Using this (homeomorphic) identification $\mathbb{C}^{n\times n} \cong \mathbb{R}^{2n^2}$, we can equip $\mathbb{C}^{n\times n}$ with a smooth structure, by simply letting it inherit the standard structure on $\mathbb{R}^{2n^2}$ from Example 1.3. This turns $\mathbb{C}^{n\times n}$ into a $2n^2$-dimensional smooth manifold.

As a final preparation, we recall from group theory that the ***general linear group*** over a field $\mathbb{F}$ is defined by

$$\mathbf{GL}_n(\mathbb{F}) = \{A \in \mathbb{F}^{n\times n} : \det(A) \neq 0\}\,,$$

whereas the ***special linear group*** over a field $\mathbb{F}$ is given by

$$\mathbf{SL}_n(\mathbb{F}) = \{A \in \mathbb{F}^{n\times n} : \det(A) = 1\}\,.$$

We are now ready to define the scope of this thesis.

Definition 2.1. A ***matrix Lie group*** is a subgroup $G \subseteq \mathbf{GL}_n(\mathbb{C})$ for some $n \in \mathbb{Z}^+$, such that G is closed in $\mathbf{GL}_n(\mathbb{C})$ (with respect to the subspace topology induced by $\mathbb{C}^{n\times n}$).

Put differently, a subgroup $G \subseteq \mathbf{GL}_n(\mathbb{C})$ is a matrix Lie group if and only if the limit of any convergent sequence in G either belongs to G or is singular.

Remark 2.2. It is natural to ask how this definition relates to the geometric definition of a general Lie group given in Chapter 1. It will turn out (see Sections 2.4 and 3.2 for precise statements) that every matrix Lie group can be given the structure of a Lie group, whereas not every Lie group can be realized as a matrix Lie group.

We will now look at a few examples of matrix Lie groups.

Example 2.3. The simplest example of a matrix Lie group is of course $\mathbf{GL}_n(\mathbb{C})$ itself, as it is clearly a subgroup of itself and closed in itself with respect to the subspace topology.

Example 2.4. Next we note that $\mathbf{GL}_n(\mathbb{R})$ is a matrix Lie group. It is clearly a subgroup of $\mathbf{GL}_n(\mathbb{C})$. To show that it is closed, we note that $\mathbb{R}^{n\times n} \subseteq \mathbb{C}^{n\times n}$ is closed. Indeed, let $(A_k)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}^{n\times n}$ that converges to some $A \in \mathbb{C}^{n\times n}$. Clearly all the entries of A must be real; if not, $A_k - A$ would have a constant nonzero component for all $k \in \mathbb{Z}^+$ when seen as an element of $\mathbb{R}^{2n^2}$, which would contradict the fact that $\|A_k - A\| \to 0$ as $k \to \infty$. As a consequence, if $(A_k)_{k=1}^{\infty}$ is a sequence in $\mathbf{GL}_n(\mathbb{R})$ that converges to some $A \in \mathbf{GL}_n(\mathbb{C})$, we must have $A \in \mathbf{GL}_n(\mathbb{R}) = \mathbf{GL}_n(\mathbb{C}) \cap \mathbb{R}^{n\times n}$.

Example 2.5. The real and complex special linear groups, $\mathbf{SL}_n(\mathbb{C})$ and $\mathbf{SL}_n(\mathbb{R})$, are also matrix Lie groups. Since all their elements have non-zero determinant, it is clear that they are subsets of $\mathbf{GL}_n(\mathbb{C})$. Since $\det(AB) = \det(A)\det(B)$, it is also clear that they both are closed under multiplication and inverses, and thus are subgroups of $\mathbf{GL}_n(\mathbb{C})$. To show that $\mathbf{SL}_n(\mathbb{C})$ is closed in $\mathbf{GL}_n(\mathbb{C})$, we note that the map $\det\colon \mathbb{C}^{n\times n} \to \mathbb{C}$ is continuous. Together with the fact that $\{1\}$ is closed in $\mathbb{C}$, this gives that $\mathbf{SL}_n(\mathbb{C}) = \det^{-1}(\{1\})$ is closed in the whole space $\mathbb{C}^{n\times n}$ and thus also in $\mathbf{GL}_n(\mathbb{C})$. That $\mathbf{SL}_n(\mathbb{R})$ is closed follows from the fact that $\mathbf{SL}_n(\mathbb{R}) = \mathbf{SL}_n(\mathbb{C}) \cap \mathbf{GL}_n(\mathbb{R})$.

Example 2.6. The ***unitary group*** in dimension n is defined by

$$\mathbf{U}(n) = \{A \in \mathbb{C}^{n\times n} : A^*A = I\}\,.$$

For any $A \in \mathbf{U}(n)$, it holds that

$$1 = \det(A^*A) = \det(\overline{A^\top})\det(A) = \overline{\det(A)}\det(A)\,,$$

and thus, $|\det(A)| = 1$. This shows that $\mathbf{U}(n) \subseteq \mathbf{GL}_n(\mathbb{C})$. Furthermore, if $A, B \in \mathbf{U}(n)$, then $AB \in \mathbf{U}(n)$ since

$$(AB)^*AB = B^*A^*AB = B^*B = I\,.$$

Since $A^*A = I$ implies $AA^* = I$, it is also clear that for $A \in \mathbf{U}(n)$ it holds that $A^{-1} = A^* \in \mathbf{U}(n)$. Hence, $\mathbf{U}(n)$ is a subgroup of $\mathbf{GL}_n(\mathbb{C})$. Finally, note that the map $f\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ defined by $f(A) = A^*A$ is continuous. Since $\{I\}$ is closed in $\mathbb{C}^{n\times n}$ we conclude that $\mathbf{U}(n) = f^{-1}(\{I\})$ is closed in $\mathbb{C}^{n\times n}$ and hence also closed in $\mathbf{GL}_n(\mathbb{C})$. Note that we can also view S^1 as a matrix Lie group, since it is isomorphic to $\mathbf{U}(1)$, both as a group and as a topological space.

Example 2.7. The ***orthogonal group*** in dimension n is defined by

$$\mathbf{O}(n) = \{A \in \mathbb{C}^{n\times n} : A^\top A = I\}\,.$$

Showing that $\mathbf{O}(n)$ is a subgroup of $\mathbf{GL}_n(\mathbb{C})$ is analogous to the argument we just made for $\mathbf{U}(n)$. Furthermore, it is easy to see that $\mathbf{O}(n) = \mathbf{U}(n) \cap \mathbb{R}^{n\times n}$, so that $\mathbf{O}(n)$ is closed in $\mathbb{C}^{n\times n}$, and therefore also closed in $\mathbf{GL}_n(\mathbb{C})$.

Example 2.8. The ***special unitary group*** and the ***special orthogonal group*** in dimension n are defined by

$$\mathbf{SU}(n) = \mathbf{SL}_n(\mathbb{C}) \cap \mathbf{U}(n) \ \text{ and } \ \mathbf{SO}(n) = \mathbf{SL}_n(\mathbb{R}) \cap \mathbf{O}(n)\,,$$

respectively. They are easily seen to be matrix Lie groups, since intersections of closed subgroups are closed subgroups themselves.

Example 2.9. The ***Heisenberg group*** is defined to be

$$H = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} : a, b, c \in \mathbb{R} \right\}.$$

By observing that

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a' & b' \\ 0 & 1 & c' \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a + a' & b + b' + ac' \\ 0 & 1 & c + c' \\ 0 & 0 & 1 \end{pmatrix},$$

and that

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -a & ac - b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}$$

we conclude that H is a subgroup of $\mathbf{GL}_3(\mathbb{R})$. To show that H is a closed in $\mathbf{GL}_3(\mathbb{R})$ (and therefore also closed in $\mathbf{GL}_3(\mathbb{C})$) we let $(A_k)_{k=1}^{\infty}$ be a sequence in H that converges to some $A \in \mathbf{GL}_3(\mathbb{R})$. If A is not upper triangular, with ones on the diagonal, then $A_k - A$ would have some constant non-zero entry for all $k \in \mathbb{Z}^+$, which would contradict the fact that $\|A_k - A\| \to 0$ as $k \to \infty$.

Example 2.10. The torus $T^2 = S^1 \times S^1$ can be realized as a matrix Lie group by setting

$$T^2 = \left\{ \begin{pmatrix} e^{is} & 0 \\ 0 & e^{it} \end{pmatrix} : s, t \in \mathbb{R} \right\}.$$

Clearly, T^2 is a subgroup of $\mathbf{U}(2)$, and by an argument similar to that in the previous example, T^2 must be closed in $\mathbf{U}(2)$. Hence, T^2 is a closed subgroup of $\mathbf{GL}_2(\mathbb{C})$.

Remark 2.11. It is worth noting that not every matrix group is a matrix *Lie* group, or in other words: not every subgroup of $\mathbf{GL}_n(\mathbb{C})$ is closed. One example is the ***skew line on a torus***, given by

$$G = \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{i\alpha t} \end{pmatrix} : t \in \mathbb{R} \right\}$$

for some fixed $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, which clearly is a subgroup of $\mathbf{GL}_2(\mathbb{C})$. However, G is not closed in $\mathbf{GL}_2(\mathbb{C})$. To see this, we note that $-I \in \mathbf{GL}_2(\mathbb{C})$ and that $-I \notin G$, since $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. However, α can be arbitrarily well approximated by rational numbers, and it thus follows from elementary number theory that for appropriately chosen odd integers k, $\pi\alpha k$ can be made to be sufficiently close to an odd multiple of π, so that $e^{i\pi\alpha k}$ becomes arbitrarily close to -1, while $e^{i\pi k} = -1$. Hence, there exists a sequence in G that converges to $-I$.

2.2 The matrix exponential and logarithm

To translate between a matrix Lie group and its so-called Lie algebra (which we will define momentarily), we will need a notion of an exponential map. The following definition, inspired by the Taylor series for e^x in the one-variable case, will turn out to be exactly what we need.

Definition 2.12. The ***matrix exponential*** for $n \times n$ matrices is the map $\exp\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ defined by

$$\exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!}\,. \tag{2.1}$$

For this definition to make sense, we need to show that the power series (2.1) actually converges for all $X \in \mathbb{C}^{n\times n}$. To do that, we first need the following submultiplicative property of the Hilbert–Schmidt norm.

Proposition 2.13. *For all $X, Y \in \mathbb{C}^{n\times n}$, $\|XY\| \leqslant \|X\|\|Y\|$.*

Proof. The triangle inequality combined with Cauchy–Schwarz inequality gives that for any $i, j \in \{1, \ldots, n\}$,

$$\begin{aligned} |(XY)_{ij}|^2 &= \left|\sum_{k=1}^{n} X_{ik}Y_{kj}\right|^2 \leqslant \left(\sum_{k=1}^{n} |X_{ik}Y_{kj}|\right)^2 \\ &= \left(\sum_{k=1}^{n} |X_{ik}||Y_{kj}|\right)^2 \leqslant \left(\sum_{k=1}^{n} |X_{ik}|^2\right)\left(\sum_{l=1}^{n} |Y_{lj}|^2\right). \end{aligned}$$

Summing over all $i, j \in \{1, \ldots, n\}$ then gives

$$\begin{aligned} \|XY\|^2 = \sum_{i,j=1}^{n} |(XY)_{ij}|^2 &\leqslant \sum_{i,j=1}^{n} \left(\left(\sum_{k=1}^{n} |X_{ik}|^2\right)\left(\sum_{l=1}^{n} |Y_{lj}|^2\right)\right) \\ &= \left(\sum_{i,k=1}^{n} |X_{ik}|^2\right)\left(\sum_{l,j=1}^{n} |Y_{lj}|^2\right) = \|X\|^2\|Y\|^2\,. \end{aligned}$$ □

Theorem 2.14. *The power series $\sum_{k=0}^{\infty} X^k/k!$ converges absolutely for all $X \in \mathbb{C}^{n\times n}$.*

Proof. As a consequence of Proposition 2.13, we have that $\|X^k\| \leqslant \|X\|^k$ for all $k \in \mathbb{Z}^+$ (but not necessarily $k = 0$). This gives

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{\|X^k\|}{k!} &= \|I\| + \sum_{k=1}^{\infty} \frac{\|X^k\|}{k!} \leqslant \|I\| + \sum_{k=1}^{\infty} \frac{\|X\|^k}{k!} \\ &= \sqrt{n} + \sum_{k=1}^{\infty} \frac{\|X\|^k}{k!} = (\sqrt{n} - 1) + e^{\|X\|}\,, \end{aligned}$$

and absolute convergence follows. □

Proposition 2.15. *The matrix exponential is a continuous map.*

Proof. We will use the Weierstrass M-test, combined with the uniform limit theorem. Fix $R > 0$ and form the set $D = \{X \in \mathbb{C}^{n\times n} : \|X\| < R\}$. For each $k \in \mathbb{Z}_0^+$, let $f_k\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ be defined by $f_k(X) = X^k/k!$, and set $M_0 = \sqrt{n}$ and $M_k = R^k/k!$ for $k \in \mathbb{Z}^+$. Then for any $X \in D$, $\|f_0(X)\| = M_0$, and by Proposition 2.13, we have $\|f_k(X)\| \leqslant \|X\|^k/k! < M_k$ for every $k \in \mathbb{Z}^+$. Also note that $\sum_{k=0}^{\infty} M_k = \sqrt{n} - 1 + e^R < \infty$, so by the Weierstrass M-test, $\exp = \sum_{k=0}^{\infty} f_k$ is uniformly convergent on D. By the uniform limit theorem, this implies that exp is continuous on D, and since $R > 0$ was chosen arbitrarily (and continuity is a local property), we conclude that exp is continuous on all of $\mathbb{C}^{n\times n}$. □

Proposition 2.16. *For all $X, Y \in \mathbb{C}^{n\times n}$ the following statements hold:*

(i) $\exp(0) = I$.

(ii) $[\exp(X)]^* = \exp(X^*)$.

(iii) *If* $XY = YX$, *then* $\exp(X+Y) = \exp(X)\exp(Y)$.

(iv) *The matrix* $\exp(X)$ *is invertible, with* $[\exp(X)]^{-1} = \exp(-X)$.

(v) *If* $C \in \mathbf{GL}_n(\mathbb{C})$, *then* $\exp(CXC^{-1}) = C\,\exp(X)\,C^{-1}$.

Proof.

(i) This follows from direct computation.

(ii) Taking the adjoint is linear and continuous, so for any $X \in \mathbb{C}^{n\times n}$ we have

$$[\exp(X)]^* = \left(\lim_{N\to\infty}\sum_{k=0}^{N}\frac{X^k}{k!}\right)^* = \lim_{N\to\infty}\left(\sum_{k=0}^{N}\frac{X^k}{k!}\right)^* = \lim_{N\to\infty}\sum_{k=0}^{N}\frac{(X^k)^*}{k!} = \lim_{N\to\infty}\sum_{k=0}^{N}\frac{(X^*)^k}{k!} = \exp(X^*)\,.$$

(iii) Due to the absolute convergence, the product $\exp(X)\exp(Y)$ can be evaluated as a Cauchy product, which gives

$$\exp(X)\exp(Y) = \sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{X^k}{k!}\frac{Y^{m-k}}{(m-k)!} = \sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}X^kY^{m-k} = \sum_{m=0}^{\infty}\frac{(X+Y)^m}{m!} = \exp(X+Y)\,,$$

where we in the third equality used the fact that X and Y commute.

(iv) Since X and $-X$ commute, this follows directly from (iii) and (i).

(v) Note that $(CXC^{-1})^k = CX^kC^{-1}$ for any $k \in \mathbb{Z}_0^+$, and that conjugation by C is a linear, continuous operator on $\mathbb{C}^{n\times n}$. The statement can now be proved the same way as we proved (ii). □

The following result, with clear resemblance to a result from calculus will be useful in the future.

Proposition 2.17. *Let $X \in \mathbb{C}^{n\times n}$. Then the map $\varphi\colon \mathbb{R} \to \mathbb{C}^{n\times n}$, defined by $\varphi(t) = \exp(tX)$, is smooth with*

$$\varphi'(t) = X\,\exp(tX) = \exp(tX)\,X\,.$$

In particular, $\varphi'(0) = X$.

Proof. The (i, j)-th entry of $\varphi(t)$ is given by

$$(\varphi(t))_{ij} = (\exp(tX))_{ij} = \sum_{k=0}^{\infty} \frac{t^k (X^k)_{ij}}{k!}\,,$$

which is a power series in t with complex coefficients that converges for all $t \in \mathbb{R}$. It is a well-known fact that a power series is term-wise differentiable within its radius of convergence. This implies that

$$\begin{aligned}(\varphi'(t))_{ij} = \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi(t))_{ij} &= \frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=0}^{\infty} \frac{t^k (X^k)_{ij}}{k!} = \sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t^k (X^k)_{ij}}{k!}\right)\\ &= \sum_{k=0}^{\infty} \frac{t^{k-1}(X^k)_{ij}}{(k-1)!} = \sum_{k=0} \frac{t^k (X^{k+1})_{ij}}{k!}\,,\end{aligned}$$

which in turn leads to

$$\varphi'(t) = \sum_{k=0}^{\infty} \frac{t^k X^{k+1}}{k!} = \left(\sum_{k=0}^{\infty} \frac{t^k X^k}{k!}\right) X = X\left(\sum_{k=0}^{\infty} \frac{t^k X^k}{k!}\right).$$

The desired result follows. Note that we in the two last equalities used the fact that multiplication by X from the left and the right, respectively, are linear, continuous operators on $\mathbb{C}^{n\times n}$, and therefore can be applied term-wise to any convergent series in $\mathbb{C}^{n\times n}$. □

Just as with the complex exponential map, the matrix exponential is not invertible on all of its domain. It is, however, locally invertible. To show this, we will first define a logarithm function, inspired by the Taylor series for $\ln(x)$ in the one-variable case.

