Lie Groups, Lie Algebras and Representation Theory

Elías Guisado Villalgordo

Lie Groups, Lie Algebras and Representation Theory

Elías Guisado Villalgordo

Memoria presentada como parte de los requisitos para la obtención de los títulos de Grado en Matemáticas y Grado en Física por la Universidad de Sevilla.

Tutorizada por

Juan Carlos García Vázquez

Contents

Abstract

In Chapter 1, the concepts of a Lie group and a matrix Lie group are introduced, and we construct and study the group homomorphism $\mathsf{SU}(2) \to \mathsf{SO}(3)$. In Chapter 2 we define the notion of a matrix power series, and we do so in a general setting, which allows to deduce properties of which the matrix case is a particular case. The idea of a matrix exponential and matrix logarithm is also defined, and we give proofs of their most important properties. To close the chapter, we give the general statement and the proof of the differentiability of a matrix power series. In Chapter 3, we start with the concepts of an abstract Lie algebra and the Lie algebra of a matrix Lie group, providing the reader with examples of the Lie algebras of the typical matrix Lie groups. Then, we present the idea of a category and a functor, with lots of examples which intend to hint at the value of category theory as a unifying language in mathematics. We do this because in the following section, we prove that there exists a functor from the category of matrix Lie groups into the category of real Lie algebras, which condenses the relation between these two families of mathematical objects. After defining the concept of the complexification of a real Lie algebra, we prove the main general results for matrix Lie groups, and in particular, that matrix Lie groups are embedded submanifolds of $\mathsf{GL}(n;\mathbb{C})$. In the beginning of Chapter 4, we show that representations and actions, of a group or a Lie algebra, are just two sides of the same coin. We later explain how to understand the class of representations of a Lie group or a Lie algebra as a category, and show that the representations of a matrix Lie group can be related with those of its Lie algebra through a functor. After that, we proceed to classify the finite-dimensional irreducible representations of the Lie algebra of $\mathsf{SU}(2)$, and from that, we determine which finite-dimensional irreducible representations of the Lie algebra of $\mathsf{SO}(3)$ come from representations of $\mathsf{SO}(3)$ itself. In the final Chapter 5, we state the Lie group–Lie algebra correspondence, and we use it to show that the category of the finite-dimensional representations of any simply connected matrix Lie group is isomorphic to the category of the finite-dimensional representations of its Lie algebra.

Preface

The idea of this end-of-degree project was suggested to me by my advisor on the summer of 2020. Originally, my advisor proposed to study the theory of Lie groups in order to later study its applications to mathematical physics, and in particular, to quantum mechanics. The idea was to study the Lie groups book [Hall1] of Brian C. Hall so as to later study the applications of the mathematical theory to physics, through the study of the book [Hall2], of same author. Specifically, our intention was to after study Chapter 17 of [Hall2], which address the mathematical physics behind angular momentum in quantum mechanics. As it usually happens with these things, I ended up being able to study only half of what was firstly planned. While I was learning through the material and taking time to solve all the exercises of [Hall1] of the chapters I read, the mathematics part of the original project turned out to be big enough on its own. Or at least, the mathematics themselves turned out to be more than enough material for an EoDP with my study approach. When I study mathematics, I can't help myself trying to achieve the most thorough understanding possible of what I am learning. I usually find myself carefully reading and pondering each piece of text, not being content with the result until I have convinced myself that I know what is really happening underneath on each case. For better or for worse, this system consumes much more time and produces more written text than other study philosophies.

Since the original idea was to apply the learned mathematics to physics, a part of the contents of this thesis are of physical interest. These matters are the homomorphism $\mathsf{SU}(2) \to \mathsf{SO}(3)$, given in Sect. 1.2, and the classification of the finite-dimensional irreducible complex representations of $\mathsf{SU}(2)$ and of $\mathsf{SO}(3)$, given in sects. 4.4 and 4.5, respectively.

The final text of this EoDP you are reading now is based on the very well-written book by Brian C. Hall, [Hall1]. Following the book, the five chapters of this thesis mimic the order and contents of the first five chapters of [Hall1]. Most of the text of the thesis is taken directly from the book of Hall, and either no changes at all or only some minor changes have been added to the text parts of the thesis which are taken from it. Some of these minor changes have been introduced to adapt the text to the scope of the thesis, and some others to add more detail to mathematical arguments.

In comparison with the book of Hall, the novelty that this thesis proposes is the formulation of the results in the language of category theory, whenever such results are suitable for a categorical description. This decision was made because the category theory that we had to introduce to achieve this was minimal: only the notions of a category and a functor. This way, the momentary detour from Lie groups we make in Sect. 3.4 to explain these two concepts does not interrupt the continuity of the rest of the text. Just with the concepts of a category and a functor at hand, we are readily able to categorically describe a lot of things that happen at the interplay between Lie groups and Lie algebras. Nevertheless, it must be noticed that with the categorical formulation we are not really producing new content different from that which was already in [Hall1], and that we include here. Instead, the new thing—at least in this thesis—is the way in which we phrase already known results, and eventually, the way in which we phrase its proofs. Here, the categorical approach allows to understand mathematical phenomena from a more general point of view. Lastly, we shall highlight that the terminology we have chosen in the definitions of the categorical concepts is that of [Rieh].

In the following list, we specify the text parts of each section of the thesis that are taken from [Hall1].

(1.1) The part of examples is taken from Sect. 1.2.
(1.2) The material from the beginning until the statement of Proposition 1.11 is taken from Sect. 1.4.
(2.2) The statement and proof of Proposition 2.12 is taken from Sect. 2.1.
(2.3) is taken from Sect. 2.3, except lemmas 2.16 and 2.17.
(2.4) is taken from Sect. 2.4.
(3.n) is taken from Sect. 3.n, for $n = 1, 2, 5, 6, 7, 8$.
(3.3) is taken from Sect. 3.4, except Lemma 3.17.
(4.1) Text after Lemma 4.1 and before Definition 4.3 is from Sect. 4.1.
(4.2) All text which is not phrased in categorical terms is taken from Sect. 4.1.
(4.3) is taken from Sect. 4.2.
(4.4) is taken from Sect. 4.6.
(4.5) I wrote the part of the three lemmas. The rest of the text is taken from Sect. 4.7.
(5.1) is taken from Sect. 5.1.
(5.3) is taken from Sect. 5.9.
(5.4) is taken from Sect. 5.10.

Whenever any other text has been taken from other book, the source is indicated before the corresponding piece of text.