Definition 2.18. Let $U = \{A \in \mathbb{C}^{n\times n} : \|A - I\| < 1\}$. The ***matrix logarithm*** for $n \times n$ matrices is the map $\log\colon U \to \mathbb{C}^{n\times n}$ defined by

$$\log(A) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(A-I)^k}{k}\,.$$

Theorem 2.19. *The series* $\sum_{k=1}^{\infty}(-1)^{k+1}(A-I)^k/k$ *converges absolutely for all* $A \in U$.

Proof. The submultiplicativity of the norm gives

$$\sum_{k=1}^{\infty} \left\| (-1)^{k+1} \frac{(A-I)^k}{k} \right\| = \sum_{k=1}^{\infty} \frac{\|(A-I)^k\|}{k} \leqslant \sum_{k=1}^{\infty} \frac{\|A-I\|^k}{k} \,,$$

which is bounded above by the geometric series $\sum_{k=1}^{\infty}\|A-I\|^k$, which converges when $\|A-I\| < 1$. □

Proposition 2.20. *The matrix logarithm* $\log\colon U \to \mathbb{C}^{n\times n}$ *is continuous.*

Proof. We prove this the same way we proved Proposition 2.15. For any fixed $R \in (0,1)$, let $D = \{A \in \mathbb{C}^{n\times n} : \|A-I\| < R\}$. For any $k \in \mathbb{Z}^+$, let f_k be the k-th term in the power series of log, and set $M_k = R^k/k$. Then proceed as in the proof of Proposition 2.15. □

We now show that the matrix logarithm is a local inverse of the matrix exponential. The elegant idea behind the proof can be traced back to [vN29].

Theorem 2.21.

(i) For every $X \in \mathbb{C}^{n\times n}$ *such that* $\|X\| < \ln(2)$, *we have* $\|\exp(X) - I\| < 1$ *(so that* $\log(\exp(X))$ *is defined) and* $\log(\exp(X)) = X$.

(ii) For every $A \in \mathbb{C}^{n\times n}$ *such that* $\|A-I\| < 1$, *we have* $\exp(\log(A)) = A$.

Proof. To prove (i), we first note that for any $X \in \mathbb{C}^{n\times n}$,

$$\|\exp(X) - I\| = \left\| \sum_{k=1}^{\infty} \frac{X^k}{k!} \right\| \leqslant \sum_{k=1}^{\infty} \frac{\|X^k\|}{k!} \leqslant \sum_{k=1}^{\infty} \frac{\|X\|^k}{k!} = e^{\|X\|} - 1 \,,$$

where we in the last inequality used Proposition 2.13. From this we conclude that $\|\exp(X) - I\| < 1$ for all $X \in \mathbb{C}^{n\times n}$ such that $\|X\| < \ln(2)$. Next, we observe that for every such $X \in \mathbb{C}^{n\times n}$,

$$\begin{aligned} \log(\exp(X)) &= \log\left(I + X + \tfrac{1}{2!}X^2 + \cdots\right) \\ &= \left(X + \tfrac{1}{2!}X^2 + \cdots\right) - \tfrac{1}{2}\left(X + \tfrac{1}{2!}X^2 + \cdots\right)^2 + \cdots \end{aligned}$$

Due to the absolute convergence, we can freely rearrange the terms to collect terms that contain the same power of X. This gives

$$\begin{aligned} \log(\exp(X)) &= X + \left(\tfrac{1}{2!} - \tfrac{1}{2}\right)X^2 + \left(\tfrac{1}{3!} - \tfrac{1}{2} + \tfrac{1}{3}\right)X^3 + \cdots \\ &= X + 0 + 0 + \cdots \end{aligned}$$

To convince ourselves that the coefficients of X^k sum to 0 for all $k > 1$, we note that the coefficients are the same in the real scalar case, and since we know that $\ln(e^x) = x$, we conclude that the coefficients of X^k for $k > 1$ must indeed add up to 0. The proof of (ii) is similar. □

A consequence of Theorem 2.21 is that the matrix exponential maps the open ball $B_{\ln(2)}(0)$ of radius $\ln(2)$ in $\mathbb{C}^{n\times n}$ homeomorphically onto its image $\exp(B_{\ln(2)}(0)) \subseteq \mathbf{GL}_n(\mathbb{C})$. The next theorem shows that this mapping is actually a diffeomorphism.

Theorem 2.22. *Both the matrix exponential and the matrix logarithm are smooth maps.*

We omit the proof, but note that this is a consequence of a more general result, which states that every complex power series $\sum_{k=0}^{\infty} c_k z^k$, of a certain radius of convergence $R > 0$, gives rise to a smooth map $X \mapsto \sum_{k=0}^{\infty} c_k X^k$ defined on the open ball of radius R in $\mathbb{C}^{n\times n}$ (or, more generally, any complex unital Banach algebra). See Section 3.1 in [HN11] for details.

We conclude this section by collecting a couple of useful identities involving the matrix exponential.

Proposition 2.23. *For any* $X \in \mathbb{C}^{n\times n}$, $\det(\exp(X)) = e^{\mathrm{trace}(X)}$.

Proof. Let $\{\lambda_1, \dots, \lambda_n\}$ be the multiset of eigenvalues of X (included according to multiplicity). From linear algebra, we know that $X = PJP^{-1}$, where J is in Jordan canonical form and $P \in \mathbf{GL}_n(\mathbb{C})$. Note that $J = D + T$, where $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, T is a strictly upper triangular, and $DT = TD$.

Direct computation shows that $\exp(D) = \mathrm{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$ and that $\exp(T)$ is an upper triangular matrix with ones on the diagonal. This, together with Proposition 2.16(v) and elementary linear algebra gives

$$\begin{aligned}
\det(\exp(X)) &= \det(\exp(PJP^{-1})) = \det(P\exp(J)P^{-1}) \\
&= \det(\exp(J)) = \det(\exp(D+T)) \\
&= \det(\exp(D)\exp(T)) = \det(\exp(D))\det(\exp(T)) \\
&= e^{\lambda_1}\cdots e^{\lambda_n}\cdot 1\cdots 1 = e^{\lambda_1+\cdots+\lambda_n} = e^{\mathrm{trace}(X)}. \qquad \square
\end{aligned}$$

The following formula will help us handle $\exp(X+Y)$ in the general case when $X, Y \in \mathbb{C}^{n\times n}$ do not necessarily commute.

Proposition 2.24 (Lie product formula)**.** *For all* $X, Y \in \mathbb{C}^{n\times n}$,

$$\exp(X+Y) = \lim_{N\to\infty}\left(\exp\left(\frac{X}{N}\right)\exp\left(\frac{Y}{N}\right)\right)^N.$$

To prove this, we will temporarily switch norms. This is allowed, since all norm on a finite-dimensional $\mathbb{C}$-vector space such as $\mathbb{C}^{n\times n}$ are equivalent, and therefore induce the same topology. Hence, convergence in one norm implies convergence in all other norms.

Definition 2.25. The ***operator norm*** for $n \times n$ matrices is the map $\|\cdot\|_{\mathrm{op}} \colon \mathbb{C}^{n\times n} \to \mathbb{R}$ defined by $\|A\|_{\mathrm{op}} = \max\{|Ax| : x \in \mathbb{C}^n \text{ and } |x| = 1\}$. Here, $|\cdot|$ denotes the standard Euclidean norm on $\mathbb{C}^n$.

It is easy to verify that $\|\cdot\|_{\mathrm{op}}$ is a norm, that $\|AB\|_{\mathrm{op}} \leqslant \|A\|_{\mathrm{op}}\|B\|_{\mathrm{op}}$ for all $A, B \in \mathbb{C}^{n\times n}$, and that $\|I\|_{\mathrm{op}} = 1$. It also holds that $\|\exp(X)\|_{\mathrm{op}} \leqslant e^{\|X\|_{\mathrm{op}}}$.

Proof of Proposition 2.24. We start by observing that the terms of degree 0 and 1 in the power series expansions

$$\exp\left(\tfrac{X}{N}\right)\exp\left(\tfrac{Y}{N}\right) = \left(I + \tfrac{X}{N} + \cdots\right)\left(I + \tfrac{Y}{N} + \cdots\right) = I + \tfrac{X}{N} + \tfrac{Y}{N} + \cdots$$

and

$$\exp\left(\tfrac{X+Y}{N}\right) = I + \tfrac{X+Y}{N} + \cdots$$

are identical. Since all other terms are small for large values of N, we might suspect that we do not only have

$$\exp\left(\tfrac{X}{N}\right)\exp\left(\tfrac{Y}{N}\right) - \exp\left(\tfrac{X+Y}{N}\right) \to 0$$

as $N \to \infty$, but also

$$\left(\exp\left(\tfrac{X}{N}\right)\exp\left(\tfrac{Y}{N}\right)\right)^N - \left(\exp\left(\tfrac{X+Y}{N}\right)\right)^N \to 0\,.$$

Note that $\exp((X+Y)/N)^N = \exp(X+Y)$, so this would give the desired conclusion. We will now turn this idea into a formal proof by using the big O notation*. We first observe that

$$\begin{aligned}\left\|\exp\left(\tfrac{X}{N}\right) - I - \tfrac{X}{N}\right\|_{\mathrm{op}} &= \left\|\sum_{k=2}^{\infty}\frac{X^k}{k!N^k}\right\|_{\mathrm{op}} \leqslant \sum_{k=2}^{\infty}\frac{\|X\|_{\mathrm{op}}^k}{k!N^k} \\ &\leqslant \frac{1}{N^2}\|X\|_{\mathrm{op}}^2\sum_{k=0}^{\infty}\frac{\|X\|_{\mathrm{op}}^k}{(k+2)!N^k} \leqslant \frac{1}{N^2}\|X\|_{\mathrm{op}}^2\sum_{k=0}^{\infty}\frac{\|X\|_{\mathrm{op}}^k}{k!N^k} \\ &\leqslant \frac{1}{N^2}\|X\|_{\mathrm{op}}^2 e^{\|X\|_{\mathrm{op}}/N} \leqslant \frac{1}{N^2}\|X\|_{\mathrm{op}}^2 e^{\|X\|_{\mathrm{op}}}\,,\end{aligned}$$

which shows that

$$\exp\left(\tfrac{X}{N}\right) = I + \tfrac{X}{N} + \mathscr{O}\left(\tfrac{1}{N^2}\right).$$

Similarly,

$$\begin{aligned}\exp\left(\tfrac{Y}{N}\right) &= I + \tfrac{Y}{N} + \mathscr{O}\left(\tfrac{1}{N^2}\right) \\ \exp\left(\tfrac{X+Y}{N}\right) &= I + \tfrac{X+Y}{N} + \mathscr{O}\left(\tfrac{1}{N^2}\right),\end{aligned}$$

and it is easy to verify that

$$\exp\left(\tfrac{X}{N}\right)\exp\left(\tfrac{Y}{N}\right) = I + \tfrac{X}{N} + \tfrac{Y}{N} + \mathscr{O}\left(\tfrac{1}{N^2}\right).$$

*Given a map $G\colon \mathbb{Z}^+ \to \mathbb{C}^{n\times n}$ (or a map $g\colon \mathbb{Z}^+ \to \mathbb{C}$), and a map $f\colon \mathbb{Z}^+ \to \mathbb{R}$, we will write $G(N) = \mathscr{O}(f(N))$ (or $g(N) = \mathscr{O}(f(N))$ to indicate the existence of some constant $c \in \mathbb{R}$ such that $\|G(N)\|_{\mathrm{op}} \leqslant cf(N)$ (or $|g(N)| \leqslant cf(N)$) for all sufficiently large $N \in \mathbb{Z}^+$.

For each $N \in \mathbb{Z}^+$, we now introduce the notation

$$\begin{aligned} A_N &= \exp\left(\tfrac{X}{N}\right)\exp\left(\tfrac{Y}{N}\right), \\ B_N &= \exp\left(\tfrac{X+Y}{N}\right). \end{aligned}$$

Clearly $\|A_N - B_N\|_{\text{op}} = \mathscr{O}(1/N^2) \to 0$ as $N \to \infty$. To show the desired limit $\|A_N^N - B_N^N\|_{\text{op}} \to 0$, we will use Lemma 2.26 (see below), which says that

$$\|A_N^N - B_N^N\|_{\text{op}} \leqslant N M_N^{N-1} \|A_N - B_N\|_{\text{op}},$$

where $M_N = \max\{\|A_N\|_{\text{op}}, \|B_N\|_{\text{op}}\}$. By making the estimates

$$\begin{aligned} \|A_N\|_{\text{op}} &= \left\| \exp\left(\tfrac{X}{N}\right)\exp\left(\tfrac{Y}{N}\right) \right\|_{\text{op}} = \left\| \exp\left(\tfrac{X}{N}\right) \right\|_{\text{op}} \left\| \exp\left(\tfrac{X}{N}\right) \right\|_{\text{op}} \\ &\leqslant e^{\|X\|_{\text{op}}/N} e^{\|Y\|_{\text{op}}/N} = e^{(\|X\|_{\text{op}} + \|Y\|_{\text{op}})/N}, \\ \|B_N\|_{\text{op}} &= \left\| \exp\left(\tfrac{X+Y}{N}\right) \right\|_{\text{op}} \leqslant e^{\|X+Y\|_{\text{op}}/N} \leqslant e^{(\|X\|_{\text{op}} + \|Y\|_{\text{op}})/N}. \end{aligned}$$

we conclude that $M_N \leqslant e^{(\|X\|_{\text{op}} + \|Y\|_{\text{op}})/N}$, and we thus have

$$\begin{aligned} \|A_N^N - B_N^N\|_{\text{op}} &\leqslant N e^{(\|X\|_{\text{op}} + \|Y\|_{\text{op}})(N-1)/N} \mathscr{O}\left(\tfrac{1}{N^2}\right) \\ &\leqslant N e^{\|X\|_{\text{op}} + \|Y\|_{\text{op}}} \mathscr{O}\left(\tfrac{1}{N^2}\right) = \mathscr{O}\left(\tfrac{1}{N}\right), \end{aligned}$$

i.e. $\|A_N^N - B_N^N\|_{\text{op}} \to 0$ as $N \to \infty$. □

Lemma 2.26. *Let $A, B \in \mathbb{C}^{n\times n}$. Then, for any $N \in \mathbb{Z}^+$, it holds that*

$$\|A^N - B^N\|_{\text{op}} \leqslant N M^{N-1} \|A - B\|_{\text{op}},$$

where $M = \max\{\|A\|_{\text{op}}, \|B\|_{\text{op}}\}$.

Proof. Consider the telescopic sum

$$A^N - B^N = \sum_{k=1}^{N} \left(A^{N-k+1}B^{k-1} - A^{N-k}B^k\right) = \sum_{k=1}^{N} \left(A^{N-k}(A-B)B^{k-1}\right).$$

Taking the operator norm on both sides, and applying the triangle inequality gives

$$\begin{aligned} \|A^N - B^N\|_{\text{op}} &\leqslant \sum_{k=1}^{N} \|A^{N-k}\|_{\text{op}} \|A-B\|_{\text{op}} \|B^{k-1}\|_{\text{op}} \\ &\leqslant \sum_{k=1}^{N} \|A\|_{\text{op}}^{N-k} \|A-B\|_{\text{op}} \|B\|_{\text{op}}^{k-1} \\ &\leqslant \sum_{k=1}^{N} M^{N-1} \|A-B\|_{\text{op}} = N M^{N-1} \|A-B\|_{\text{op}}. \end{aligned}$$

□

2.3 The Lie algebra of a matrix Lie group

In our attempts to translate problems about Lie groups into problems in linear algebra, we will associate a Lie algebra to each matrix Lie group. Together with the exponential map from the previous section, this will become a very powerful tool for understanding matrix Lie groups.

Definition 2.27. A ***Lie algebra*** is a pair $(\mathfrak{g}, [.,.])$, where $\mathfrak{g}$ is a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, and $[.,.]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is a map that satisfies the following conditions:

(i) Bilinearity: $[ax + by, z] = a[x, z] + b[y, z]$ and $[x, ay + bz] = a[x, y] + b[x, z]$ for all $x, y, z \in \mathfrak{g}$ and $a, b \in \mathbb{F}$.

(ii) Anticommutativity: $[x, y] = -[y, x]$ for all $x, y \in \mathfrak{g}$.

(iii) The Jacobi identity: $[x, [y, z]]+[z, [x, y]]+[y, [z, x]] = 0$ for all $x, y, z \in \mathfrak{g}$.

Definition 2.28. Let G be a matrix Lie group. Then the ***associated Lie algebra*** of G is given by $\mathfrak{g} = \{X \in \mathbb{C}^{n\times n} : \exp(tX) \in G \text{ for all } t \in \mathbb{R}\}$.