Acknowledgements

I want to thank all those who, directly or indirectly, helped and eased the production of this thesis. I want to specially thank my advisor, for his kind attention and assistance throughout the development of the text; as well as for his shared passion for mathematics, which always inspires students around him to strive for mathematical greatness. Without him, this text wouldn't be what it has come to be. I would like to thank Marta and Javier as well, whose mathematical conversations have stimulated the progress of this thesis, and whose enthusiasm for the discipline always encourages me to keep learning more. I also feel grateful for my friends who accompanied me throughout this unusual last undergraduate year. Lucía, Juanma, Álvaro, Alberto, Berta, Marcos, Adrián, Irene, Miguel, and more. The pandemic and the production of this text would have been much harder to cope without them.

In general, I want to thank all those who made possible to accomplish graduation after these years of university. I feel particularly indebted to my parents, for their constant support for the life project of their son, allowing me to become the person that I am today. Some people say that university degrees are never achieved alone, and mine is no exception to this rule. I feel obliged to thank all of my university fellows that enabled me to overcome these academic years. I could not say how many times I have helped others and I have been helped from others. The collective abilities of my class fellows have more than exceeded the mere additive sum of our separate strengths.

Seville, July 2021 — Elías Guisado Villalgordo

1 | Matrix Lie Groups

In this thesis, unless stated the contrary, $\mathbb{K}$ will always denote the field of real or complex numbers.

1.1 Definitions and Examples

A Lie group is, roughly speaking, a continuous group, in the sense that the group elements are parameterized by real parameters. The first examples of Lie groups are $(\mathbb{R}^n, +)$ and the circumference $(S^1, \cdot)$, where $S^1 \subset \mathbb{C}$ is the set of complex numbers of modulus 1. The way of mathematically concretise what do we mean by "continuous" in a Lie group is by introducing the notion of a differentiable manifold. So a Lie group is a group that is at the same time either a curve, a surface, or a hypersurface. On the contrary, there are the *discrete* groups, and amongst them we find, for example, the finite groups, but also infinite groups such as $(\mathbb{Z}^n, +)$.

In the same way that an equation like $x^2 + y^2 + z^2 = 1$ defines a surface in $\mathbb{R}^3$, an equation like

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc = 1 \tag{1.1}$$

defines a hypersurface in $\mathbb{R}^4$. At the same time, the set of 2×2 real matrices that satisfy equation (1.1) has a group structure with matrix multiplication, and so, the set of 2×2 real matrices with determinant one turns out to be a Lie group.[1] This is a Lie group made up of matrices, and it constitutes our first example of a *matrix Lie group*,[2] a type of Lie group which find amongst the most studied ones. They are also the ones that we will treat in this thesis.

Groups in general can be understood as the mathematical tool needed to study the concept of symmetry. In this picture, the objective of group theory would be to study the properties and types of symmetries that there can exist, and different types

[1] We will rigorously prove this later on.

[2] In Spanish, *grupo de Lie matricial.*

of groups would give rise to different types of symmetries. In particular, two different general classes of symmetries can be distinguished: *discrete* and *continuous* ones. Examples of the former would be the symmetries of regular polyhedra, or the kind of symmetry of an infinite hexagonal plane honeycomb pattern. Examples of the latter would be the rotational symmetry of a sphere, or the translational invariance which is sometimes postulated in physics in certain physical systems.

The study of Lie groups would then correspond to the study of the *continuous* symmetries of things.

Definition 1.1. A **Lie group** is a set G which is both a group and a manifold[3] and such that these two structures satisfy a compatibility condition: the group operation $G \times G \longrightarrow G$ and the group inverse element map $(\cdot)^{-1} : G \longrightarrow G$ are both differentiable.[4]

We can reduce the two conditions above to a single one: it is not difficult to show that a group G which is also a manifold is a Lie group if and only if the map $(x, y) \in G \times G \mapsto x^{-1}y \in G$ is differentiable.

Definition 1.2. For $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and given $n \in \mathbb{N}$, we define the **general linear group over $\mathbb{K}$ of degree** n, denoted $\mathsf{GL}(n; \mathbb{K})$, as the set of $n \times n$ invertible matrices with entries in $\mathbb{K}$.

We will denote the set of $n \times n$ matrices with entries in $\mathbb{K}$ as $M_n(\mathbb{K})$.

Proposition 1.3. *The general linear group is a Lie group.*

Proof. We can identify $M_n(\mathbb{K})$ with $\mathbb{K}^{n^2}$ in an obvious way, and $\mathbb{K}^{n^2}$ has the standard real manifold structure, which is n^2-dimensional if $\mathbb{K} = \mathbb{R}$ and $2n^2$-dimensional if $\mathbb{K} = \mathbb{C}$. Thus, since the determinant is a continuous function, $\mathsf{GL}(n; \mathbb{K}) = \det^{-1}(\mathbb{K} \setminus \{0\}) \subset M_n(\mathbb{K})$ is an open set of $M_n(\mathbb{K})$, and therefore it inherits a manifold structure from $M_n(\mathbb{K})$, of the same dimension. On the other hand, matrix multiplication is differentiable and also is the map $A \in \mathsf{GL}(n; \mathbb{K}) \mapsto A^{-1} \in \mathsf{GL}(n; \mathbb{K})$, since A^{-1} equals $\dfrac{1}{\det A}$ times the adjugate matrix of A, and both are differentiable functions of $A \in \mathsf{GL}(n; \mathbb{K})$. □

Definition 1.4. A **matrix Lie group** is a closed subgroup of $\mathsf{GL}(n; \mathbb{C})$.

$\mathsf{GL}(n; \mathbb{R})$ is a matrix Lie group, for if $A_m \subset \mathsf{GL}(n; \mathbb{R})$ and A_m converges to some invertible matrix A, its entries must be real.

There are two reasons behind Definition 1.4. The principal one is that all interesting groups which are made up of invertible matrices turn out to be closed in the

[3]We will be always considering C^∞-differentiable real manifolds. That is, a topological manifold (a locally euclidean, Hausdorff and second-countable topological space) with a C^∞-differentiable atlas.

[4]By **differentiable**, from now on, we will always mean C^∞-differentiability. We will occasionally speak of a **smooth** function to refer to the same concept.

complex general linear group (this does not mean that they are closed in $M_n(\mathbb{K})$), amongst of these "interesting matrix groups" there are the matrix groups we will be considering in this thesis. The other reason comes from the closed-subgroup theorem, which asserts that every closed subgroup of a Lie group G is also an embedded submanifold of G and thus also a Lie group by its own. We will be proving the closed-subgroup theorem for the case of $\mathsf{GL}(n;\mathbb{C})$. That is, we will prove that every matrix Lie group is indeed a Lie group.