The term "Lie algebra" is justified, since for any matrix Lie group G, the associated Lie algebra $\mathfrak{g}$ really is a Lie algebra over $\mathbb{R}$ in the sense of Definition 2.27, if equipped with the usual matrix commutator $[.,.]$, defined by $[X, Y] = XY - YX$, as its Lie bracket. We formulate this as a theorem.

Theorem 2.29. *Let $G \subseteq \mathbf{GL}_n(\mathbb{C})$ be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then*

(i) $sX \in \mathfrak{g}$ *for all* $s \in \mathbb{R}$ *and* $X \in \mathfrak{g}$.

(ii) $X + Y \in \mathfrak{g}$ *for all* $X, Y \in \mathfrak{g}$.

(iii) $AXA^{-1} \in \mathfrak{g}$ *for all* $A \in G$ *and* $X \in \mathfrak{g}$.

(iv) $[X, Y] \in \mathfrak{g}$ *for all* $X, Y \in \mathfrak{g}$.

(v) $[.,.]$ *is bilinear, anticommutative and satisfies the Jacobi identity.*

Proof.

(i) For any $X \in \mathfrak{g}$ and $s \in \mathbb{R}$, we have $\exp(t(sX)) = \exp((ts)X) \in G$ for all $t \in \mathbb{R}$. This shows $sX \in \mathfrak{g}$.

(ii) Let $X, Y \in \mathfrak{g}$. For any $t \in \mathbb{R}$, Proposition 2.24 gives

$$\exp(t(X + Y)) = \lim_{N\to\infty} \left(\exp\left(\tfrac{tX}{N}\right) \exp\left(\tfrac{tY}{N}\right)\right)^N ,$$

from which we conclude that $\exp(t(X + Y))$ is an invertible limit of elements in G. Since G is closed in $\mathbf{GL}_n(\mathbb{C})$, we have $\exp(t(X+Y)) \in G$, i.e. $X + Y \in \mathfrak{g}$.

(iii) Let $A \in G$ and $X \in \mathfrak{g}$. By Proposition 2.16(v) and the fact that $X \in \mathfrak{g}$, we have

$$\exp(tAXA^{-1}) = \exp(A(tX)A^{-1}) = A\exp(tX)A^{-1} \in G$$

for all $t \in \mathbb{R}$. Hence, $AXA^{-1} \in \mathfrak{g}$.

(iv) The product rule for matrix-valued functions gives that

$$\begin{aligned} XY - YX &= \tfrac{\mathrm{d}}{\mathrm{d}t}\Big(\exp(tX)Y\exp(-tX)\Big)\Big|_{t=0} \\ &= \lim_{h\to 0} \frac{\exp(hX)Y\exp(-hX) - Y}{h} \,. \end{aligned}$$

According to (iii),

$$\exp(hX)Y\exp(-hX) = \exp(hX)Y\exp(hX)^{-1} \in \mathfrak{g} \,.$$

Together with (i) and (ii), this shows that $XY - YX$ can be expressed as a limit of elements in $\mathfrak{g}$. We also know that $\mathfrak{g}$ is an $\mathbb{R}$-linear subspace of the finite-dimensional $\mathbb{R}$-vector space $\mathbb{C}^{n\times n}$. Consequently $\mathfrak{g}$ is closed in $\mathbb{C}^{n\times n}$, and we conclude that $XY - YX \in \mathfrak{g}$.

(v) This can be verified by direct computation. □

A useful interpretation of the Lie algebra $\mathfrak{g}$ of a matrix Lie group G, is as the set of tangents at the identity for a special kind of curves in G called one-parameter subgroups.

Definition 2.30. Let G be a Lie group. A ***one-parameter subgroup*** of G is a continuous group homomorphism $\varphi\colon (\mathbb{R}, +) \to G$.

Remark 2.31. Note that the name one-parameter *subgroup* is a bit of a misnomer, since it is $\varphi(\mathbb{R})$ that is a subgroup of G, not the map φ itself.

It is easy to show that $t \mapsto \exp(tX)$ is a one-parameter subgroup of G for every X in the Lie algebra of G. We will now show that *all* one-parameter subgroups are of this form. The proof is an adaptation of that in [Hal15].

Theorem 2.32. *Let $G \subseteq \mathbf{GL}_n(\mathbb{C})$ be a matrix Lie group with Lie algebra $\mathfrak{g}$, and let $\varphi\colon \mathfrak{g} \to G$ be a one-parameter subgroup of G. Then there exists a unique $X \in \mathbb{C}^{n\times n}$ such that $\varphi(t) = \exp(tX)$ for all $t \in \mathbb{R}$.*

Proof. The uniqueness follows from Proposition 2.17; if $X, Y \in \mathfrak{g}$ both are such that $\exp(tX) = \varphi(t) = \exp(tY)$ for all $t \in \mathbb{R}$, then differentiation at $t = 0$ gives $X = Y$.

To show the existence, we pick some $\varepsilon \in (0, \ln(2))$ and form the open ball $B_\varepsilon = \{X \in \mathbb{C}^{n\times n} : \|X\| < \varepsilon\}$. Note that by Theorem 2.21, $\exp|_{B_\varepsilon}$ is a homeomorphism with inverse $\log|_{\exp(B_\varepsilon)}$. Next, form the set $U = \exp(B_{\varepsilon/2})$. Note that $U \subseteq \exp(B_\varepsilon)$ is open in $\mathbf{GL}_n(\mathbb{C})$, which implies that $\varphi^{-1}(U \cap G)$ is open in $\mathbb{R}$. A consequence of this is that there exists some $t_0 > 0$ such that $\varphi(t) \in U$ for all $t \in [-t_0, t_0]$. In particular, $\log(\varphi(t_0)) \in B_{\varepsilon/2}$, so by taking $X = \frac{1}{t_0}\log(\varphi(t_0))$, we get $t_0 X \in B_{\varepsilon/2}$ and $\varphi(t_0) = \exp(t_0 X)$.

Our goal now will be to show that $\varphi(t) = \exp(tX)$ for all $t \in \mathbb{R}$, not just $t = t_0$. We start by considering $t = t_0/2$. We want to show that $\varphi(t_0/2) = \exp(t_0 X/2)$, or equivalently: $\log(\varphi(t_0/2)) = t_0 X/2$. To do this, we start by writing

$$\begin{aligned}\exp\Big(2\log(\varphi(t_0/2)\Big) &= \exp\Big(\log(\varphi(t_0/2)\Big)^2 \\ &= \varphi(t_0/2)^2 = \varphi(t_0) = \exp(t_0 X)\,.\end{aligned}$$

Notice that $t_0 X \in B_{\varepsilon/2} \subseteq B_\varepsilon$ and that we also must have $2\log(\varphi(t_0/2)) \in B_\varepsilon$ since $\varphi(t_0/2) \in U$. The local bijectivity of the matrix exponential therefore gives $2\log(\varphi(t_0/2)) = t_0 X$, which is equivalent to what we wanted to show.

Next, we consider $t = t_0/4$. We want to show $\varphi(t_0/4) = \exp(t_0 X/4)$, or equivalently: $\log(\varphi(t_0/4)) = t_0 X/4$. We therefore observe that

$$\begin{aligned}\exp\Big(2\log(\varphi(t_0/4)\Big) &= \exp\Big(\log(\varphi(t_0/4)\Big)^2 \\ &= \varphi(t_0/4)^2 = \varphi(t_0/2) = \exp(t_0 X/2)\,.\end{aligned}$$

As above, we note that $t_0 X/2 \in B_\varepsilon$ and that $\varphi(t_0/4) \in U$, so that $2\log(\varphi(t_0/4)) \in B_\varepsilon$, and we then use the local bijectivity of the matrix exponential to conclude that $2\log(\varphi(t_0/4)) = t_0 X/2$, i.e. $\log(\varphi(t_0/4)) = t_0 X/4$.

Continuing in this fashion, we can inductively show that $\varphi(t) = \exp(tX)$ for all $t = t_0/2^k$ for $k \in \mathbb{Z}^+$. From there, it is easy to go on and show that $\varphi(t) = \exp(tX)$ for all $t = mt_0/2^k$ for $m \in \mathbb{Z}$ and $k \in \mathbb{Z}^+$; the fact that φ is a group homomorphism gives

$$\varphi(mt_0/2^k) = \varphi(t_0/2^k)^m = \exp(t_0 X/2^k)^m = \exp(mt_0 X/2^k)\,.$$

Finally, because any $t \in \mathbb{R}$ can be arbitrarily well approximated by real numbers of the form $mt_0/2^k$, with $m \in \mathbb{Z}$ and $k \in \mathbb{Z}^+$, and because the maps $t \mapsto \varphi(t)$ and $t \mapsto \exp(tX)$ are continuous, we conclude that $\varphi(t) = \exp(tX)$ for all $t \in \mathbb{R}$. □

Remark 2.33. There are two important consequences of Theorem 2.32. One is that every one-parameter subgroup is smooth, by Proposition 2.17. Another is that there is a one-to-one correspondence between one-parameter subgroups and elements of $\mathfrak{g}$, in the sense that

$$\mathfrak{g} = \left\{ \varphi'(0) : \begin{array}{c}\varphi\colon \mathbb{R} \to G \text{ is a} \\ \text{one-parameter subgroup}\end{array} \right\}.$$

We will later take this one step further, and show that $\mathfrak{g}$ is the set of tangents at the identity for *all* smooth curves in G through the identity (see Theorem 2.52).

We now compute the Lie algebras for a few well-known matrix Lie groups.

Example 2.34. Let $\mathfrak{gl}_n(\mathbb{C})$ be the Lie algebra of $\mathbf{GL}_n(\mathbb{C})$. Then

$$\mathfrak{gl}_n(\mathbb{C}) = \mathbb{C}^{n\times n}.$$

Proof. This is a direct consequence of Proposition 2.16(iv). □

Example 2.35. Let $\mathfrak{gl}_n(\mathbb{R})$ be the Lie algebra of $\mathbf{GL}_n(\mathbb{R})$. Then

$$\mathfrak{gl}_n(\mathbb{R}) = \mathbb{R}^{n\times n}.$$

Proof. $\boxed{\subseteq}$ Let $X \in \mathbb{C}^{n\times n}$ be such that $\exp(tX) \in \mathbf{GL}_n(\mathbb{R})$ for all $t \in \mathbb{R}$. Then Proposition 2.17 gives that

$$X = \tfrac{\mathrm{d}}{\mathrm{d}t}\exp(tX)\Big|_{t=0} = \lim_{h\to 0}\frac{\exp(hX) - I}{h},$$

i.e. X can be expressed as a limit of elements in $\mathbb{R}^{n\times n}$. Since $\mathbb{R}^{n\times n}$ is closed in $\mathbb{C}^{n\times n}$, this implies that $X \in \mathbb{R}^{n\times n}$.

$\boxed{\supseteq}$ Proposition 2.16(iv), together with the fact that $\mathbb{R}^{n\times n}$ is closed in $\mathbb{C}^{n\times n}$, gives that $\exp(tX) \in \mathbf{GL}_n(\mathbb{C}) \cap \mathbb{R}^{n\times n} = \mathbf{GL}_n(\mathbb{R})$ for all $X \in \mathbb{R}^{n\times n}$ and $t \in \mathbb{R}$. □

Example 2.36. Let $\mathfrak{sl}_n(\mathbb{C})$ be the Lie algebra of $\mathbf{SL}_n(\mathbb{C})$. Then

$$\mathfrak{sl}_n(\mathbb{C}) = \{X \in \mathbb{C}^{n\times n} : \mathrm{trace}(X) = 0\}.$$

Proof. Note that $X \in \mathfrak{sl}_n(\mathbb{C})$ is equivalent to $\det(\exp(tX)) = 1$ for all $t \in \mathbb{R}$, which by Proposition 2.23 is equivalent to

$$e^{t\,\mathrm{trace}(X)} = e^{\mathrm{trace}(tX)} = 1 \tag{2.2}$$

for all $t \in \mathbb{R}$. By differentiating at $t = 0$ we see that this implies $\mathrm{trace}(X) = 0$. It is also clear that (2.2) is satisfied by $X \in \mathbb{C}^{n\times n}$ such that $\mathrm{trace}(X) = 0$. □

Proposition 2.37. *Let $G, H, K \subseteq \mathbf{GL}_n(\mathbb{C})$ be matrix Lie groups with Lie algebras $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{k}$, respectively. Suppose that $G = H \cap K$. Then $\mathfrak{g} = \mathfrak{h} \cap \mathfrak{k}$.*

Proof. This follows immediately from Definition 2.28. □

Example 2.38. Let $\mathfrak{sl}_n(\mathbb{R})$ be the Lie algebra of $\mathbf{SL}_n(\mathbb{R})$. Then

$$\mathfrak{sl}_n(\mathbb{R}) = \{X \in \mathbb{R}^{n\times n} : \mathrm{trace}(X) = 0\}.$$

Proof. Note that $\mathbf{SL}_n(\mathbb{R}) = \mathbf{SL}_n(\mathbb{C}) \cap \mathbf{GL}_n(\mathbb{R})$, so by Proposition 2.37,

$$\mathfrak{sl}_n(\mathbb{R}) = \mathfrak{sl}_n(\mathbb{C}) \cap \mathfrak{gl}_n(\mathbb{R}) = \{X \in \mathbb{R}^{n\times n} : \mathrm{trace}(X) = 0\}.$$ □

Example 2.39. Let $\mathfrak{u}(n)$ be the Lie algebra of $\mathbf{U}(n)$. Then

$$\mathfrak{u}(n) = \{X \in \mathbb{C}^{n\times n} : X^* + X = 0\} .$$

Proof. Note that $X \in \mathfrak{u}(n)$ if and only if $\exp(tX)^* = \exp(tX)^{-1}$ for all $t \in \mathbb{R}$. By Proposition 2.16(ii), this is equivalent to

$$\exp(tX^*) = \exp(-tX) \tag{2.3}$$

for all $t \in \mathbb{R}$. Differentiation at $t = 0$ according to Proposition 2.17 gives $X^* = -X$. It is also clear that (2.3) is satisfied by all $X \in \mathbb{C}^{n\times n}$ such that $X^* = -X$. □

Example 2.40. Let $\mathfrak{su}(n)$ be the Lie algebra of $\mathbf{SU}(n)$. Then

$$\mathfrak{su}(n) = \{X \in \mathbb{C}^{n\times n} : X^* + X = 0 \text{ and } \operatorname{trace}(X) = 0\} .$$

Proof. This follows from Proposition 2.37 together with

$$\mathbf{SU}(n) = \mathbf{U}(n) \cap \mathbf{SL}_n(\mathbb{C}) .$$

□

Example 2.41. Let $\mathfrak{o}(n)$ be the Lie algebra of $\mathbf{O}(n)$, and let $\mathfrak{so}(n)$ be the Lie algebra of $\mathbf{SO}(n)$. Then

$$\mathfrak{o}(n) = \mathfrak{so}(n) = \{X \in \mathbb{R}^{n\times n} : X^\top + X = 0\} .$$

Proof. This follows from Proposition 2.37 and the equalities

$$\mathbf{O}(n) = \mathbf{U}(n) \cap \mathbf{GL}_n(\mathbb{R}) \text{ and } \mathbf{SO}(n) = \mathbf{O}(n) \cap \mathbf{SL}_n(\mathbb{R}) ,$$

together with the fact that $X^\top + X = 0$ implies $\operatorname{trace}(X) = 0$. □

Example 2.42. The Lie algebra of the Heisenberg group is given by

$$\mathfrak{h} = \left\{ \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} : a, b, c \in \mathbb{R} \right\} .$$

Proof. $\boxed{\subseteq}$ Let $X \in \mathbb{C}^{n\times n}$ be such that $\exp(tX) \in H$ for all $t \in \mathbb{R}$. Then the diagonal and lower diagonal entries of $\exp(tX)$ are constantly equal to 1 and 0, respectively. Hence, $X = \frac{\mathrm{d}}{\mathrm{d}t} \exp(tX)\Big|_{t=0}$ is of the desired form.

$\boxed{\supseteq}$ If X is strictly upper diagonal, it is not difficult to see that for all $t \in \mathbb{R}$, the series expansion of $\exp(tX)$ terminates after finitely many terms, and only involves I and finitely many strictly upper diagonal terms. Hence, $\exp(tX) \in H$. □

We are now ready to define the exponential map of a matrix Lie group.

Definition 2.43. Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then the ***exponential map*** for G is the restriction $\exp|_{\mathfrak{g}} \colon \mathfrak{g} \to G$ of the matrix exponential.

It is easy to see that the exponential map of a matrix Lie group in general is not bijective. For example, we have that

$$\exp|_{\mathfrak{su}(2)}\left(\begin{pmatrix} 2\pi i & 0 \\ 0 & -2\pi i \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \exp|_{\mathfrak{su}(2)}\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right),$$

so $\exp|_{\mathfrak{su}(2)}$ is not injective. We also have that for any $X \in \mathbb{R}^{n\times n}$, $\exp(X) = \exp(X/2)\exp(X/2)$, so $\det(\exp(X)) = \det(\exp(X/2))^2 \geqslant 0$. This shows that $\exp|_{\mathfrak{gl}_n(\mathbb{R})}$ is not surjective.