An example of a subgroup of $\mathsf{GL}(n;\mathbb{C})$ which is not closed is the subset of invertible matrices with rational coefficients. Another interesting example is the "irrational line in a torus." See Fig. 1.1 of [Hall1] and Exercise 10 of Chapter 1 of same book.

Mastering the subject of Lie groups involves not only learning the general theory but also familiarizing oneself with examples. We now introduce some of the most important examples of (matrix) Lie groups.

Definition 1.5. The **special linear group**, denoted $\mathsf{SL}(n;\mathbb{K})$, is the group of $n\times n$ invertible matrices with entries in $\mathbb{K}$ which have determinant one. It is a subgroup of $\mathsf{GL}(n;\mathbb{C})$.

If A_n is a sequence of matrices with determinant one and A_n converges to A, then A also has determinant one, because the determinant is a continuous function. Thus, $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$ are matrix Lie groups.

Recall that an $n \times n$ complex matrix A is said to be **unitary** if the column vectors of A are orthonormal, that is, if

$$\sum_{l=1}^{n} \overline{A_{lj}} A_{lk} = \delta_{jk}. \tag{1.2}$$

We may rewrite (1.2) as

$$\sum_{l=1}^{n} (A^*)_{jl} A_{lk} = \delta_{jk}, \tag{1.3}$$

where δ_{jk} is the Kronecker delta, equal to 1 if $j = k$ and equal to zero if $j \neq k$. Here A^* is the **adjoint** of A, defined by

$$(A^*)_{jk} = \overline{A_{kj}}.$$

Equation (1.3) says that $A^*A = I$; thus, we see that A is unitary if and only if $A^* = A^{-1}$. In particular, every unitary matrix is invertible.

The adjoint operation on matrices satisfies $(AB)^* = B^*A^*$. From this, we can see that if A and B are unitary, then

$$(AB)^*(AB) = B^*A^*AB = B^{-1}A^{-1}AB = I,$$

showing that AB is also unitary. Furthermore, since $(AA^{-1})^* = I^* = I$, we see that $(A^{-1})^*A^* = I$, which shows that $(A^{-1})^* = (A^*)^{-1}$. Thus, if A is unitary, we have

$$(A^{-1})^*A^{-1} = (A^*)^{-1}A^{-1} = (AA^*)^{-1} = I,$$

showing that A^{-1} is again unitary.

Thus, the collection of unitary matrices is a subgroup of $\mathsf{GL}(n;\mathbb{C})$.

Definition 1.6. The **unitary group of degree** n, denoted $\mathsf{U}(n)$, is the set of all unitary $n \times n$ matrices. The **special unitary group of degree** n, denoted $\mathsf{SU}(n)$, is defined to be $\mathsf{U}(n) \cap \mathsf{SL}(n;\mathbb{C})$, the set of unitary $n \times n$ matrices with determinant one.

It is easy to check that both $\mathsf{U}(n)$ and $\mathsf{SU}(n)$ are closed subgroups of $\mathsf{GL}(n;\mathbb{C})$ and thus matrix Lie groups.

Meanwhile, let $\langle \cdot, \cdot \rangle$ denote the standard inner product on $\mathbb{C}^n$, given by

$$\langle x, y \rangle = \sum_{j=1}^{n} \overline{x}_j y_j.$$

(Note that we put the conjugate on the *first* factor of the inner product.) For any $x, y \in \mathbb{C}^n$, we have

$$\begin{aligned}\langle x, Ay \rangle &= \sum_{j=1}^{n} \overline{x}_j (Ay)_j = \sum_{j=1}^{n} \overline{x}_j \sum_{k=1}^{n} A_{jk} y_k \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} (\overline{A^*})_{kj} \overline{x}_j y_k = \sum_{k=1}^{n} (\overline{A^* x})_k y_k = \langle A^* x, y \rangle .\end{aligned}$$

Thus,

$$\langle Ax, Ay \rangle = \langle A^*Ax, y \rangle ,$$

from which we can see that if A is unitary, then A preserves the inner product on $\mathbb{C}^n$, that is,

$$\langle Ax, Ay \rangle = \langle x, y \rangle$$

for all x and y. Conversely, if A preserves the inner product, we must have $\langle A^*Ax, y \rangle = \langle x, y \rangle$ for all x, y. It is not hard to see that this condition holds only if $A^*A = I$. Thus, an equivalent characterization of unitarity is that A is unitary if and only if A preserves the standard inner product on $\mathbb{C}^n$.

Finally, for any matrix A, we have that $\det A^* = \overline{\det A}$. Thus, if A is unitary, we have

$$\det(A^*A) = |\det A|^2 = \det I = 1.$$

Hence, for all unitary matrices A, we have $|\det A| = 1$.

In a similar fashion, an $n \times n$ real matrix A is said to be **orthogonal** if the column vectors of A are orthonormal. As in the unitary case, we may give equivalent versions of this condition. The only difference is that if A is real, A^* is the same as the transpose A^{tr} of A, given by

$$(A^{\text{tr}})_{jk} = A_{kj}.$$

Thus, A is orthogonal if and only if $A^{\text{tr}} = A^{-1}$, and this holds if and only if A preserves the inner product in $\mathbb{R}^n$. Since $\det(A^{\text{tr}}) = \det A$, if A is orthogonal, we have

$$\det(A^{\text{tr}}A) = \det(A)^2 = \det I = 1,$$

so that $\det A = \pm 1$. From the same argument as before, the product of orthogonal matrices is orthogonal and the inverse of and orthogonal matrix is also orthogonal.

Definition 1.7. The **orthogonal group of degree** n, denoted $\mathsf{O}(n)$, is the set of all real orthogonal $n \times n$ matrices. The **special unitary group of degree** n, denoted $\mathsf{SO}(n)$, is defined to be $\mathsf{O}(n) \cap \mathsf{SL}(n; \mathbb{R})$, the set of real orthogonal $n \times n$ matrices with determinant one.

It is easy to check that both $\mathsf{O}(n)$ and $\mathsf{SO}(n)$ are closed subgroups of $\mathsf{GL}(n; \mathbb{C})$ and thus matrix Lie groups.

Geometrically, elements of $\mathsf{SO}(n)$ are rotations, while the elements of $\mathsf{O}(n)$ are either rotations or combinations of rotations and reflections.

As a final example, we observe that several important groups which are not defined as groups of matrices can be thought as such. The group $\mathbb{R} \setminus \{0\}$ of non-zero real numbers under multiplication is isomorphic to $\mathsf{GL}(1; \mathbb{R})$. Similarly, the group $\mathbb{C} \setminus \{0\}$ of non-zero complex numbers under multiplication is isomorphic to $\mathsf{GL}(1; \mathbb{C})$ and the group S^1 of complex numbers with absolute value one is isomorphic to $\mathsf{U}(1)$.