The following results, which can be traced back to [vN29], show that the exponential map is at least locally bijective. The proofs we give here follow the more modern approaches found in e.g. [Hal15] and [Lee13].

Lemma 2.44. *Let $G \subseteq \mathbf{GL}_n(\mathbb{C})$ be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Then there exists* an $\varepsilon \in (0, \ln(2))$ such that for all $A \in \exp(B_\varepsilon(0))$, $A \in G$ if and only if $\log(A) \in \mathfrak{g}$. Here, $B_\varepsilon(0) = \{X \in \mathbb{C}^{n\times n} : \|X\| < \varepsilon\}$.*

Proof. $\boxed{\Longleftarrow}$ This is a direct consequence of Theorem 2.21. For any $\varepsilon < \ln(2)$ and any $A \in \exp(B_\varepsilon(0))$ such that $\log(A) \in \mathfrak{g}$, it holds that $A = \exp(\log(A))$, and hence $A \in G$ by the definition of $\mathfrak{g}$.

$\boxed{\Longrightarrow}$ Assume towards a contradiction that for any $\varepsilon \in (0, \ln(2))$, there exists an $A \in G \cap \exp(B_\varepsilon(0))$ such that $\log(A) \notin \mathfrak{g}$. We can then construct a sequence $(A_k)_{k=1}^\infty$ in $G \cap \exp(B_{\ln(2)}(0))$, with $A_k \in G \cap \exp(B_{1/(k+1)}(0))$ and $\log(A_k) \notin \mathfrak{g}$. Since exp is continuous, we must have $A_k \to I$ as $k \to \infty$.

The idea will now be to refine this into another sequence, with log-values lying entirely in $\mathfrak{g}^\perp$ (the orthogonal complement of $\mathfrak{g}$ in $\mathbb{C}^{n\times n}$ with respect to the usual inner product). From this we will construct a non-zero $Y \in \mathfrak{g}^\perp$ such that $Y \in \mathfrak{g}$, which will give us the desired contradiction, since $\mathfrak{g} \cap \mathfrak{g}^\perp = \{0\}$.

For this end, we define a map $\Phi \colon \mathfrak{g} \oplus \mathfrak{g}^\perp = \mathbb{C}^{n\times n} \to \mathbf{GL}_n(\mathbb{C})$ with $\Phi(X+Y) = \exp(X)\exp(Y)$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}^\perp$. Note that Φ is smooth since exp is smooth, and that we, by the product rule, have

$$\mathrm{d}\Phi_0(X+Y) = \tfrac{\mathrm{d}}{\mathrm{d}t}\Phi(t(X+Y))\Big|_{t=0} = \tfrac{\mathrm{d}}{\mathrm{d}t}\exp(tX)\exp(tY)\Big|_{t=0} = X + Y\,,$$

for any $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}^\perp$. This shows that $\mathrm{d}\Phi_0$ is the identity map in $\mathbb{C}^{n\times n}$, and in particular, that $\mathrm{d}\Phi_0$ is invertible. By the inverse function theorem, we conclude that Φ has a smooth inverse defined in some open neighbourhood U of I.

*Note that the requirement $\varepsilon < \ln(2)$ ensures that $\log(A)$ is well-defined for all $A \in \exp(B_\varepsilon(0))$.

Since $A_k \to I$, there exists some $K \in \mathbb{Z}^+$ such that $A_k \in U$ for all $k \geqslant K$. For each such $k \geqslant K$, apply Φ^{-1} to A_k to obtain $X_k \in \mathfrak{g}$ and $Y_k \in \mathfrak{g}^\perp$ such that $A_k = \exp(X_k)\exp(Y_k)$. Since Φ^{-1} is continuous, $X_k, Y_k \to 0$ as $k \to \infty$. We also note that $Y_k \neq 0$ for all $k \geqslant K$, since otherwise we would have $\log(A_k) = X_k \in \mathfrak{g}$. We can now form a sequence $(C_k)_{k=K}^\infty$ by setting $C_k = \exp(Y_k) = \exp(-X_k)A_k$, where clearly $C_k \in G$, since it is a product of two elements of G. Since exp is continuous, we have that $C_k \to I$. If we normalize the Y_k's, we obtain a sequence $(Y_k/\|Y_k\|)_{k=K}^\infty$ in the unit sphere of $\mathfrak{g}^\perp$. Since this sphere is a compact set, there must exist some subsequence $(Y_{k_i}/\|Y_{k_i}\|)_{i=1}^\infty$ such that $Y_{k_i}/\|Y_{k_i}\| \to Y$ as $i \to \infty$ for some $Y \in \mathfrak{g}^\perp$ with $\|Y\| = 1$.

To show that $Y \in \mathfrak{g}$, we fix an arbitrary $t \in \mathbb{R}$, and try to show that $\exp(tY) \in G$. Since G is closed in $\mathbf{GL}_n(\mathbb{C})$ by assumption, and since it is clear that $\exp(tY) \in \mathbf{GL}_n(\mathbb{C})$, it will suffice to show that $\exp(tY)$ can be approximated arbitrarily well by elements from G. For this end, we use that $Y_{k_i}/\|Y_{k_i}\| \to Y$ and that* $m_{k_i}\|Y_{k_i}\| \to t$, where $m_{k_i} = \lfloor t/\|Y_{k_i}\| \rfloor$. Using the continuity of both multiplication by a scalar and exp, we get

$$\begin{aligned}\exp(tY) &= \lim_{i\to\infty} \exp\left(m_{k_i}\|Y_{k_i}\| \cdot \frac{Y_{k_i}}{\|Y_{k_i}\|} \right) = \lim_{i\to\infty} \exp(m_{k_i} Y_{k_i}) \\ &= \lim_{i\to\infty} \exp(Y_{k_i})^{m_{k_i}} = \lim_{i\to\infty} C_{k_i}^{m_{k_i}} ,\end{aligned}$$

where $C_{k_i}^{m_{k_i}} \in G$ since it is a product of elements of G. From this we conclude that $\exp(tY) \in G$, and since $t \in \mathbb{R}$ was arbitrary, that $Y \in \mathfrak{g}$. But at the same time, $Y \in \mathfrak{g}^\perp$ with $\|Y\| = 1$, which gives us the desired contradiction. □

Theorem 2.45. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then there exists an open neighbourhood U of 0 in $\mathfrak{g}$ and an open neighbourhood V of I in G, such that $\exp|_U : U \to V$ is a homeomorphism.*

Proof. Let $\varepsilon \in (0, \ln(2))$ be such that Lemma 2.44 holds, and set $U = B_\varepsilon(0) \cap \mathfrak{g}$ and $V = \exp(B_\varepsilon(0)) \cap G$. Then $\exp|_U : U \to V$ is a bijection, with inverse $\log|_V$. Since both $\exp|_U$ and $\log|_V$ are continuous (they are restrictions of continuous maps), we conclude that $\exp|_U$ is a homeomorphism. □

We close this section by giving a first hint on how the Lie algebra of a matrix Lie group controls the properties of the matrix Lie group.

Theorem 2.46. *Let G be a path-connected matrix Lie group. Then for any $A \in G$, there exists some $X_1, \ldots, X_m \in \mathfrak{g}$ such that*

$$A = \exp(X_1)\exp(X_2)\cdots\exp(X_m) .$$

*Note that $m_{k_i} \leqslant t/\|Y_{k_i}\| \leqslant m_{k_i} + 1$, which implies $m_{k_i}\|Y_{k_i}\| \leqslant t \leqslant m_{k_i}\|Y_{k_i}\| + \|Y_{k_i}\|$, and hence $|t - m_{k_i}\|Y_{k_i}\|| \leqslant \|Y_{k_i}\| \to 0$ as $i \to \infty$.

Remark 2.47. If G is not connected, the theorem holds for the connected component of G that contains the identity, which can be shown to be a matrix Lie group in its own right.

Before giving the proof, we first establish the following lemma.

Lemma 2.48. *Let G be a matrix Lie group, and let $\gamma\colon [0,1] \to G$ be a continuous map. Then for every $\varepsilon > 0$ there exists some $\delta > 0$ such that $s, t \in [0,1]$ and $|s-t| < \delta$ implies*

$$\|\gamma(s)\gamma(t)^{-1} - I\| < \varepsilon\,.$$

Proof. We make the following estimate:

$$\|\gamma(s)\gamma(t)^{-1} - I\| = \|(\gamma(s) - \gamma(t))\gamma(t)^{-1}\| \leqslant \|\gamma(s) - \gamma(t)\|\|\gamma(t)^{-1}\|\,.$$

Note that the map $t \mapsto \|\gamma(t)^{-1}\|$ is the composition of continuous functions and thus continuous itself. Since it is defined on the compact set $[0,1]$, the extreme value theorem guarantees the existence of $M = \max_{t\in[0,1]}\|\gamma(t)^{-1}\|$. Furthermore, the uniform continuity theorem (see Theorem 27.6 in [Mun00]) shows that γ is uniformly continuous, so there exists some $\delta > 0$ such that $\|\gamma(s) - \gamma(t)\| < \varepsilon/M$ whenever $s, t \in [0,1]$ and $|s-t| < \delta$. The desired conclusion follows. □

Proof of Theorem 2.46. Let $\varepsilon \in (0, \ln(2))$ be as in Lemma 2.44, and let $\alpha\colon [0,1] \to G$ be a path from I to A. Since $\exp(B_\varepsilon(0))$ is an open neighbourhood of I, we can use Lemma 2.48 to obtain some $\delta > 0$ such that $|s-t| < \delta$ and $s, t \in [0,1]$ implies $\alpha(s)\alpha(t)^{-1} \in \exp(B_\varepsilon(0))$. If we now pick $N \in \mathbb{Z}^+$ such that $1/N < \delta$, we obtain

$$A = \left(\alpha(1)\,\alpha\Big(\tfrac{N-1}{N}\Big)^{-1}\right)\left(\alpha\Big(\tfrac{N-1}{N}\Big)\,\alpha\Big(\tfrac{N-2}{N}\Big)^{-1}\right)\cdots\left(\alpha\Big(\tfrac{1}{N}\Big)\,\alpha(0)^{-1}\right).$$

Because each factor belongs to $G \cap \exp(B_\varepsilon(0))$, we can define

$$X_k = \log\left(\alpha\Big(\tfrac{N-k+1}{N}\Big)\,\alpha\Big(\tfrac{N-k}{N}\Big)^{-1}\right) \in \mathfrak{g}$$

for every $k \in \{1, \ldots, N\}$, and thus obtain

$$A = \exp(X_1)\exp(X_2)\cdots\exp(X_N)\,.$$ □

2.4 Matrix Lie groups as manifolds

In this section, we will prove two geometric results for matrix Lie groups: that every matrix Lie group can be given the structure of a Lie group, and that the Lie algebra of a matrix Lie group is the tangent space at the identity.

We begin by showing that $\mathbf{GL}_n(\mathbb{C})$ is a Lie group, and recall from Section 2.1 that we have given $\mathbb{C}^{n\times n}$ the structure of a $2n^2$-dimensional smooth manifold via the identification $\mathbb{C}^{n\times n} \cong \mathbb{R}^{2n^2}$.

Proposition 2.49. *The set* $\mathbf{GL}_n(\mathbb{C})$ *is a submanifold of* $\mathbb{C}^{n\times n}$*, and if equipped with matrix multiplication, the subspace topology and the induced smooth structure, it becomes a Lie group.*

Proof. Note that $\det\colon \mathbb{C}^{n\times n} \to \mathbb{C}$ is continuous and that $\mathbb{C}^\times$ is an open subset of $\mathbb{C}$, which implies that $\mathbf{GL}_n(\mathbb{C}) = \det^{-1}(\mathbb{C}^\times)$ is open. An open subset of a smooth manifold is a submanifold of the same dimension, so $\mathbf{GL}_n(\mathbb{C}) \subseteq \mathbb{C}^{n\times n}$ is a $2n^2$-dimensional submanifold of $\mathbb{C}^{n\times n}$. Next, note that both matrix multiplication and matrix inversion are smooth maps on $\mathbf{GL}_n(\mathbb{C})$; if $\mathbb{C}^{n\times n}$ is identified with $\mathbb{R}^{2n^2}$, each component of both matrix multiplication and matrix inversion will be rational expressions in the entries of the inputs. □

Theorem 2.50. *Let* $G \subseteq \mathbf{GL}_n(\mathbb{C})$ *be a matrix Lie group with Lie algebra* $\mathfrak{g}$*, such that* $\dim_{\mathbb{R}}(\mathfrak{g}) = k$*. Then* G *is a* k*-dimensional submanifold of* $\mathbf{GL}_n(\mathbb{C})$*.*

Proof. We start by finding a slice chart at the identity. Let $\varepsilon \in (0, \ln(2))$ be such that Lemma 2.44 holds. Then $U = \exp(B_\varepsilon(0))$ is an open neighbourhood of I in $\mathbf{GL}_n(\mathbb{C})$, and since $\varepsilon < \ln(2)$, the map $\log\big|_U\colon U \to B_\varepsilon(0)$ will be a homeomorphism. Pick an orthonormal $\mathbb{R}$-basis $\{e_1, \ldots, e_{2n^2}\}$ for $\mathbb{C}^{n\times n} = \mathfrak{g} \oplus \mathfrak{g}^\perp$, such that the subset $\{e_1, \ldots, e_k\}$ is an orthonormal basis for $\mathfrak{g}$. Let $\psi\colon \mathbb{C}^{n\times n} \to \mathbb{R}^{2n^2}$ be the associated vector space isomorphism. Then setting $V = \psi(B_\varepsilon(0))$ and $x = \psi \circ \log\big|_U\colon U \to V$ turns (U, x) into a slice chart of G in $\mathbf{GL}_n(\mathbb{C})$. Indeed, it is clear that (U, x) is a chart of $\mathbf{GL}_n(\mathbb{C})$, and by Lemma 2.44, we will have $\log(U \cap G) = B_\varepsilon(0) \cap \mathfrak{g}$, which implies that $x(U \cap G) = x(U) \cap (\mathbb{R}^k \oplus \{0\})$.

We can now use this chart to construct a slice chart at an arbitrary element $p \in G$. For any $X \in \mathbb{C}^{n\times n}$, let $L_X\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ denote multiplication by X from the left, and set $U_p = L_p(U)$, $V_p = V$ and $x_p = x \circ L_{p^{-1}}\colon U_p \to V_p$. Since multiplication by an invertible matrix from the left is a homeomoprhism in $\mathbb{C}^{n\times n}$, it is clear that U_p will be open and that x_p will be a homeomorphism. It is also clear that we have $x_p(U_p \cap G) = x_p(U_p) \cap (\mathbb{R}^k \oplus \{0\})$. □

The desired result, that every matrix Lie group can be given the structure of a Lie group, now follows readily. This establishes that "matrix Lie groups" is indeed an appropriate term for the closed linear groups we are studying.

Corollary 2.51. Let $G \subseteq \mathbf{GL}_n(\mathbb{C})$ be a matrix Lie group. Then $(G, \cdot, \tau, \widehat{\mathcal{A}})$, where $\cdot$ is matrix multiplication on G, τ is the subspace topology in $\mathbf{GL}_n(\mathbb{C})$, and $\widehat{\mathcal{A}}$ is the induced structure in $\mathbf{GL}_n(\mathbb{C})$, is a Lie group.

Proof. It is clear from Theorem 2.50 that $(G, \tau, \widehat{\mathcal{A}})$ is a smooth manifold. The smoothness of the maps $(g, h) \mapsto g \cdot h$ and $g \mapsto g^{-1}$ is a direct consequence of Theorem 1.6 and Proposition 2.49. □

Next, we show that the Lie algebra of a matrix Lie group can be thought of as the tangent space at the identity.

Theorem 2.52. *Let $G \subseteq \mathbf{GL}_n(\mathbb{C})$ be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then $\mathfrak{g} = T_I G$, where*

$$T_I G = \left\{ \gamma'(0) : \begin{array}{c} \gamma\colon (a,b) \to G \text{ smooth curve} \\ \text{with } 0 \in (a,b) \text{ and } \gamma(0) = I \end{array} \right\}.$$

Proof. $\boxed{\subseteq}$ Let $X \in \mathfrak{g}$. Clearly, $\gamma\colon \mathbb{R} \to G$ with $\gamma(t) = \exp(tX)$ is a smooth curve in G that satisfies $\gamma(0) = I$ and $\gamma'(0) = X$. This implies that $X \in T_I G$.

$\boxed{\supseteq}$ Let $\gamma\colon (a,b) \to G$ be a smooth curve in G such that $0 \in (a,b)$ and $\gamma(0) = I$. We want to show that $\gamma'(0) \in \mathfrak{g}$. Since γ is continuous, there exists some $\varepsilon > 0$ such that $(-\varepsilon, \varepsilon) \subseteq (a,b)$ and $\gamma((-\varepsilon,\varepsilon))$ is contained in a neighbourhood of I where Lemma 2.44 holds. Note that for $|t| < \varepsilon$, we can write $\gamma(t) = \exp(\delta(t))$, where $\delta = \log \circ \gamma\colon (-\varepsilon,\varepsilon) \to \mathfrak{g}$ is a well-defined map by Lemma 2.44. By Theorem 2.22, δ is smooth and since $\mathfrak{g}$ is a subspace of the finite-dimensional vector space $\mathbb{C}^{n\times n}$ we must have $\delta'(t) \in \mathfrak{g}$ for all $|t| < \varepsilon$. By using the product rule, and by noticing that $\delta(0) = 0$, we get

$$\begin{aligned} \gamma'(0) &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{k=0}^{\infty} \frac{\delta(t)^k}{k!} \right) \Bigg|_{t=0} = \sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}t} \frac{\delta(t)^k}{k!} \Bigg|_{t=0} \\ &= \sum_{k=1}^{\infty} \sum_{l=0}^{k-1} \frac{\delta(0)^l \cdot \delta'(0) \cdot \delta(0)^{k-l-1}}{k!} = \delta'(0)\,, \end{aligned}$$

from which we conclude that $\gamma'(0) \in \mathfrak{g}$. □

2.5 Lie group and Lie algebra homomorphisms

The following gives a notion of morphisms which turns the class of Lie groups into a category.

Definition 2.53. Let G and H be Lie groups. A map $\Phi\colon G \to H$ that is both continuous and a group homomorphism is called a ***Lie group homomorphism***. An invertible Lie group homomorphism Φ such that Φ^{-1} is continuous is called a ***Lie group isomorphism***.

Remark 2.54. It can be shown (see Proposition 3.12 in [Bt85]) that every Lie group homomorphism is smooth, and thus a morphism not only in the categories of groups and topological spaces, but also in the category of smooth manifolds. In Theorem 2.59 we give an elementary proof of this in the context of matrix Lie groups.

We also define a notion of morphisms for Lie algebras.

Definition 2.55. Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras over the same field. A linear map $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ is said to be a ***Lie algebra homomorphism*** if $\varphi([x,y]_{\mathfrak{g}}) = [\varphi(x), \varphi(y)]_{\mathfrak{h}}$ for all $x, y \in \mathfrak{g}$. An invertible Lie algebra homomorphism is called a ***Lie algebra isomorphism***.

Theorem 2.56. *Let G and H be matrix Lie groups, with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Let $\Phi\colon G \to H$ be a Lie group homomorphism. Then there exists a unique $\mathbb{R}$-linear map $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ that makes the following diagram commute:*

$$\begin{array}{ccc} G & \xrightarrow{\ \Phi\ } & H \\ {\scriptstyle\exp}\uparrow & & \uparrow{\scriptstyle\exp} \\ \mathfrak{g} & \xrightarrow{\ \varphi\ } & \mathfrak{h} \end{array}$$

Furthermore, the map φ will satisfy the following properties:

(i) $\varphi(X) = \frac{\mathrm{d}}{\mathrm{d}t}\Phi(\exp(tX))\Big|_{t=0}$ *for all* $X \in \mathfrak{g}$.

(ii) $\varphi(AXA^{-1}) = \Phi(A)\varphi(X)\Phi(A)^{-1}$ *for all* $A \in G$, $X \in \mathfrak{g}$.

(iii) $\varphi([X,Y]) = [\varphi(X), \varphi(Y)]$ *for all* $X, Y \in \mathfrak{g}$.

Proof. We will divide the proof into several smaller steps.

Step 1: Uniqueness.
If $\varphi, \psi\colon \mathfrak{g} \to \mathfrak{h}$ both were $\mathbb{R}$-linear maps such that the diagram commutes, we would, for every $X \in \mathfrak{g}$ and every $t \in \mathbb{R}$, have

$$\exp(t\varphi(X)) = \exp(\varphi(tX)) = \Phi(\exp(tX)) = \exp(\psi(tX)) = \exp(t\psi(X))\,.$$

Differentiation at $t = 0$ on both sides would then give $\varphi(X) = \psi(X)$.

Step 2: Existence of a map $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ that makes the diagram commute.
Let $X \in \mathfrak{g}$ and form the map $t \mapsto \Phi(\exp(tX))$. Notice that this is a one-parameter subgroup of H, which by Theorem 2.32, implies that there exists a unique $Z \in \mathfrak{h}$ such that $\Phi(\exp(tX)) = \exp(tZ)$ for all $t \in \mathbb{R}$. This gives us a well-defined mapping $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ with $X \mapsto Z$, with the property that

$$\Phi(\exp(tX)) = \exp(t\varphi(X)) \tag{2.4}$$

for all $X \in \mathfrak{g}$ and $t \in \mathbb{R}$. Setting $t = 1$ shows that φ makes the diagram commute.

Step 3: Linearity of φ.
We start by showing that φ is homogeneous. Let $s \in \mathbb{R}$ and $X \in \mathfrak{g}$. Then for any $t \in \mathbb{R}$, the property (2.4) of $\varphi(X)$ gives

$$\Phi\Big(\exp(t(sX))\Big) = \Phi\Big(\exp((ts)X)\Big) = \exp\Big((ts)\varphi(X)\Big) = \exp\Big(t(s\varphi(X))\Big)\,,$$

i.e. $s\varphi(X)$ has the uniquely defining property of $\varphi(sX)$.

Next, we show that φ is additive. Let $X, Y \in \mathfrak{g}$. Using Proposition 2.24 twice, as well as the fact that Φ is a continuous group homomorphism, we get

$$\begin{aligned}
\exp\Big(t\varphi(X+Y)\Big) &= \Phi(\exp(t(X+Y)) = \Phi(\exp(tX+tY)) \\
&= \Phi\left(\lim_{N\to\infty}\left(\exp\left(\tfrac{tX}{N}\right)\exp\left(\tfrac{tY}{N}\right)\right)^N\right) \\
&= \lim_{N\to\infty}\left(\Phi\left(\exp\left(\tfrac{tX}{N}\right)\right)\Phi\left(\exp\left(\tfrac{tY}{N}\right)\right)\right)^N \\
&= \lim_{N\to\infty}\left(\exp\left(\tfrac{t\varphi(X)}{N}\right)\exp\left(\tfrac{t\varphi(Y)}{N}\right)\right)^N \\
&= \exp(t\varphi(X)+t\varphi(Y)) = \exp(t(\varphi(X)+\varphi(Y)))\,.
\end{aligned}$$

Differentiation at $t=0$ gives $\varphi(X+Y) = \varphi(X)+\varphi(Y)$.

Step 4: Additional properties of φ.

(i) This follows directly from differentiation of (2.4) at $t=0$.

(ii) Let $A \in G$ and $X \in \mathfrak{g}$. Then, for any $t \in \mathbb{R}$, Proposition 2.16(v) and the fact that Φ is a group homomorphism, gives

$$\begin{aligned}
\exp(t\varphi(AXA^{-1})) &= \exp(\varphi(A(tX)A^{-1})) \\
&= \Phi(\exp(A(tX)A^{-1})) \\
&= \Phi(A\exp(tX)A^{-1}) \\
&= \Phi(A)\Phi(\exp(tX))\Phi(A)^{-1} \\
&= \Phi(A)\exp(t\varphi(X))\Phi(A)^{-1}\,.
\end{aligned}$$

Differentiation at $t=0$ gives the desired equality.

(iii) Let $X, Y \in \mathfrak{g}$. Using the product rule, the fact that φ is linear (and hence also continuous), as well as (iv), we then get

$$\begin{aligned}
\varphi([X,Y]) &= \varphi\Big(\tfrac{\mathrm{d}}{\mathrm{d}t}\exp(tX)Y\exp(-tX)\Big|_{t=0}\Big) \\
&= \varphi\left(\lim_{h\to 0}\frac{\exp(hX)Y\exp(-hX)-Y}{h}\right) \\
&= \lim_{h\to 0}\frac{\varphi\left(\exp(hX)Y\exp(-hX)\right)-\varphi(Y)}{h} \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\,\varphi(\exp(tX)Y\exp(-tX))\Big|_{t=0} \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\Phi(\exp(tX))\varphi(Y)\Phi(\exp(-tX))\Big|_{t=0} \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\exp(t\varphi(X))\varphi(Y)\exp(-t\varphi(X))\Big|_{t=0} \\
&= [\varphi(X),\varphi(Y)]\,.
\end{aligned}$$

□

A consequence of (iii) in the theorem above, is that the map $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. We will call it the ***induced Lie algebra homomorphism*** associated to Φ, and inspired by (i), we will give it the differential notation $\mathrm{d}\Phi_I$.

Theorem 2.57. *Let G, H and K be matrix Lie groups, let $\Phi\colon G \to H$ and $\Psi\colon H \to K$ be Lie group homomorphisms with $\varphi = \mathrm{d}\Phi_I$ and $\psi = \mathrm{d}\Psi_I$. Let $\Lambda = \Psi \circ \Phi$. Then $\mathrm{d}\Lambda_I = \psi \circ \varphi$.*

Proof. Let X belong to the Lie algebra of G. Then, for any $t \in \mathbb{R}$, it holds that

$$\Lambda(\exp(tX)) = \Psi(\Phi(\exp(tX))) = \Psi(\exp(t\varphi(X))) = \exp(t\psi(\varphi(X)))\,.$$

Differentiation at $t = 0$ then gives $\mathrm{d}\Lambda_I(X) = \psi(\varphi(X))$, as required. □

Remark 2.58. A consequence of the above theorem is that we have a functor from the category of matrix Lie groups to the category of real Lie algebras, that sends every matrix Lie group to its Lie algebra, and every Lie group homomorphism to the induced Lie algebra homomorphism.

We end this section by proving that for group homomorphisms between matrix Lie groups, continuity is enough to guarantee smoothness.

Theorem 2.59. *Let $\Phi\colon G \to H$ be a continuous group homomorphism between matrix Lie groups $G \subseteq \mathbf{GL}_n(\mathbb{C})$ and $H \subseteq \mathbf{GL}_m(\mathbb{C})$. Then Φ is smooth.*

Proof. Let $U \subseteq G$ be as in Lemma 2.45. Then for any $A \in U$, it holds that $A = \exp(\log(A))$, so that

$$\Phi(A) = \Phi(\exp(\log(A))) = \exp(\varphi(\log(A)))\,,$$

where $\varphi = \mathrm{d}\Phi_I$. This shows that

$$\Phi\big|_U = \exp \circ \varphi \circ \log\big|_U\,,$$

which is a composition of smooth maps and thus smooth.

Now, let $p \in G$ be arbitrary. Since left-multiplication by p is a homeomorphism, it is clear that pU is open, and for any $A \in pU$, it holds that

$$\Phi(A) = \Phi(pp^{-1}A) = \Phi(p)\Phi(p^{-1}A) = \Phi(p)\Phi\big|_U(p^{-1}A)\,,$$

so that

$$\Phi\big|_{pU} = L_{\Phi(p)} \circ \Phi\big|_U \circ L_{p^{-1}}\,,$$

where $L_{\Phi(p)}$ is left-translation by $\Phi(p)$ in H, and $L_{p^{-1}}$ is left translation by p^{-1} in G. Hence, $\Phi|_{pU}$ is also a composition of smooth maps and therefore smooth.

Because smoothness is a local property, and because we have shown that Φ is smooth on every element of a smooth atlas on G, we conclude that Φ is smooth on all of G. □

2.6 Lifting Lie algebra homomorphisms

Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. We know from Theorem 2.56 that every Lie group homomorphism $\Phi\colon G \to H$ induces a Lie algebra homomorphism $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ such that $\Phi(\exp(X)) = \exp(\varphi(X))$. This makes it possible to translate knowledge and questions about Lie groups to the corresponding Lie algebras. However, to be able to translate back, we need the following partial converse of Theorem 2.56.

Theorem 2.60. *Let G and H be as above. If G is path-connected and simply connected, then for every Lie algebra homomorphism $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ there exists a unique Lie group homomorphism $\Phi\colon G \to H$ such that $\mathrm{d}\Phi_I = \varphi$, i.e. $\Phi(\exp(X)) = \exp(\varphi(X))$ for every $X \in \mathfrak{g}$.*

The proof that we give here is inspired by Section 5.7 of [Hal15]. It relies on the so-called Baker–Campbell–Hausdorff formula, which tells us something about $\log(\exp(X)\exp(Y))$ in the general case when X and Y do not necessarily commute, so that $\log(\exp(X)\exp(Y)) = X + Y$ does not necessarily hold.

Theorem 2.61. *There exists some $\varepsilon > 0$ such that if $X, Y \in \mathbb{C}^{n\times n}$ are such that $\|X\|, \|Y\| < \varepsilon$, then $\log(\exp(X)\exp(Y))$ can be expressed as a convergent series:*

$$\log(\exp(X)\exp(Y)) = X + Y + \tfrac{1}{2}[X,Y] + \cdots \tag{2.5}$$

where each term consists of nested Lie brackets of X and Y.

We omit the proof of this fact, and refer to Chapter 7 of [Sti08].

Proof of Theorem 2.60. We will divide the proof into several smaller steps.

Step 1: Define Φ locally.
Let $\varepsilon > 0$ be small enough so that for any $A, B \in U$, where

$$U = \{A \in G : \|A - I\| < 1 \text{ and } \|\log(A)\| < \varepsilon\},$$

we have that $\log(A) \in \mathfrak{g}$ (by Lemma 2.44). Also make $\varepsilon > 0$ small enough for U to be path-connected and for formula (2.5) to hold for $X = \log(A)$ and $Y = \log(B)$, as well as for $\varphi(X)$ and $\varphi(Y)$.

We can then define a map $f\colon U \to H$ by setting $f = \exp \circ \varphi \circ \log|_U$. For each $A, B \in U$ with $X = \log(A)$ and $Y = \log(B)$, such that $AB \in U$, we then get $f(A) = \exp(\varphi(X))$, $f(B) = \exp(\varphi(Y))$ and

$$\begin{aligned} f(AB) &= f(\exp(X)\exp(Y)) = \exp(\varphi(\log(\exp(X)\exp(Y)))) \\ &= \exp(\varphi(X + Y + \tfrac{1}{2}[X,Y] + \cdots)) \\ &= \exp(\varphi(X) + \varphi(Y) + \tfrac{1}{2}[\varphi(X),\varphi(Y)] + \cdots) \\ &= \exp(\varphi(X))\exp(\varphi(Y)) = f(A)f(B), \end{aligned}$$

which shows that f locally is a group homomorphism. Note that we applied (2.5) two times: one time in G and one time in H. Also note that we in the third row used the fact that φ is linear (and thus also continuous) and therefore can be applied term-wise to the series.

Step 2: Define Φ globally.
We will now try to extend f to a global Lie group homomorphism $\Phi\colon G \to H$. For this end, fix $A \in G$. Since G is path-connected, there exists a path $\alpha\colon [0,1] \to G$ such that $\alpha(0) = I$ and $\alpha(1) = A$. Let $\{t_i\}_{i=0}^N$ be a partition of $[0,1]$ with

$$0 = t_0 < t_1 < \cdots < t_N = 1$$

such that, for any $i \in \{0, \ldots, N-1\}$, it holds that $s, t \in [t_i, t_{i+1}]$ implies $\alpha(s)\alpha(t)^{-1} \in U$. We will call a partition of the domain of a curve in G that satisfies this criterion a "good" partition. By Lemma 2.48, there exists at least one such good partition. Given a good partition, we can express A as a product of elements in U:

$$A = [\alpha(1)\alpha(t_{N-1})^{-1}][\alpha(t_{N-1})\alpha(t_{N-2})^{-1}]\cdots[\alpha(t_1)\alpha(0)^{-1}]\,,$$

where each factor indeed belongs to U. We can then try to define $\Phi(A)$ by

$$\Phi(A) = f(\alpha(1)\alpha(t_{N-1})^{-1})f(\alpha(t_{N-1})\alpha(t_{N-2})^{-1})\cdots f(\alpha(t_1)\alpha(0)^{-1})\,.$$

For this procedure to give a well-defined map $\Phi\colon G \to H$, we need to check that $\Phi(A)$ neither depends on the curve nor the good partition that we used. Until we know that, we will use the notation $\Phi(A, \alpha, \{t_i\}_{i=0}^N)$ to highlight these possible dependencies.

Step 3: $\Phi(A)$ is independent of the good partition.
We now show that the choice of a good partition does not affect the value of $\Phi(A)$, i.e. that given a path $\alpha\colon [0,1] \to G$ from I to A and any two good partitions $\{t_i\}_{i=0}^N$ and $\{s_j\}_{j=0}^M$ of $[0,1]$, $\Phi(A, \alpha, \{t_i\}_{i=0}^N) = \Phi(A, \alpha, \{s_j\}_{j=0}^M)$.