The group $\mathbb{R}$ under addition is isomorphic to $\mathsf{GL}(1; \mathbb{R})^+$ (1×1 real matrices with positive determinant) via the map $x \mapsto [e^x]$. The group $\mathbb{R}^n$ (with vector addition) is isomorphic to the group of diagonal real matrices with positive diagonal entries, via the map

$$(x_1, \ldots, x_n) \longmapsto \begin{pmatrix} e^{x_1} & & 0 \\ & \ddots & \\ 0 & & e^{x_n} \end{pmatrix}.$$

1.2 Homomorphisms

Definition 1.8. Let G and H be matrix Lie groups. A map $\Phi : G \to H$ is called a **Lie group homomorphism** if (1) Φ is a group homomorphism and (2) Φ is continuous.

If, in addition, Φ is bijective and the inverse map Φ^{-1} is continuous, then Φ is called a **Lie group isomorphism**.

Given G and H arbitrary Lie groups, it is customary to call a map Φ between two Lie groups a Lie group homomorphism if Φ is a group homomorphism and Φ is *smooth*, whereas in the previous definition we have only required that Φ be continuous. In Sect. 3.7 we will show that every continuous homomorphism between matrix Lie groups is automatically smooth, so that there is no conflict in terminology.

Recall from elementary group theory that the inverse of a bijective group homomorphism is also a group homomorphism. Thus, if Φ is a Lie group isomorphism, then so is Φ^{-1}. Any two matrix Lie groups G and H between which there exists a Lie group isomorphism are said to be **isomorphic** (as Lie groups), and in that case we write $G \cong H$.

The simplest interesting example of a Lie group homomorphism is the determinant, which is a homomorphism of $\mathsf{GL}(n;\mathbb{K})$ into $\mathsf{GL}(1;\mathbb{K}) = \mathbb{K}\setminus\{0\}$. Another simple example is the map $\Phi : \mathbb{R} \longrightarrow \mathsf{SO}(2)$ given by

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

This map is clearly continuous, and a calculation (using standard trigonometric identities) shows that it is a homomorphism.

An important topic in the theory of matrix Lie groups and specially in the physics applications is the relationship between $\mathsf{SU}(2)$ and $\mathsf{SO}(3)$, which are almost, but not quite, isomorphic. Specifically, we now construct a Lie group homomorphism $\Phi : \mathsf{SU}(2) \longrightarrow \mathsf{SO}(3)$ that is two-to-one and onto.[5] To construct the homomorphism and further study its properties, the next two results will be practical.

Proposition 1.9. $\mathsf{SU}(2)$ *is homeomorphic to* S^3.

Proof. We show that every matrix $A \in \mathsf{SU}(2)$ is of the form

$$A = \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix}, \tag{1.4}$$

where $\alpha, \beta \in \mathbb{C}$ are such that $|\alpha|^2 + |\beta|^2 = 1$.

On the one hand, observe that the columns of (1.4) are an orthonormal basis of $\mathbb{C}^2$, so the matrix of (1.4) is in $\mathsf{SU}(2)$.

[5] Any function $f : X \longrightarrow Y$, where X and Y are sets, induces an equivalence relationship in its domain X by the rule: for $x, x' \in X$, we have $x \sim_f x'$ if and only if $f(x) = f(x')$. By definition, for any integer $n \geq 1$, a function $f : X \longrightarrow Y$ is then said to be n-to-one if all the equivalence classes from $X/{\sim_f}$ have cardinal equal to n. That is, f is n-to-one if $\#f^{-1}(y) = n$ for all $y \in \operatorname{Im} f$; exactly n elements of X are mapped to each $y \in \operatorname{Im} f$. A one-to-one function is the same as an injective function.

Conversely, suppose

$$A = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \in \mathsf{SU}(2).$$

If we call $W = \operatorname{span}\{(\alpha, \beta)\}$, then $\mathbb{C}^2 = W \oplus W^\perp$, so $W^\perp$ is one-dimensional and therefore $W^\perp = \operatorname{span}\{(-\overline{\beta}, \overline{\alpha})\}$. Since the columns of A are an orthonormal basis of $\mathbb{C}^2$, we have $(\gamma, \delta) \in W^\perp$ and thus $(\gamma, \delta) = c(-\overline{\beta}, \overline{\alpha})$, for some complex number c, so

$$A = \begin{pmatrix} \alpha & -c\overline{\beta} \\ \beta & c\overline{\alpha} \end{pmatrix}.$$

Taking determinant, $1 = \det A = c(|\alpha|^2 + |\beta|^2) = c\,\|(\alpha, \beta)\|^2 = c$.

We can define a map

$$\begin{aligned} \mathsf{SU}(2) &\longrightarrow S^3 \\ \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} &\longmapsto (a_1, a_2, b_1, b_2), \end{aligned}$$

where $\alpha = a_1 + ia_2$ and $\beta = b_1 + ib_2$, with $a_j, b_j \in \mathbb{R}$. This map is surjective and from the fact that all matrices of $\mathsf{SU}(2)$ are of the form (1.4), it also follows that it is injective. Moreover, it is continuous and so is the inverse $(a_1, a_2, b_1, b_2) \mapsto \begin{pmatrix} a_1 + ia_2 & -b_1 + ib_2 \\ b_1 + ib_2 & a_1 - ia_2 \end{pmatrix}$. Therefore $\mathsf{SU}(2) \cong S^3$. □

It deduces that $\mathsf{SU}(2)$ is path-connected, since S^3 is path-connected (in general $\mathsf{SU}(n)$ is always path-connected, see Proposition 1.13 from [Hall1]). Furthermore, $\mathsf{SU}(2)$ is simply connected, since S^3 is (see for example Theorem 59.3 in [Mun]). In chapter 5, we will see that the topological property of simple-connectedness is a remarkable one in the theory of Lie groups.