To do this, suppose that we add a point to the partition $\{t_i\}_{i=0}^N$, say $s \in (t_i, t_{i+1})$ for some $i \in \{0, 1, \ldots, N-1\}$. Clearly the new partition will still be good. Furthermore, notice that $\Phi(A, \alpha, \{t_i\} \cup \{s\})$ would be obtained by replacing the factor $f(\alpha(t_{i+1})\alpha(t_i)^{-1})$ in $\Phi(A, \alpha, \{t_i\})$ by

$$f(\alpha(t_{i+1})\alpha(s)^{-1})f(\alpha(s)\alpha(t_i)^{-1})\,.$$

But since clearly $\alpha(t_{i+1})\alpha(s)^{-1}$, $\alpha(s)\alpha(t_i)^{-1}$ and $\alpha(t_{i+1})\alpha(t_i)^{-1}$ belong to U, we can use the fact that f is locally a group homomorphism to deduce that

$$f(\alpha(t_{i+1})\alpha(s)^{-1})f(\alpha(s)\alpha(t_i)^{-1}) = f(\alpha(t_{i+1})\alpha(t_i)^{-1})\,.$$

Hence, $\Phi(A, \alpha, \{t_i\} \cup \{s\}) = \Phi(A, \alpha, \{t_i\})$, so refining a good partition by adding a point (or for that matter, any finite number of points) does not change the value of $\Phi(A)$. Since $\{t_i\}$ and $\{s_j\}$ have a common refinement, namely $\{t_i\} \cup \{s_j\}$, we conclude that $\Phi(A, \alpha, \{t_i\}_{i=0}^N) = \Phi(A, \alpha, \{s_j\}_{j=0}^M)$.

Step 4: $\Phi(A)$ is independent of the curve.
Suppose that $\alpha_0, \alpha_1 \colon [0,1] \to G$ both are paths from I to A. We want to show that $\Phi(A, \alpha_0) = \Phi(A, \alpha_1)$.

Since G is simply connected, we know that α_0 and α_1 are path-homotopic (see for example §52 in [Mun00]), i.e. there exists a continuous map $H \colon [0,1] \times [0,1] \to G$ such that $H(0, \cdot) = \alpha_0$, $H(1, \cdot) = \alpha_1$, $H(\cdot, 0) \equiv I$ and $H(\cdot, 1) \equiv A$. We will now use this map H to create a finite sequence of curves that represent stepwise deformations of α_0 into α_1, such that each deformation takes place on a sufficiently small portion of the curve for $\Phi(A)$ to remain unchanged.

To accomplish this, we note (by an analogous argument to that in the proof of Lemma 2.48) that there exists a $\delta > 0$ such that $|s - s'|, |t - t'| < \delta$ and $s, s', t, t' \in [0,1]$ implies $H(s,t)H(s',t')^{-1} \in U$. Let $N \in \mathbb{Z}^+$ be such that $2/N < \delta$. We will now use this N to form a family of curves $\beta_{i,j} \colon [0,1] \to G$, where for each $i \in \{0, \ldots, N-1\}$ and $j \in \{0, \ldots, N\}$ we set

$$\beta_{i,j}(t) = \begin{cases} H((i+1)/N, t) & \text{for } t \in [0, (j-1)/N] \\ H(i/N - t + j/N, t) & \text{for } t \in [(j-1)/N, j/N] \\ H(i/N, t) & \text{for } t \in [j/N, 1]\,. \end{cases}$$

We will deform α_0 to α_1 via the scheme

$$\begin{aligned} \alpha_0 &= \beta_{0,0} \to \beta_{0,1} \to \beta_{0,2} \to \cdots \to \beta_{0,N} \to \\ &\to \beta_{1,0} \to \beta_{1,1} \to \beta_{1,2} \to \cdots \to \beta_{1,N} \to \\ &\to \cdots \to \\ &\to \beta_{N-1,0} \to \beta_{N-1,1} \to \beta_{N-1,2} \to \cdots \to \beta_{N-1,N} \to \alpha_1\,. \end{aligned}$$

We now claim that $\Phi(A, \beta_{i,j}) = \Phi(A, \beta_{i,j+1})$ for all $i \in \{0, \ldots, N-1\}$ and $j \in \{0, \ldots, N-1\}$. This follows from the fact that

$$\left\{0, \tfrac{1}{N}, \tfrac{2}{N}, \ldots, \tfrac{j-1}{N}, \tfrac{j+1}{N}, \ldots, 1\right\}$$

is a good partition for both $\beta_{i,j}$ and $\beta_{i,j+1}$, and that $\beta_{i,j}$ and $\beta_{i,j+1}$ coincide for all these points. Since we, by Step 3, are free to choose whatever good partition we like, this shows that $\Phi(A, \beta_{i,j}) = \Phi(A, \beta_{i,j+1})$. Similarly, we note that $\Phi(A, \beta_{i,N}) = \Phi(A, \beta_{i+1,0})$, since

$$\left\{0, \tfrac{1}{N}, \tfrac{2}{N}, \ldots, \tfrac{N-1}{N}, 1\right\}$$

is a good partition for both $\beta_{i,N}$ and $\beta_{i+1,0}$, such that $\beta_{i,N}$ and $\beta_{i+1,0}$ coincide on all partition points. This shows that in every step of our deformation scheme, $\Phi(A)$ remains constant, and we therefore conclude that $\Phi(A, \alpha_0) = \Phi(A, \alpha_1)$, and that the procedure from Step 1 gives a well-defined map $\Phi \colon G \to H$.

Step 5: Φ is a group homomorphism.
Let $A, B \in G$. Let $\alpha\colon [0,1] \to G$ be a path from I to A with a good partition $\{s_i\}_{i=0}^{M}$, and let $\beta\colon [0,1] \to G$ be a path from I to B with a good partition $\{t_j\}_{j=0}^{N}$. We then have

$$\begin{aligned}\Phi(A) &= f(\alpha(1)\alpha(t_{M-1})^{-1}) \cdots f(\alpha(t_1)\alpha(0)^{-1})\\ \Phi(B) &= f(\beta(1)\beta(t_{N-1})^{-1}) \cdots f(\beta(t_1)\beta(0)^{-1})\,.\end{aligned}$$

We can also form the map $\gamma\colon [0,1] \to G$ defined by

$$\gamma(t) = \begin{cases} \beta(2t) & \text{for } t \in [0,1/2] \\ \alpha(2t-1)B & \text{for } t \in [1/2,1]\,, \end{cases}$$

which will be a path from I to AB. It is easy to see that the following partition for γ is good:

$$\left\{ \tfrac{s_0}{2}, \ldots, \tfrac{s_M}{2}, \tfrac{1+t_0}{2}, \ldots, \tfrac{1+t_N}{2} \right\}.$$

Hence,

$$\begin{aligned}\Phi(AB) &= f\Big(\gamma(\tfrac{1+t_N}{2})\gamma(\tfrac{1+t_{N-1}}{2})^{-1}\Big) \cdots f\Big(\gamma(\tfrac{1+t_1}{2})\gamma(\tfrac{1+t_0}{2})^{-1}\Big)\\ &\qquad \cdot f\Big(\gamma(\tfrac{s_M}{2})\gamma(\tfrac{s_{M-1}}{2})^{-1}\Big) \cdots f\Big(\gamma(\tfrac{s_1}{2})\gamma(\tfrac{s_0}{2})^{-1}\Big)\\ &= f(\beta(t_N)\beta(t_{N-1})^{-1}) \cdots f(\beta(t_1)\beta(t_0)^{-1})\\ &\qquad \cdot f(\alpha(s_M)\alpha(s_{M-1})^{-1}) \cdots f(\alpha(s_1)\alpha(s_0)^{-1})\\ &= \Phi(A)\Phi(B)\,.\end{aligned}$$

Step 6: $\Phi|_U = f$.
Let $A \in U$. Since U is path-connected, we can find a path $\alpha\colon [0,1] \to G$ from I to A, that lies entirely in U. Let $\{t_i\}_{i=0}^{N}$ be a good partition of this path α. We now claim that $\Phi(\alpha(t_j)) = f(\alpha(t_j))$ for all $j \in \{0, \ldots, N\}$, which for $j = N$ gives $\Phi(A) = f(A)$.

The case $j = 0$ is clear. For $j \in \{1, \ldots, N\}$, we note that the subset $\{t_i\}_{i=0}^{j}$ of our partition is itself a good partition of the path $\alpha_j = \alpha|_{[0,t_j]}\colon [0,t_j] \to G$ that goes from I to $\alpha(t_j)$. Hence,

$$\Phi(\alpha(t_j)) = f(\alpha(t_j)\alpha(t_{j-1})^{-1}) \cdots f(\alpha(t_1)\alpha(t_0)^{-1})\,.$$

We now proceed by induction on j. For $j = 1$, note that

$$\Phi(\alpha(t_1)) = f(\alpha(t_1)\alpha(0)^{-1}) = f(\alpha(t_1))\,.$$

Now assume that $\Phi(\alpha(t_j)) = f(\alpha(t_j))$. Then it holds that

$$\begin{aligned}\Phi(\alpha(t_{j+1})) &= f(\alpha(t_{j+1})\alpha(t_j)^{-1}) \cdots f(\alpha(t_1)\alpha(0)^{-1}) \\ &= f(\alpha(t_{j+1})\alpha(t_j)^{-1})\Phi(\alpha(t_j)) \\ &= f(\alpha(t_{j+1})\alpha(t_j)^{-1})f(\alpha(t_j)) = f(\alpha(t_{j+1}))\,.\end{aligned}$$

Here, we used the inductive hypothesis in the second equality, and the fact that f is a local homomorphism in the last equality. The desired result follows.

Step 7: Φ is continuous.
We already know that Φ is continuous on the open set $U \subseteq G$, since $\Phi|_U = f$, which we recall is defined as a composition of continuous functions. Next, note that for any $p \in G$, the map $L_p \colon G \to G$ defined by $L_p(A) = pA$ is a homeomorphism. This implies that $pU = L_p(U)$ is open. It is furthermore clear that for $A \in pU$,

$$\Phi|_{pU}(A) = \Phi(p \cdot (p^{-1}A)) = \Phi(p)\Phi(p^{-1}A) = \Phi(p)\Phi|_U(p^{-1}A)\,,$$

i.e. $\Phi|_{pU} = L_{\Phi(p)} \circ f \circ L_{p^{-1}}|_{pU}$ is a composition of continuous maps and must therefore be continuous. We have now shown that Φ is continuous on an open neighborhood of each $p \in G$. Since continuity is a local property, we conclude that Φ is continuous on all of G.

Step 8: $\Phi(\exp(X)) = \exp(\varphi(X))$ for all $X \in \mathfrak{g}$.
For any $X \in \mathfrak{g}$, we can find some sufficiently large $m \in \mathbb{Z}^+$ such that $X/m \in U$. This gives

$$\begin{aligned}\Phi(\exp(X)) &= \Phi(\exp(X/m)^m) = \Phi(\exp(X/m))^m \\ &= f(\exp(X/m))^m = \exp(\varphi(X)/m)^m = \exp(\varphi(X))\,.\end{aligned}$$

Step 9: Φ is unique.
Suppose that $\Phi_1, \Phi_2 \colon G \to H$ both are Lie group homomorphisms such that $\Phi_1(\exp(X)) = \exp(\varphi(X)) = \Phi_2(\exp(X))$ for all $X \in \mathfrak{g}$. Let $A \in G$. We want to show that $\Phi_1(A) = \Phi_2(A)$. Since G is path-connected, Theorem 2.46 gives that $A = \exp(X_1) \cdots \exp(X_m)$ for some $X_1, \ldots, X_m \in \mathfrak{g}$. Hence,

$$\begin{aligned}\Phi_1(A) &= \Phi_1(\exp(X_1) \cdots \exp(X_m)) \\ &= \Phi_1(\exp(X_1)) \cdots \Phi_1(\exp(X_m)) \\ &= \exp(\varphi(X_1)) \cdots \exp(\varphi(X_m)) = \Phi_2(A)\,.\end{aligned}$$

This concludes the proof. ☐

Chapter 3

Basic representation theory

3.1 Definitions and examples

In many applications where Lie groups arise, they act continuously and linearly on some vector space. This kind of actions can be described by the notion of a Lie group representation.

Definition 3.1. A ***representation*** of a Lie group G is a pair (V, Π), where V is a finite-dimensional vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, and $\Pi \colon G \to \mathbf{GL}(V)$ is a smooth group homomorphism. The dimension $\dim(V)$ of the underlying vector space is called the ***dimension*** of (V, Π). If the map Π is injective, the representation is said to be ***faithful***.

Remark 3.2. For this definition to make sense, we need to have a topology and smooth structure on $\mathbf{GL}(V)$. Since we have limited ourselves to finite-dimensional real and complex vector spaces, we can identify V with $\mathbb{F}^n$, and $\mathbf{GL}(V)$ with the matrix Lie group $\mathbf{GL}_n(\mathbb{F})$, by picking a basis. (The choice of basis will not matter, since a change of basis corresponds to a diffeomorphism $\mathbf{GL}_n(\mathbb{F}) \to \mathbf{GL}_n(\mathbb{F})$.) Note that in the case of matrix Lie groups, this identification, together with Theorem 2.59, allows us to replace the requirement of Π being smooth by only requiring continuity.

Remark 3.3. Given a representation (V, Π) of a Lie group G, we will often use the notation $\Pi_g(v) = (\Pi(g))(v)$ for $g \in G$ and $v \in V$. Another common notation is $g.v = \Pi_g(v)$. This is motivated by the fact that the representation (V, Π) gives rise to a continuous linear group action $G \times V \to V$, defined by $g.v = \Pi_g(v)$.

Example 3.4. Every Lie group G admits at least one representation, namely the ***trivial representation***, given by (V, Π), where $V = \mathbb{C}$ and $\Pi \colon G \to \mathbf{GL}(\mathbb{C}) \cong \mathbb{C}^\times$ is defined by setting $\Pi(g) = 1$ for all $g \in G$.

Example 3.5. For any matrix Lie group $G \subseteq \mathbf{GL}_n(\mathbb{C})$, we can construct the (clearly faithful) ***standard representation*** (V, Π), where $V = \mathbb{C}^n$ and $\Pi \colon G \to \mathbf{GL}(\mathbb{C}^n) \cong \mathbf{GL}_n(\mathbb{C})$ is defined by setting $\Pi(A) = A$ for all $A \in G$.

Example 3.6. Sometimes it is useful to let a Lie group act on the underlying vector space of its own Lie algebra. Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. The ***adjoint representation*** of G is then given by the pair $(\mathfrak{g}, \mathrm{Ad})$, where the map $\mathrm{Ad}\colon G \to \mathbf{GL}(\mathfrak{g})$ sends each $A \in G$ to the map $\mathrm{Ad}_A\colon \mathfrak{g} \to \mathfrak{g}$, which in turn is defined by $\mathrm{Ad}_A(X) = AXA^{-1}$.

Proposition 2.29(iii) shows that Ad_A is well-defined, and it is easy to see that Ad_A is an invertible linear map, with inverse given by $\mathrm{Ad}_{A^{-1}}$. Hence, $\mathrm{Ad}_A \in \mathbf{GL}(\mathfrak{g})$. It is furthermore easy to see that $\mathrm{Ad}_A \circ \mathrm{Ad}_B = \mathrm{Ad}_{AB}$. Continuity follows from the fact that, for any $X \in \mathfrak{g}$, it holds that

$$\begin{aligned}\|\mathrm{Ad}_A(X) - \mathrm{Ad}_B(X)\| &= \|AXA^{-1} - BXB^{-1}\| \\ &= \|AXA^{-1} - BXA^{-1} + BXA^{-1} - BXB^{-1}\| \\ &\leqslant \|A - B\|\|X\|\|A^{-1}\| + \|B\|\|X\|\|A^{-1} - B^{-1}\|\,.\end{aligned}$$

Definition 3.7. An ***intertwining map*** (or ***morphism***) between representations (V, Π) and (W, Σ) of a Lie group G is a linear map $\psi\colon V \to W$, such that the diagram

$$\begin{array}{ccc} V & \xrightarrow{\ \psi\ } & W \\ {\scriptstyle \Pi_g}\downarrow & & \downarrow{\scriptstyle \Sigma_g} \\ V & \xrightarrow{\ \psi\ } & W \end{array}$$

commutes for all $g \in G$. An invertible intertwining map is called an ***isomorphism*** of representations.

Definition 3.8. Let G be a Lie group, and let (V, Π) be a representation of G. Let W be a linear subspace of V such that $\Pi_g(w) \in W$ for every $w \in W$ and $g \in G$. Then W is said to be a ***G-invariant*** with respect to (V, Π), and the pair $(W, \Pi|_W)$ is said to be a ***subrepresentation*** of (V, Π). A G-invariant subspace W is said to be ***nontrivial*** if $W \notin \{\{0\}, V\}$. A representation (V, Π) of a Lie group G is said to be ***irreducible*** if it gives rise to no nontrivial G-invariant subspaces.

Irreducible representations are particularly interesting, since they can be used as building blocks for more complicated representations. Indeed, for compact matrix Lie groups, we have the following deep result.

Theorem 3.9. *Let G be a compact matrix Lie group, and let (V, Π) be a finite-dimensional complex representation. Then there exists a finite family of irreducible complex representations $\{(V_i, \Pi_i)\}_{i=1}^n$ such that*

$$V = \bigoplus_{i=1}^{n} V_i \quad \text{and} \quad \Pi = \bigoplus_{i=1}^{n} \Pi_i\,.$$

We omit the proof, and refer to Section 4.4 of [Hal15].

We now go on to define a notion of Lie algebra representations.

Definition 3.10. A ***representation*** of a (real or complex) Lie algebra $\mathfrak{g}$ is a pair (V, π), where V is a (real or complex) vector space and $\pi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie algebra homomorphism. Here, $\mathfrak{gl}(V) = \mathrm{End}(V)$ is a Lie algebra with Lie bracket given by $[X, Y] = XY - YX$. The dimension $\dim(V)$ of V is called the ***dimension*** of (V, π). If π is injective, the representation is said to be ***faithful***. The concepts of intertwining maps, subrepresentations and irreducibility are defined analogously to the case of Lie groups.

The next theorem gives us a one-to-one correspondance between the representations of a matrix Lie group and the representations of its Lie algebra, provided that the Lie group is path-connected and simply connected.

Theorem 3.11. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then every representation (V, Π) of G gives rise to a Lie algebra representation (V, π) of $\mathfrak{g}$, where $\pi = \mathrm{d}\Pi_I$. Furthermore, if G is path-connected and simply connected, every representation (V, π) of $\mathfrak{g}$ gives rise to a representation (V, Π) of G, with $\mathrm{d}\Pi_I = \pi$.*

Proof. This is a direct consequence of Theorem 2.56 and Theorem 2.60. □

Example 3.12. Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$, and let $\mathrm{ad} = \mathrm{d}(\mathrm{Ad})_I\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ be the Lie algebra homomorphism induced by $\mathrm{Ad}\colon G \to \mathbf{GL}(\mathfrak{g})$, defined as above. It then holds that

$$\begin{aligned}
\mathrm{ad}_X(Y) &= \tfrac{\mathrm{d}}{\mathrm{d}t} \mathrm{Ad}_{\exp(tX)} \Big|_{t=0}(Y) \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t} \mathrm{Ad}_{\exp(tX)}(Y)\Big|_{t=0} \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t} \exp(tX) Y \exp(-tX)\Big|_{t=0} \\
&= [X, Y] \,.
\end{aligned}$$

This shows that $\mathrm{ad}_X = [X, \cdot]$. The pair $(\mathfrak{g}, \mathrm{ad})$ is called the ***adjoint representation*** of $\mathfrak{g}$.

3.2 A non-matrix Lie group

From Section 2.4, we know that all matrix Lie groups can be thought of as general Lie groups. Here, we prove that the converse is *not* true. The counterexample we will give is due to Birkhoff [Bir36], but our presentation and proof will follow the representation-theoretic approach used in [Hal15]. We start by considering the the product manifold $G = \mathbb{R} \times \mathbb{R} \times S^1$, equipped with the binary operation $*\colon G \times G \to G$ defined by

$$(x_1, y_1, u_1) * (x_2, y_2, u_2) = (x_1 + x_2, y_1 + y_2, e^{ix_1y_2}u_1u_2) \,.$$

Proposition 3.13. *The smooth manifold G with the operation $*$ defined as above is a Lie group.*

Proof. We first verify that $*$ is a group operation on G. Direct computation shows that it is associative:

$$\begin{aligned}
&[(x_1, y_1, u_1) * (x_2, y_2, u_2)] * (x_3, y_3, u_3) \\
&= (x_1 + x_2, y_1 + y_2, e^{ix_1 y_2} u_1 u_2) * (x_3, y_3, u_3) \\
&= \Big((x_1 + x_2) + x_3, (y_1 + y_2) + y_3, e^{i(x_1+x_2)y_3}(e^{ix_1 y_2} u_1 u_2) u_3\Big) \\
&= \Big(x_1 + x_2 + x_3, y_1 + y_2 + y_3, e^{ix_1 y_3} e^{ix_2 y_3} e^{ix_1 y_2} u_1 u_2 u_3\Big) \\
&= \Big(x_1 + (x_2 + x_3), y_1 + (y_2 + y_3), e^{ix_1(y_2+y_3)} u_1 e^{ix_2 y_3} u_2 u_3\Big) \\
&= (x_1, y_1, u_1) * (x_2 + x_3, y_2 + y_3, e^{ix_2 y_3} u_2 u_3) \\
&= (x_1, y_1, u_1) * [(x_2, y_2, u_2) * (x_3, y_3, u_3)] \,.
\end{aligned}$$

It is furthermore easy to show that $(0, 0, 1)$ is the identity element of G, and that the inverse of an arbitrary element (x, y, u) will be given by $(-x, -y, e^{ixy} u^{-1})$. That the group multiplication and the group inversion are smooth follows from Theorem 1.6 and the fact that they are restrictions of smooth maps $(\mathbb{R} \times \mathbb{R} \times \mathbb{C}) \times (\mathbb{R} \times \mathbb{R} \times \mathbb{C}) \to (\mathbb{R} \times \mathbb{R} \times \mathbb{C})$ and $\mathbb{R} \times \mathbb{R} \times \mathbb{C}^\times \to \mathbb{R} \times \mathbb{R} \times \mathbb{C}^\times$, respectively. □

Remark 3.14. It can be shown that G is isomorphic (as an abstract group) to the factor group H/N, where H denotes the Heisenberg group from Example 2.9, and

$$N = \left\{ \begin{pmatrix} 1 & 0 & n \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}.$$

We will now show that G cannot be realized as a matrix Lie group.

Proposition 3.15. *There exists no matrix Lie group that is isomorphic to G as Lie groups.*

Proof. The idea of this proof will be to use Example 3.5, where we saw that every matrix Lie group has a faithful finite-dimensional complex representation. As a consequence of this, any Lie group that is isomorphic to a matrix Lie group must also have faithful finite-dimensional complex representation. This, we will claim, is not the case for our Lie group G. To prove this, we let $\Sigma\colon G \to \mathbf{GL}(V)$ be an arbitrary finite-dimensional complex representation of G, and we will attempt to show that $\ker(\Sigma)$ is non-trivial. To get access to the tools that we have developed for representations of matrix Lie groups, we will employ the Heisenberg group H, as well as the map $\Phi\colon H \to G$ defined by

$$\Phi\colon \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \mapsto (a, c, e^{2i\pi \cdot b}) .$$

It is easy to verify that Φ is a continuous group homomorphism (it can, in fact, be identified with the natural projection $H \to H/N$). Hence, we can use our representation Σ to obtain a finite-dimensional representation $\Pi = \Sigma \circ \Phi\colon H \to \mathbf{GL}(V)$ of H. Clearly, $\ker(\Pi) = \Phi^{-1}(\ker(\Sigma)) \supseteq \ker(\Phi)$, so if we are able to show that $\ker(\Pi) \supsetneq \ker(\Phi)$, it will follow that $\ker(\Sigma)$ is non-trivial.

It is easy to see that

$$\ker(\Phi) = \left\{ \begin{pmatrix} 1 & 0 & n \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\} .$$

To investigate $\ker(\Pi)$, we pass to the Lie algebra level, and consider the induced Lie algebra homomorphism $\pi\colon \mathfrak{h} \to \mathfrak{gl}(V)$. It is easily seen that

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \; Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } Z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is a basis for $\mathfrak{h}$, and that $[X, Y] = Z$ whereas $[X, Z] = [Y, Z] = 0$. Based on this, we now make the following three claims.

Claim 1: $\exp(n\pi(Z)) = I$ *for all* $n \in \mathbb{Z}$.
This follows from the fact that

$$\exp(nZ) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & n \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & n \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \ker(\Phi) ,$$

so that $\exp(n\pi(Z)) = \Pi(\exp(nZ)) = I$.

Claim 2: $\pi(Z)$ *is nilpotent.*
It can be shown (for instance by using the Jordan canonical form) that nilpotency of a matrix is equivalent to 0 being the only eigenvalue. Let $\lambda \in \mathbb{C}$ be an eigenvalue of $\pi(Z)$. Then the associated eigenspace V_λ is invariant, not only under $\pi(Z)$ but also under $\pi(X)$ and $\pi(Y)$. To see this, note that for any $v \in V_\lambda$,

$$\pi(Z)\pi(X)v = \pi(X)\pi(Z)v = \pi(X)\lambda v = \lambda\pi(X)v ,$$

so $\pi(X)v \in V_\lambda$. Hence, V_λ is invariant under $\pi(X)$. That V_λ is invariant under $\pi(Y)$ is shown similarly.

Together with the fact that

$$\pi(Z) = \pi([X,Y]) = [\pi(X), \pi(Y)] = \pi(X)\pi(Y) - \pi(Y)\pi(X)\,,$$

we now have

$$\lambda\,\mathrm{id}_{V_\lambda} = \pi(Z)\big|_{V_\lambda} = \pi(X)\big|_{V_\lambda}\pi(Y)\big|_{V_\lambda} - \pi(Y)\big|_{V_\lambda}\pi(X)\big|_{V_\lambda}\,.$$

Taking the trace on both sides (recall that $\mathrm{trace}(AB) = \mathrm{trace}(BA)$) gives $\lambda \dim_{\mathbb{C}}(V_\lambda) = 0$, i.e. $\lambda = 0$.

Claim 3: $\Pi(\exp(t\pi(Z))) = I$ *for all* $t \in \mathbb{R}$.
Since $\pi(Z)$ is nilpotent, we know that the series expansion of $\exp(t\pi(Z))$ will terminate after a finite number of terms. Hence, if we pick a basis for V, the (i,j)-th entry of the matrix corresponding to $\exp(t\pi(Z))$ will be some polynomial $p_{ij}(t)$ in t. But from Claim 1, we know that for every $n \in \mathbb{Z}$ $\exp(n\pi(Z)) = I$, i.e. $p_{ij}(n) = \delta_{ij}$. Only constant polynomials can take a certain value infinitely many times, so this must mean that $p_{ij}(t) = \delta_{ij}$ for all $t \in \mathbb{R}$. From this we conclude that $\Pi(\exp(tZ)) = \exp(t\pi(Z)) = I$.

But then

$$\ker(\Pi) \supseteq \left\{ \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : t \in \mathbb{R} \right\} \supsetneq \left\{ \begin{pmatrix} 1 & 0 & n \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\} = \ker(\Phi)\,,$$

and we conclude that Σ cannot be faithful, and hence, that G cannot be isomorphic to a matrix Lie group. □

Chapter 4

More on SO(3) and SU(2)

In this chapter we will use the groups $\mathbf{SO}(3)$ and $\mathbf{SU}(2)$ as a case study to show how some of the theory developed in earlier chapters can be applied in practice. We first note a similarity between $\mathbf{SO}(3)$ and $\mathbf{SU}(2)$, namely that they have the same Lie algebra.

Proposition 4.1. *The Lie algebras* $\mathfrak{su}(2)$ *and* $\mathfrak{so}(3)$ *are isomorphic.*

Proof. From Example 2.40, we get that

$$\mathfrak{su}(2) = \left\{ \begin{pmatrix} xi & -a+bi \\ a+bi & -xi \end{pmatrix} : x, a, b \in \mathbb{R} \right\} .$$

It is easy to verify that a basis for $\mathfrak{su}(2)$ is given by $\{E_1, E_2, E_3\}$, where

$$E_1 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \ E_2 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \ \text{and} \ E_3 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

and that the following relations are satisfied:

$$[E_1, E_2] = E_3\,, \ [E_2, E_3] = E_1 \ \text{ and } \ [E_3, E_1] = E_2\,. \tag{4.1}$$

From Example 2.41, on the other hand, we get that

$$\mathfrak{so}(3) = \left\{ \begin{pmatrix} 0 & -a & b \\ a & 0 & -c \\ -b & c & 0 \end{pmatrix} : a, b, c \in \mathbb{R} \right\}$$

from which we can obtain a basis $\{F_1, F_2, F_3\}$ for $\mathfrak{so}(3)$, where

$$F_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \ F_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \ F_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The following commutation relations are easily verified:

$$[F_1, F_2] = F_3\,, \ [F_2, F_3] = F_1 \ \text{and} \ [F_3, F_1] = F_2\,. \tag{4.2}$$

As a consequence of the bilinearity and anticommutativity of Lie brackets, the commutation relations (4.1) and (4.2) completely determine the Lie bracket for $\mathfrak{su}(2)$ an $\mathfrak{so}(3)$, respectively. It is thus clear that we obtain a Lie algebra isomorphism $\mathfrak{su}(2) \to \mathfrak{so}(3)$ by sending E_i to F_i for $i \in \{1, 2, 3\}$ and extending linearly. □

A consequence of what we just showed, together with Theorem 2.50, is that both $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$ are three-dimensional manifolds. However, despite having isomorphic Lie algebras, they are not isomorphic as Lie groups. In fact, they are not even homeomorphic as topological spaces. In order to prove this, we will compare $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$ to two more easily understood topological "model spaces."

Proposition 4.2. *The topological space* $\mathbf{SU}(2)$ *is homeomorphic to* S^3.

Proof. Note that

$$\mathbf{SU}(2) = \left\{ \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix} : \begin{matrix} z, w \in \mathbb{C}\,, \\ |z|^2 + |w|^2 = 1 \end{matrix} \right\},$$

and that

$$S^3 = \{(a, b, c, d \in \mathbb{R} : a^2 + b^2 + c^2 + d^2 = 1\}\,.$$

It is then natural to form the map $f\colon S^3 \to \mathbf{SU}(2)$ with

$$f(a, b, c, d) = \begin{pmatrix} a + di & -b + ci \\ b + ci & a - di \end{pmatrix},$$

which clearly is a homeomorphism. □

Proposition 4.3 (Euler's rotation theorem)**.** *Every element of* $\mathbf{SO}(3)$ *corresponds to a rotation around some rotation axis.*

Proof. We start by showing that for each $A \in \mathbf{SO}(3)$, there exists some line through the origin that is fixed by A (this will turn out to be the rotation axis), i.e. some $v \in \mathbb{R}^3 \setminus \{0\}$ such that $Av = v$. This is equivalent to showing that 1 is an eigenvalue of A, i.e. that $\det(A - I) = 0$. This follows from the following computation:

$$\begin{aligned} \det(A - I) &= \det((A - I)^\top) = \det(A^\top - I) = \det(A^\top - A^\top A) \\ &= \det(A^\top (I - A)) = \det(A^\top) \det(I - A) \\ &= \det(A) \det(I - A) = \det(I - A) = -\det(A - I)\,. \end{aligned}$$

Let $v \in \mathbb{R}^3$ be such that $Av = v$, and assume that $|v| = 1$. We now let $u_1, u_2 \in (\mathbb{R}v)^\perp$ be such that $\{v, u_1, u_2\}$ is a positively oriented orthonormal basis for $\mathbb{R}^3$. Since orthogonality and orientation is preserved by A, it is clear that in this new basis, the linear transformation associated with A corresponds to a block matrix of the form

$$A' = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} ,$$

where the columns of $R \in \mathbb{R}^{2\times 2}$ must be orthogonal. Furthermore, since the determinant is basis independent, we have

$$1 = \det(A) = \det(A') = \det(R) ,$$

and so we conclude that $R \in \mathbf{SO}(2)$. It then follows from elementary linear algebra that

$$R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

for some $\theta \in \mathbb{R}$, so that

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

Thus it is clear that A corresponds to a rotation around $\mathbb{R}v$ by an angle of θ. Since we chose u_1 and u_2 so that $\{v, u_1, u_2\}$ is positively oriented, the direction of the rotation in $(\mathbb{R}v)^\perp$ will be given by the right-hand screw rule. □

We now construct a topological model of $\mathbf{SO}(3)$ by letting B^3 be the ball in $\mathbb{R}^3$ of radius π, and $\sim$ be the equivalence relation on B^3 defined by $u \sim u$ for all $u \in B^3$, and $u \sim -u$ if and only if $|u| = \pi$. Then the quotient space $B^3/\!\sim$ is the ball B^3 with antipodal points identified.

Proposition 4.4. *The topological space* $\mathbf{SO}(3)$ *is homeomorphic to* $B^3/\!\sim$*.*

Proof. We start by introducing a bit of notation. For each $v \in \mathbb{R}^3$ with $|v| = 1$ and each $\theta \in \mathbb{R}$, we will let $R_{v,\theta}$ denote the matrix in $\mathbb{R}^{3\times 3}$ that corresponds to a rotation around the axis $\mathbb{R}v$ by an angle of θ in the direction given by the right-hand screw rule. It is clear that every such $R_{v,\theta}$ belongs to $\mathbf{SO}(3)$, and together with Proposition 4.3 this gives that

$$\mathbf{SO}(3) = \{R_{v,\theta} : v \in \mathbb{R}^3,\ |v| = 1 \text{ and } \theta \in [-\pi, \pi]\} .$$

Since $R_{v,-\theta} = R_{-v,\theta}$ for every $v \in \mathbb{R}^3$ and $\theta \in \mathbb{R}$, we get that

$$\mathbf{SO}(3) = \{R_{v,\theta} : v \in \mathbb{R}^3,\ |v| = 1 \text{ and } \theta \in [0, \pi]\} .$$

Note that $I = R_{v,0}$ and $R_{v,\pi} = R_{-v,\pi}$ for all $v \in \mathbb{R}^3$, and that for each fixed $\theta \in (0, \pi)$, $R_{v,\theta}$ corresponds to a unique element in $\mathbf{SO}(3)$ for every choice of $v \in \mathbb{R}^3$ with $|v| = 1$. This makes it possible to construct a bijection $f \colon B^3/\!\sim \to \mathbf{SO}(3)$ given by

$$f(u) = \begin{cases} I & \text{if } u = 0 \\ R_{u/\|u\|, \|u\|} & \text{if } u \neq 0\,. \end{cases}$$

It can easily be verified that f is continuous. This, together with the fact that $B^3/\!\sim$ is compact (a quotient space of a compact space is always compact), and the fact that $\mathbf{SO}(3)$ is Hausdorff (a subspace of a Hausdorff space is always Hausdorff), implies that f is a homeomorphism (see Theorem 26.6 in [Mun00]). $\square$

Since S^3 is simply connected (see for instance [Lee12] for an elementary proof), it follows that $\mathbf{SU}(2)$ is simply connected. This is not the case for $\mathbf{SO}(3)$, as it can be shown (see Section 1.3 in [Hal15]) that $B^3/\!\sim$ is path-connected but not simply connected*. Since simple connectedness is a topological invariant, we conclude that $\mathbf{SO}(3)$ and $\mathbf{SU}(2)$ are not homeomorphic, and in particular not isomorphic as Lie groups.