Lemma 1.10. *The trace of a product of matrices is invariant under cyclic permutations. That is, if $A_1, \ldots, A_n \in M_n(\mathbb{C})$, then*

$$\begin{aligned} \operatorname{trace}(A_1 A_2 \cdots A_n) &= \operatorname{trace}(A_2 A_3 \cdots A_n A_1) \\ &= \operatorname{trace}(A_3 A_4 \cdots A_n A_1 A_2) \\ &\ \vdots \\ &= \operatorname{trace}(A_n A_1 A_2 \cdots A_{n-1}). \end{aligned}$$

Proof. It is sufficient to prove that $\operatorname{trace}(AB) = \operatorname{trace}(BA)$, and indeed

$$\operatorname{trace}(AB) = \sum_{j=1}^{n} (AB)_{jj} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} b_{kj} = \sum_{k=1}^{n} \sum_{j=1}^{n} b_{kj} a_{jk} = \sum_{k=1}^{n} (BA)_{kk} = \operatorname{trace}(BA).$$

□

We now build the homomorphism $\Phi : \mathsf{SU}(2) \longrightarrow \mathsf{SO}(3)$. Consider the space V of all 2×2 complex matrices X which are Hermitian (i.e., $X^* = X$) and have trace zero. Elements of V are precisely the matrices of the form

$$X = \begin{pmatrix} x_1 & x_2 + ix_3 \\ x_2 - ix_3 & -x_1 \end{pmatrix}, \tag{1.5}$$

with $x_1, x_2, x_3 \in \mathbb{R}$. If we identify V with $\mathbb{R}^3$ by means of the coordinates x_1, x_2 and x_3 in (1.5), then the standard inner product on $\mathbb{R}^3$ can be computed as

$$\langle X_1, X_2 \rangle = \frac{1}{2}\operatorname{trace}(X_1 X_2).$$

That is to say,

$$\begin{aligned} &\frac{1}{2}\operatorname{trace}\left(\begin{pmatrix} x_1 & x_2 + ix_3 \\ x_2 - ix_3 & -x_1 \end{pmatrix}\begin{pmatrix} x_1' & x_2' + ix_3' \\ x_2' - ix_3' & -x_1' \end{pmatrix}\right) \\ &\quad = x_1x_1' + x_2x_2' + x_3x_3' \end{aligned}$$

as one may easily check by direct calculation.

For each $U \in \mathsf{SU}(2)$, define a linear map $\Phi_U : V \longrightarrow V$ by

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

Then, by Lemma 1.10, UXU^{-1} still has trace zero and since U is unitary,

$$(UAU^{-1})^* = (U^{-1})^* AU^* = UAU^{-1},$$

showing that UAU^{-1} is again in V.

It is easy to see that $\Phi_{U_1U_2} = \Phi_{U_1}\Phi_{U_2}$. Furthermore,

$$\begin{aligned} \frac{1}{2}\operatorname{trace}\left(\left(UX_1U^{-1}\right)\left(UX_2U^{-1}\right)\right) &= \frac{1}{2}\operatorname{trace}\left(UX_1X_2U^{-1}\right) \\ &= \frac{1}{2}\operatorname{trace}\left(X_1X_2\right), \end{aligned}$$

by Lemma 1.10. Thus, each Φ_U preserves the inner product $\frac{1}{2}\operatorname{trace}(X_1X_2)$ on V. It follows that the map $U \mapsto \Phi_U$ is a homomorphism of $\mathsf{SU}(2)$ into the group of orthogonal linear transformations of $V \cong \mathbb{R}^3$, that is, into $\mathsf{O}(3)$.

We next show the continuity of $U \mapsto \Phi_U$. Let $f : V \longrightarrow \mathbb{R}^3$ be the linear isomorphism $V \cong \mathbb{R}^3$, explicitly written as

$$f(X) = \begin{pmatrix} X_{11} \\ \operatorname{Re} X_{12} \\ \operatorname{Im} X_{12} \end{pmatrix},$$

and let

$$X_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \tag{1.6}$$

be the canonical basis of V. This way we have $f(X_j) = e_j$, where e_j are the canonical basis vectors of $\mathbb{R}^3$. For each $U \in \mathsf{SU}(2)$, the transformation matrix of the linear map Φ_U with respect to the basis (1.6) is $\big(f(UX_1U^{-1}) \mid f(UX_2U^{-1}) \mid f(UX_3U^{-1})\big)$. Since the inverse matrix is a continuous operation, the entries of this matrix depend continuously on $U \in \mathsf{SU}(2)$. That is, the map $\Phi : U \in \mathsf{SU}(2) \mapsto \Phi_U \in \mathsf{O}(3)$ is continuous.

Since $\mathsf{SU}(2)$ is connected, Φ_U must actually lie in $\mathsf{SO}(3)$ for all $U \in \mathsf{SU}(2)$.[6] Thus, Φ (i.e., the map $U \mapsto \Phi_U$) is a Lie group homomorphism of $\mathsf{SU}(2)$ into $\mathsf{SO}(3)$.

Since $(-I)X(-I)^{-1} = X$, we see that Φ_{-I} is the identity element of $\mathsf{SO}(3)$.

Proposition 1.11. *The map $U \mapsto \Phi_U$ is a two-to-one and onto map of $\mathsf{SU}(2)$ to $\mathsf{SO}(3)$, with kernel equal to $\{I, -I\}$.*

We do not give here the proof of the proposition, for it is elementary and is found in [Hall1], Proposition 1.19, and we will not use it further in this thesis in a critical way.

Proposition 1.11 has an interesting application. Using quotient map theory we can use it to prove that $\mathsf{SO}(3)$ is homeomorphic to $\mathbb{RP}^3$, the real 3-dimensional projective space. Before giving the result, we give a brief overview of quotient map theory. For more details regarding quotient maps, one may consult §22 from [Mun].

Recall that a **quotient map** between topological spaces $f : A \longrightarrow B$ is a continuous and surjective map such that the codomain B is equipped with the final topology with respect to f. That is, the topology on B is the finest one among those topologies of B for which f becomes continuous. Equivalently, f is is a quotient map if it is (i) surjective and (ii) **strongly continuous**; where (ii) means that $G \subset B$ is open if and only if $f^{-1}(G) \subset A$ is open; this last thing is the same as saying that $F \subset B$ is closed if and only if $f^{-1}(F) \subset A$ is closed. Thus, a sufficient (although not necessary) condition for a continuous map to be strongly continuous is to be an open or closed map.

If $f : A \longrightarrow B$ is *any* continuous function between topological spaces, by the universal property of the quotient topology there is an induced continuous map $\hat{f} : A/{\sim_f} \longrightarrow B$ (see footnote 5 on p. 10 for the definition of $\sim_f$), which on this case is also injective. Namely, $\hat{f}$ is the unique map which fits on the following diagram while

[6] In more detail: if we compose $\det \circ \Phi$, we obtain a continuous function from the connected topological space $\mathsf{SU}(2)$ to the set of the possible determinant values of matrices of $\mathsf{SO}(3)$, that is, into $\{1, -1\}$. Since every continuous function from a connected topological space into a discrete topological space is constant and $(\det \circ \Phi)(I) = 1$, we have that $\det \circ \Phi$ is a map constantly equal to 1. That is, $\operatorname{Im} \Phi \subset \mathsf{SO}(3)$.

making it commutative

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d] & B \\ \frac{A}{\sim_f} \arrow[ur, "\hat{f}"'] & \end{tikzcd}.$$

It turns out that the function f is a quotient map if and only if $\hat{f}$ is a homeomorphism (Corollary 22.3 [Mun]). This is an important characterization of quotient maps. We could define the concept of a **quotient space** of a topological space X to be a pair (Q, p), where Q is a topological space and $p : X \to Q$ is a quotient map. If $\sim$ is any equivalence relationship on X, then $p : X \to X/\sim$ is always a quotient map and therefore $(X/\sim, p)$ is a quotient space of X. With this terminology, the previous statement "f is a quotient map if and only if $\hat{f}$ is a homeomorphism" can be interpreted as saying that quotient spaces of a topological space X are *essentially the same thing* as spaces $X/\sim$ with the quotient topology for some equivalence relation $\sim$ on X.

Lastly, recall the *closed map lemma*, which states that every continuous map f from a compact topological space A to a Hausdorff space B is also closed: let $F \subset A$ be closed. Since closed sets of a compact space inherit the compactness of the space, F is also compact. Since continuity preserves compactness, $f(F)$ is also compact. Finally, since every compact subspace of a Hausdorff space is also closed, $f(F)$ is closed.

Corollary 1.12. $\mathsf{SO}(3)$ *is homeomorphic to* $\mathbb{RP}^3$*, the real projective space of dimension 3.*

Proof. The map $\Phi : U \in \mathsf{SU}(2) \mapsto \Phi_U \in \mathsf{SO}(3)$ is a quotient map: by Proposition 1.11 it is surjective and it is strongly continuous, for it is continuous and also a closed map by the closed map lemma, as $\mathsf{SU}(2) \cong S^3$ is compact and $\mathsf{SO}(3) \subset M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ is Hausdorff, since topological subspaces of a Hausdorff space are also Hausdorff.

We recall from elementary group theory that the cosets of $\ker\Phi$ in $\mathsf{SU}(2)$ are precisely the equivalence classes of $\sim_\Phi$. We thus have $\mathsf{SU}(2)/\ker\Phi = \mathsf{SU}(2)/\sim_\Phi \cong \mathsf{SO}(3)$, where the "$\cong$" is both a group isomorphism, by the first isomorphism theorem for groups; and a homeomorphism, by quotient map theory, Corollary 22.3 of [Mun]. Since $\ker\Phi = \{I, -I\}$ by Proposition 1.11, the cosets of $\ker\Phi$ in $\mathsf{SU}(2)$ are of the form $\{A, -A\}$, with $A \in \mathsf{SU}(2)$. By the homeomorphism $\mathsf{SU}(2) \cong S^3$ given in proof of Proposition 1.9, quotienting $\mathsf{SU}(2)$ by the equivalence relationship $\sim_\Phi$ (whose equivalence classes are $\{A, -A\}$) amounts to identifying antipodes in S^3. That is, there is a homeomorphism $\mathbb{RP}^3 \cong \mathsf{SU}(2)/\sim_\Phi \cong \mathsf{SO}(3)$. □

The homeomorphism $\mathsf{SO}(3) \cong \mathbb{RP}^3$ constructed in the previous proof can be shown to be really a diffeomorphism, although we will not be using this fact further on and we have not shown yet that $\mathsf{SO}(3)$ is a differentiable manifold.

An alternative proof of Corollary 1.12 is found in Proposition 1.17 of [Hall1].

It follows that $\mathsf{SO}(3)$ is not simply connected, as the fundamental group of $\mathbb{RP}^3$ is the group of two elements. In order to prove that the fundamental group of $\mathbb{RP}^3$ is indeed of order two, one needs to show that $S^3 \to \mathbb{RP}^3$ is a **covering map** (see §53 from [Mun] for the definition of a covering map) and then use Theorem 54.4 from [Mun].

2 | The Matrix Exponential

2.1 Normed and Banach algebras

The primary objective of this chapter is to define and to study the properties of the matrix exponential and the matrix logarithm. These two are particular instances of *matrix power series* in $M_n(\mathbb{K})$. So to study them, we move to a more general setting, which abstracts the structure of the space $M_n(\mathbb{K})$ which is at play when investigating the properties of a matrix power series; namely, that of a *Banach algebra.*

Definition 2.1. An **algebra over** $\mathbb{K}$ or a $\mathbb{K}$**-algebra** is a pair $(\mathcal{A}, [\cdot,\cdot])$ where $\mathcal{A}$ is a $\mathbb{K}$-vector space and $[\cdot,\cdot] : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ is a $\mathbb{K}$-bilinear map, which is referred to as the **product** of the algebra. If the product is implicitly understood, we will refer to the algebra $(\mathcal{A}, [\cdot,\cdot])$ simply as $\mathcal{A}$. Whenever the product is either associative or commutative, we will refer to $\mathcal{A}$ as an associative or commutative algebra, respectively. Lastly, if $\mathcal{A}$ has a multiplicative identity, that is, an element $1 \in \mathcal{A}$ such that $[1, X] = [X, 1] = X$ for every $X \in \mathcal{A}$, we will say that $\mathcal{A}$ is a unital algebra.

One could also speak of an algebra as a vector space which is at the same a ring (not necessarily commutative, associative or unital) and such the ring and vector space structures satisfy a compatibility condition: scalar and ring multiplication are "associative," $\lambda(XY) = (\lambda X)Y = X(\lambda Y)$, for $\lambda \in \mathbb{K}$ and $X, Y \in \mathcal{A}$. Here we have represented the bracket $[X, Y]$ simply as XY. This notation is primarily preferred for associative algebras since in this case it makes the associativity explicit (it is nicer to write XYZ rather than $(XY)Z$ or $X(YZ)$) and we will be using it for these algebras. In any case, it can be also used in non-associative algebras if one is careful with the corresponding parentheses.

If $\mathcal{A}$ is a non-zero unital $\mathbb{K}$-algebra, in this case we can identify $\mathbb{K}$ with a subring of $\mathcal{A}$, namely, $\{\lambda 1_{\mathcal{A}} \mid \lambda \in \mathbb{K}\}$, for $\lambda \in \mathbb{K} \mapsto \lambda 1_{\mathcal{A}} \in \mathcal{A}$ is a ring homomorphism from a field to a non-zero ring, and thus it is a ring embedding. So an equivalent definition for a non-zero unital $\mathbb{K}$-algebra is a ring extension $\mathbb{K} \subset \mathcal{A}$, where $\mathcal{A}$ is a non-commutative, non-associative ring.