Despite this fact, there is still a close relationship between $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$, both algebraically and topologically, originating from the fact that we can let $\mathbf{SU}(2)$ act as rotations on $\mathbb{R}^3$ in a 2-to-1 fashion.

Proposition 4.5. *There exists a surjective Lie group homomorphism* $\mathbf{SU}(2) \to \mathbf{SO}(3)$ *with kernel* $\{\pm I\}$.

Proof. In this proof we will use the 3-dimensional real vector space $\mathfrak{su}(2)$ as a model for $\mathbb{R}^3$, and we will consider the Lie group homomorphism $\mathrm{Ad} \colon \mathbf{SU}(2) \to \mathbf{GL}(\mathfrak{su}(2))$ with $\mathrm{Ad}_A(X) = AXA^{-1}$ that we introduced in Example 3.6. By picking $\mathscr{B} = \{2E_1, 2E_2, 2E_3\}$, with notation as in Proposition 4.1, as a basis for $\mathfrak{su}(2)$, we can identify Ad_A, where

$$A = \begin{pmatrix} a + di & -b + ci \\ b + ci & a - di \end{pmatrix}, \quad a^2 + b^2 + c^2 + d^2 = 1\,,$$

with the matrix

$$[\mathrm{Ad}_A]_{\mathscr{B}} = \begin{pmatrix} a^2 - b^2 + c^2 - d^2 & -2ad + 2bc & 2ab + 2cd \\ 2ad + 2bc & a^2 + b^2 - c^2 - d^2 & -2ac + 2bd \\ -2ab + 2cd & 2ac + 2bd & a^2 - b^2 - c^2 + d^2 \end{pmatrix}. \quad (4.3)$$

It is an easy computation to show that $[\mathrm{Ad}_A]_{\mathscr{B}} \in \mathbf{SO}(3)$. From this we conclude that $\mathrm{Ad} \colon \mathbf{SU}(2) \to \mathbf{GL}(\mathfrak{su}(2))$ corresponds to a Lie group homomorphism $\Phi \colon \mathbf{SU}(2) \to \mathbf{SO}(3)$ defined by $\Phi(A) = [\mathrm{Ad}_A]_{\mathscr{B}}$. From (4.3) it is

*To understand this intuitively, one can consider the loop in $B^3/\!\sim$ that goes from the north pole to the south pole in B^3, and attempt to come up with a contraction.

immediately clear that $\ker(\Phi) = \{I, -I\}$. To show that Φ is surjective, we use Proposition 4.3, which tells us that any element in $\mathbf{SO}(3)$ corresponds to a right-handed rotation by some angle θ around some axis $\mathbb{R}u$, where $u = (u_x, u_y, u_z)$ can be assumed to be of unit length. One can now verify that the corresponding rotation is given by $\Phi(A)$, where

$$A = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) + \sin\left(\frac{\theta}{2}\right) u_z i & -\sin\left(\frac{\theta}{2}\right) u_x + \sin\left(\frac{\theta}{2}\right) u_y i \\ \sin\left(\frac{\theta}{2}\right) u_x + \sin\left(\frac{\theta}{2}\right) u_y i & \cos\left(\frac{\theta}{2}\right) - \sin\left(\frac{\theta}{2}\right) u_z i \end{pmatrix}.$$

This can be done by showing first that Ad_A fixes the axis $\mathbb{R}u$, where u can be identified with the matrix

$$\begin{pmatrix} u_z i & -u_x + u_y i \\ u_x + u_y i & -u_z i \end{pmatrix} \in \mathfrak{su}(2)\,,$$

and then that Ad_A acts as a rotation in the 2-dimensional orthogonal complement $(\mathbb{R}u)^{\perp}$ in $\mathfrak{su}(2)$. The lengthy computations are left to the reader. □

Remark 4.6. It is worth noting that for the Lie group homomorphism $\Phi\colon \mathbf{SU}(2) \to \mathbf{SO}(3)$ that we just found, the induced Lie algebra isomorphism $\varphi\colon \mathfrak{su}(2) \to \mathfrak{so}(3)$ coincides with the Lie algebra isomorphism we constructed in Proposition 4.1. One way to show this is by direct computation, using the formula (4.3) above. Another approach is to note that

$$\mathrm{ad}_{E_1}(2E_1) = 0\,, \ \ \mathrm{ad}_{E_1}(2E_2) = 2E_3 \ \text{ and } \ \mathrm{ad}_{E_1}(2E_3) = -2E_2\,,$$

which implies $\varphi(E_1) = [\mathrm{ad}_{E_1}]_{\mathscr{B}} = F_1$. The equalities $\varphi(E_2) = F_2$ and $\varphi(E_3) = F_3$ can be shown in a similar fashion.

We now turn to the representation theory of $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$. For each $m \in \mathbb{Z}_0^+$, let V_m be the $(m+1)$-dimensional $\mathbb{C}$-vector space of homogeneous polynomials of degree m in two complex variables, i.e.

$$V_m = \left\{ \sum_{k=0}^{m} c_k z_1^{m-k} z_2^k : c_0, \ldots, c_m \in \mathbb{C} \right\} \subseteq \mathbb{C}[z_1, z_2]\,.$$

We will view each polynomial $p \in \mathbb{C}[z_1, z_2]$ as a function defined on $\mathbb{C}^2$, and construct a map $\Pi_m\colon \mathbf{SU}(2) \to \mathbf{GL}(V_m)$ by, for each $A \in \mathbf{SU}(2)$, letting $\Pi_m(A)\colon V_m \to V_m$ be defined by

$$[\Pi_m(A)p](z) = p(A^{-1}z)\,,$$

for $z = (z_1, z_2)^{\top} \in \mathbb{C}^2$.

Proposition 4.7. *For each $m \in \mathbb{Z}_0^+$, the pair (V_m, Π_m), defined as above, is a representation of* $\mathbf{SU}(2)$.

Proof. For an element

$$A = \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \in \mathbf{SU}(2)$$

it holds that

$$A^{-1}z = A^* z = \begin{pmatrix} \overline{\alpha} & \overline{\beta} \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \overline{\alpha} z_1 + \overline{\beta} z_2 \\ -\beta z_1 + \alpha z_2 \end{pmatrix},$$

so that for any $p(z) = \sum_{k=0}^m c_k z_1^{m-k} z_2^k \in V_m$, we get

$$(\Pi_m(A)p)(z) = p(A^{-1}z) = \sum_{k=0}^m c_k (\overline{\alpha} z_1 + \overline{\beta} z_2)^{m-k} (-\beta z_1 + \alpha z_2)^k, \qquad (4.4)$$

which (by using the binomial theorem) clearly can be seen to belong to V_m. Hence, $\Pi_m(A) \colon V_m \to V_m$ is a well-defined map. It is also clear that $\Pi_m(A)$ is linear. By noting that

$$\begin{aligned} (\Pi_m(A)\Pi_m(B)p)(z) &= (\Pi_m(B)p)(A^{-1}z) \\ &= p(B^{-1}A^{-1}z) \\ &= (\Pi_m(AB)p)(z), \end{aligned}$$

we conclude that that Π_m is a group homomorphism. This calculation also shows that for each $A \in \mathbf{SU}(2)$, $\Pi_m(A)$ is invertible with inverse $\Pi_m(A^{-1})$, and hence an element of $\mathbf{GL}(V_m)$. We finally note from (4.4) that Π is a continuous map. □

Remark 4.8. Note that (V_0, Π_0) is isomorphic to the trivial representation of $\mathbf{SU}(2)$, and that (V_1, Π_1) is isomorphic to the standard representation via the map $(c_0 z_1 + c_1 z_2) \mapsto (c_1, -c_0)$.

Given these representations of $\mathbf{SU}(2)$, we now investigate the corresponding Lie algebra representations of $\mathfrak{su}(2)$. By Theorem 3.11, we know that each representation (V_m, Π_m), defined as above, gives rise to a Lie algebra representation (V_m, π_m), where $\pi_m \colon \mathfrak{su}(2) \to \mathfrak{gl}(V_m)$ satisfies

$$\begin{aligned} (\pi_m(X)p)(z) &= \left(\frac{\mathrm{d}}{\mathrm{d}t} \Pi_m(\exp(tX))\Big|_{t=0} p \right)(z) \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \Pi_m\big(\exp(tX))p)(z)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} p(\exp(-tX)z)\Big|_{t=0} \end{aligned}$$

for all $X \in \mathfrak{su}(2)$ and $p \in V_m$.

We can obtain a more explicit formula by viewing $p(\exp(-tX)z)$ as a composition of an outer function $z \mapsto p(z)$ and an inner function $t \mapsto \exp(-tX)z$. If we let $z_1(t)$ and $z_2(t)$ denote the first and second component of $\exp(-tX)z \in \mathbb{C}^2$, respectively, the chain rule gives

$$\pi_m(X)p = \frac{\partial p}{\partial z_1} \cdot \frac{\mathrm{d}z_1}{\mathrm{d}t}\bigg|_{t=0} + \frac{\partial p}{\partial z_2} \cdot \frac{\mathrm{d}z_2}{\mathrm{d}t}\bigg|_{t=0} .$$

By the product rule, we have $\frac{\mathrm{d}}{\mathrm{d}t} \exp(-tX)z|_{t=0} = -Xz$, from which it follows that

$$\frac{\mathrm{d}z_1}{\mathrm{d}t}\bigg|_{t=0} = -X_{11}z_1 - X_{12}z_2 \quad \text{and} \quad \frac{\mathrm{d}z_2}{\mathrm{d}t}\bigg|_{t=0} = -X_{21}z_1 - X_{22}z_2 \, .$$

Hence,

$$\pi_m(X)p = -\frac{\partial p}{\partial z_1} \cdot (X_{11}z_1 + X_{12}z_2) - \frac{\partial p}{\partial z_2} \cdot (X_{21}z_1 + X_{22}z_2) \, .$$

We finally also investigate how these representations carry over to $\mathbf{SO}(3)$. As a first observation, note that each of the representations (V_m, π_m) of $\mathfrak{su}(2)$ from above gives rise to a representation (V_m, σ_m) of $\mathfrak{so}(3)$, where $\sigma_m = \pi_m \circ \varphi^{-1}$ and $\varphi\colon \mathfrak{su}(2) \to \mathfrak{so}(3)$ is the Lie algebra isomorphism discussed in Proposition 4.1. If $\mathbf{SO}(3)$ were simply connected, we would be able to lift each of these representations of $\mathfrak{so}(3)$ to representations of $\mathbf{SO}(3)$, by means of Theorem 3.11. But since $\mathbf{SO}(3)$ is not simply connected, we instead get the following result.

Proposition 4.9. *Let $m \in \mathbb{Z}_0^+$, and let $\sigma_m = \pi_m \circ \varphi^{-1}$. Then there exists a Lie group homomorphism $\Sigma_m\colon \mathbf{SO}(3) \to \mathbf{GL}(V_m)$, such that $\mathrm{d}(\Sigma_m)_I = \sigma_m$, if and only if m is even.*

Proof. We start by noting the main reason that the parity of m matters, namely that it affects the action of $-I \in \mathbf{SU}(2)$ on V_m. For any polynomial $p(z) = \sum_{k=0}^m c_k z_1^{m-k} z_2^k \in V_m$, it holds that

$$\begin{aligned}\Big(\Pi_m(-I)p\Big)(z) &= p((-I)^{-1}z) \\ &= \sum_{k=0}^{m} c_k(-z_1)^{m-k}(-z_2)^k \\ &= (-1)^m p(z) \, ,\end{aligned}$$

from which we conclude that

$$\Pi_m(-I) = \begin{cases} \mathrm{id}_{V_m} & \text{if } m \text{ is even} \\ -\mathrm{id}_{V_m} & \text{if } m \text{ is odd} \, . \end{cases}$$

We are now ready to prove the "only if" part of the proposition. Let $m \in \mathbb{Z}^+$ be odd and assume towards a contradiction that there exists a Lie group homomorphism $\Sigma_m \colon \mathbf{SO}(3) \to \mathbf{GL}(V_m)$ such that $\mathrm{d}(\Sigma_m)_I = \sigma_m$. It then holds that

$$\Sigma_m(\exp(X)) = \exp(\sigma_m(X)) = \exp(\pi_m(\varphi^{-1}(X))) = \Pi_m(\exp(\varphi^{-1}(X)))$$

for all $X \in \mathfrak{so}(3)$. In particular,

$$\Sigma_m(\exp(2\pi F_3)) = \Pi_m(\exp(\varphi^{-1}(2\pi F_3))) = \Pi_m(\exp(2\pi E_3)) \,.$$

However, direct computations show that $\exp(2\pi F_3) = I$ and $\exp(2\pi E_3) = -I$, so this would give $\Sigma_m(I) = \Pi_m(-I)$, or equivalently, $\mathrm{id}_{V_m} = -\mathrm{id}_{V_m}$, which is a contradiction.

To prove the "if" part, we let $m \in \mathbb{Z}_0^+$ be even, and attempt to construct a Lie group homomorphism $\Sigma_m \colon \mathbf{SO}(3) \to \mathbf{GL}(V_m)$ in the following way. For each $R \in \mathbf{SO}(3)$, Proposition 4.5 tells us that there exists some $A \in \mathbf{SU}(2)$ (unique up to sign) such that $\Phi(A) = \Phi(-A) = R$. Since m is even, it holds that

$$\Pi_m(-A) = \Pi_m(-I \cdot A) = \Pi_m(-I)\Pi_m(A) = \Pi_m(A) \,,$$

so we get a well-defined map by setting $\Sigma_m(R) = \Pi_m(A)$. It is easy to verify that this turns Σ_m into a group homomorphism. We also note that Σ_m is continuous in some neighbourhood U of the identity in $\mathbf{SO}(3)$, since we there can pick $A = \exp(\varphi^{-1}(\log(R)))$, so that we get $\Sigma_m|_U = \Pi_m \circ \exp \circ \varphi^{-1} \circ \log$, which is a composition of continuous maps and therefore continuous. By the same reasoning as in Step 7 of the proof of Theorem 2.60, this implies that Σ_m is continuous on all of $\mathbf{SO}(3)$. Finally note that $\Pi_m = \Sigma_m \circ \Phi$ by the construction of Σ_m. By Theorem 2.57, this gives $\pi_m = \mathrm{d}(\Sigma_m)_I \circ \varphi$, i.e. $\mathrm{d}(\Sigma_m)_I = \pi_m \circ \varphi^{-1} = \sigma_m$, as required. □

Remark 4.10. It can be shown (see Section 4.2 and 4.6 in [Hal15]) that the representations (V_m, π_m) for $m \in \mathbb{Z}_0^+$ are exactly all the irreducible representations (up to isomorphism) of $\mathfrak{su}(2)$. This, together with Theorem 2.46, can be used to show that we above have found (up to isomorphism) all the irreducible representations of $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$.

Bibliography

[Bir36] G. Birkhoff, *Lie groups simply isomorphic with no linear group*, Bull. Amer. Math. Soc. **42** (1936), no. 12, 883–888.

[Bt85] T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*, Graduate Texts in Mathematics, Springer-Verlag, 1985.

[Gud18] S. Gudmundsson, *An Introduction to Riemannian Geometry*, 2018, available at `http://www.matematik.lu.se/matematiklu/personal/sigma/Riemann.pdf`.

[Hal13] B.C. Hall, *Quantum Theory for Mathematicians*, Graduate Texts in Mathematics, Springer New York, 2013.

[Hal15] ———, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Springer International Publishing, 2015.

[HN11] J. Hilgert and K.H. Neeb, *Structure and Geometry of Lie Groups*, Springer Monographs in Mathematics, Springer New York, 2011.

[Kna02] A.W. Knapp, *Lie Groups Beyond an Introduction*, Progress in Mathematics, Birkhäuser Boston, 2002.

[Lee12] J.M. Lee, *Simply Connected Spaces*, 2012, available at `https://sites.math.washington.edu/~duchamp/courses/441/handouts/simplyconn.pdf`.

[Lee13] ———, *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics, Springer New York, 2013.

[Mun00] J.R. Munkres, *Topology*, Featured Titles for Topology Series, Prentice Hall, Incorporated, 2000.

[Sti08] J. Stillwell, *Naive Lie Theory*, Undergraduate Texts in Mathematics, Springer New York, 2008.

[vN29] J. von Neumann, *Über die analytischen Eigenschaften von Gruppen linearer Transformationen und ihrer Darstellungen*, Mathematische Zeitschrift **30** (1929), no. 1, 3–42.

Bachelor's Theses in Mathematical Sciences 2018:K19
ISSN 1654-6229

LUNFMA-4077-2018

Mathematics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lth.se/