Definition 2.2. A **normed algebra over** $\mathbb{K}$ or a **normed** $\mathbb{K}$**-algebra** is an associative $\mathbb{K}$-algebra $\mathcal{A}$ which is also a normed space and such that these two structures satisfy a compatibility condition: $\|XY\| \le \|X\| \|Y\|$ for every $X, Y \in \mathcal{A}$. That is, the norm is submultiplicative. If $\mathcal{A}$ is also a Banach space, we will say that $\mathcal{A}$ is a **Banach algebra**.

From now on and unless stated otherwise, normed algebras will always considered to be non-zero.

Example 2.3. $M_n(\mathbb{K})$ with any matrix norm (i.e., a submultiplicative norm with respect to the matrix product) is a Banach algebra. More generally, given a Banach space $\mathcal{B}$, the family of bounded operators in $\mathcal{B}$, $L(\mathcal{B})$, is known to be a Banach space with the operator norm. This norm is also submultiplicative with respect to operator composition, and so $L(\mathcal{B})$ is also a Banach algebra.

It follows from the norm submultiplicativeness and the next result that the product in a normed algebra is continuous.

Proposition 2.4. *For a bilinear map $B : X \times Y \to Z$, where X, Y, Z are normed spaces over $\mathbb{K}$, the following are equivalent:*

1. *B is continuous.*
2. *B is continuous at $(0, 0)$.*
3. *B is bounded as a bilinear map. That is, there exists $C > 0$ such that $\|B(x, y)\| \le C \|x\| \|y\|$ for every $(x, y) \in X \times Y$.*

Proof. $(1 \Rightarrow 2)$. It's clear.

$(2 \Rightarrow 3)$. Suppose that 3 is false. For each $n \in \mathbb{N}$ there exists $(x_n, y_n) \in X \times Y$ such that $\|B(x_n, y_n)\| > n^2 \|x_n\| \|y_n\|$. Since clearly $x_n \neq 0$ and $y_n \neq 0$, we can consider

$$\tilde{x}_n := \frac{x_n}{n \|x_n\|} \to 0 \quad \text{and} \quad \tilde{y}_n := \frac{y_n}{n \|y_n\|} \to 0.$$

But bilinearity of B now implies that

$$\|B(\tilde{x}_n, \tilde{y}_n)\| > n^2 \cdot \frac{1}{n} \cdot \frac{1}{n} = 1 \quad \text{for each } n$$

and thus 2 is false.

$(3 \Rightarrow 1)$. Suppose 3 holds and let $x_n \to x$ in X and $y_n \to y$ in Y. There exists $M \ge 0$ such that $\|x_n\| \le M$ and $\|y\| \le M$. Then

$$\begin{aligned} \|B(x_n, y_n) - B(x, y)\| &\le \|B(x_n, y_n) - B(x_n, y)\| + \|B(x_n, y) - B(x, y)\| \\ &= \|B(x_n, y_n - y)\| + \|B(x_n - x, y)\| \\ &\le C \|x_n\| \|y_n - y\| + C \|x_n - x\| \|y\| \\ &\le CM (\|x_n - x\| + \|y_n - y\|) \to 0, \end{aligned}$$

and we are done. □

Example 2.5. On this thesis, when dealing with complex matrices we will be working with an arbitrary matrix norm. We can do this because the norms in a finite-dimensional vector space are all equivalent. However, it is nice to have an example of such a norm. In $M_n(\mathbb{K})$, the Frobenius inner product, for each $A, B \in M_n(\mathbb{K})$ is defined as

$$\langle A, B\rangle = \operatorname{trace}(A^*B).$$

The Frobenius inner product is indeed an inner product. In fact, it coincides with the standard inner product in $\mathbb{K}^{n^2} \cong M_n(\mathbb{K})$,

$$\langle A, B\rangle = \operatorname{trace}(A^*B) = \sum_{j,k=0}^{n} \overline{a}_{jk} b_{jk}.$$

The induced norm, $\|A\| = \sqrt{\operatorname{trace}(A^*A)} = \sum_{j,k=0}^{n} |a_{jk}|^2$, is called the Frobenius norm or the Hilbert-Schmidt norm. It is a matrix norm: If $A, B \in M_n(\mathbb{K})$, and if we denote by $A_{(j)}$ and $B^{(k)}$ the j^{th} row and k^{th} column of A and B, respectively, then

$$\begin{aligned}
\|AB\|^2 &= \sum_{j=1}^{n}\sum_{k=1}^{n} \overline{A}_{(j)} B^{(k)} \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} |\langle A_{(j)}^{\text{tr}}, B^{(k)}\rangle|^2, && \text{where } \langle\cdot,\cdot\rangle \text{ is the standard inner product in } \mathbb{K}^n, \\
&\le \sum_{j=1}^{n}\sum_{k=1}^{n} \|A_{(j)}^{\text{tr}}\|_v^2 \|B^{(k)}\|_v^2, && \text{where } \|\cdot\|_v \text{ is the euclidean norm in } \mathbb{K}^n \\
& && \text{(here we have used the Cauchy-Schwarz} \\
& && \text{inequality),} \\
&= \sum_{j=1}^{n} \|A_{(j)}^{\text{tr}}\|_v^2 \sum_{k=1}^{n} \|B^{(k)}\|_v^2 \\
&= \|A\|^2 \, \|B\|^2 .
\end{aligned}$$

Definition 2.6. Let $\mathcal{A}$ be a unital normed $\mathbb{K}$-algebra. A **power series in $\mathcal{A}$** is a series of the form

$$\sum_{n=0}^{\infty} \lambda_n (X - c)^n,$$

where $\lambda_n, c \in \mathbb{K}$ and $X \in \mathcal{A}$. For the case $\mathcal{A} = M_n(\mathbb{K})$, we will speak of a **matrix power series**.

Recall that in a Banach space $\mathcal{B}$ every absolutely convergent series is convergent as well. Indeed, if $\{x_n\}_{n\ge1} \subset \mathcal{B}$ and the series $\sum_{n=1}^{\infty} x_n$ converges absolutely, it means that for every $\varepsilon > 0$ there exists n_0 such that $\varepsilon > \sum_{n=m}^{\infty} \|x_n\| \ge \sum_{n=m}^{m'} \|x_n\| = \|S_{m'} - S_m\|$ whenever $m' \ge m \ge n_0$, where $S_k = \sum_{n=1}^{k} x_n$. This way, the sequence of partial sums $\{S_k\}_{k\ge1}$ is a Cauchy sequence and thus, since $\mathcal{B}$ is Banach, $\lim_{k\to\infty} S_k = \sum_{n=0}^{\infty} x_n$ is convergent.

Proposition 2.7. *Every convergent power series in $\mathbb{K}$ induces a well-defined and continuous power series in a unital Banach algebra. Specifically, let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{K}$ and let*

$$f(x) = \sum_{n=0}^{\infty} \lambda_n (x-c)^n, \quad x \in \mathbb{K} \tag{2.1}$$

be a power series, with $\lambda_n, c \in \mathbb{K}$ and suppose that (2.1) has radius of convergence $r > 0$. Then

$$f(X) = \sum_{n=0}^{\infty} \lambda_n (X-c)^n, \quad X \in \mathcal{A} \tag{2.2}$$

is well-defined and is continuous for $X \in \mathcal{A}$, $\|X - c\| < r$.

Proof. Every power series over $\mathbb{K}$ is also absolutely convergent in its open disk of convergence, so we have $\sum_{n=0}^{\infty} |\lambda_n||x-c|^n < +\infty$ for every $x \in \mathbb{K}$, $|x - c| < r$. In other words, $\sum_{n=0}^{\infty} |\lambda_n| y^n < +\infty$ for $y \in [0, r)$. Taking $y = \|X - c\|$ makes the series (2.2) absolutely convergent for $\|X - c\| < r$ and thus also convergent for these values of X, since $\mathcal{B}$ is a Banach space.

Let's see about the continuity. Without lost of generality, we can suppose $c = 0$, for we can perform the change of variable $Y = X - c$. We have thus a series $g(Y) = \sum_{n=0}^{\infty} \lambda_n Y^n$ which converges for $Y \in \mathcal{A}$ if and only if $f(X)$ converges for $X = Y + c$. That is, $g(Y)$ is well-defined for every $Y \in \mathcal{A}$ with $\|Y\| < r$. And therefore we can restrict to proving continuity for g.

In order to prove continuity for a power series in $\mathcal{A}$, we now study how can we bound the expression $(X+Y)^k - X^k$ in terms of $\|X\|$ and $\|Y\|$.

Given a word $w = (X_1, \dots, X_n) \in \mathcal{A}^n$, we define its product $\operatorname{prod} w = X_1 \cdots X_n \in \mathcal{A}$ and given some $Y \in \mathcal{A}$, we define the *multiplicity of Y in w* as

$$\mathrm{m}_Y\, w = \#\{j \in \{1, \dots, n\} \mid X_j = Y\},$$

that is, $\mathrm{m}_Y\, w$ is the number of Y's that there is in w.

Given $X, Y \in \mathcal{A}$, let

$$p(X, Y, n, m) = \sum_{\substack{w \in \{X,Y\}^{n+m} \\ \mathrm{m}_X w = n \\ \mathrm{m}_Y w = m}} \operatorname{prod} w$$

be the sum of the different permutations of $X^n Y^m$, so that

$$(X+Y)^k = \sum_{m=0}^{k} p(X, Y, k-m, m).$$

By the submultiplicativeness of $\|\cdot\|$,

$$\begin{aligned}
\|p(X,Y,k-m,m)\| &= \left\| \sum_{\substack{w\in\{X,Y\}^k \\ \mathrm{m}_X\, w=k-m \\ \mathrm{m}_Y\, w=m}} \operatorname{prod} w \right\| \\
&\le \sum_{\substack{w\in\{X,Y\}^k \\ \mathrm{m}_X\, w=k-m \\ \mathrm{m}_Y\, w=m}} \|\operatorname{prod} w\| \\
&\le \sum_{\substack{w\in\{X,Y\}^k \\ \mathrm{m}_X\, w=k-m \\ \mathrm{m}_Y\, w=m}} \|X\|^{k-m}\|Y\|^m \\
&= \binom{k}{m}\|X\|^{k-m}\|Y\|^m;
\end{aligned}$$

so

$$\begin{aligned}
\|(X+Y)^k - X^k\| &= \left\| \sum_{m=1}^{k} p(X,Y,k-m,m) \right\| \\
&\le \sum_{m=1}^{k} \|p(X,Y,k-m,m)\| \\
&\le \sum_{m=1}^{k} \binom{k}{m}\|X\|^{k-m}\|Y\|^m \\
&= (\|X\|+\|Y\|)^k - \|X\|^k.
\end{aligned}$$

Finally, to see the continuity of $f(X) = \sum_{n=0}^{\infty} \lambda_n X^n$, if $X, Y \in \mathcal{A}$ are such that $\|X\| + \|Y\| < r$, then

$$\begin{aligned}
\|f(X+Y) - f(X)\| &\le \sum_{k=0}^{\infty} |\lambda_k| \|(X+Y)^k - X^k\| \\
&\le \sum_{k=0}^{\infty} |\lambda_k| [(\|X\|+\|Y\|)^k - \|X\|^k] \\
&= h(\|X\|+\|Y\|) - h(\|X\|)
\end{aligned}$$

where $h(x) = \sum_{k=0}^{\infty} |\lambda_k| x^k$ converges for $|x| < r$.

Hence $f(X+Y) \to f(X)$ as $\|Y\| \to 0$, that is, we have continuity. $\square$

2.2 The Exponential of a Matrix

Definition 2.8. The matrix exponential is defined as

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}, \qquad X \in M_n(\mathbb{C}),$$

and is well-defined and continuous in all $M_n(\mathbb{C})$ thanks to Proposition 2.7, for the radius of convergence of the complex power series $\sum_{m=0}^{\infty} \frac{z^m}{m!}$ equals $r = +\infty$.

Example 2.9. Let

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

We may compute that $X^2 = -I$, so that $X^3 = -X$ and $X^4 = I$. This recursion in the values of X^m allows to evaluate

$$e^{aX} = \sum_{m=0}^{\infty} (-1)^m \begin{pmatrix} \frac{a^{2m}}{(2m)!} & -\frac{a^{2m+1}}{(2m+1)!} \\ \frac{a^{2m+1}}{(2m+1)!} & \frac{a^{2m}}{(2m)!} \end{pmatrix} = \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}$$

We next state in Proposition 2.12 the properties of the matrix exponential; amongst of them it finds $e^X e^Y = e^{X+Y}$ for two matrices X, Y that commute. For its proof, we will need a result which allows us to multiply series "term by term," so that we can multiply out the series of e^X and e^Y in order to obtain that of e^{X+Y}.

Definition 2.10. Given two series $\sum_{n=0}^{\infty} X_n$, $\sum_{n=0}^{\infty} Y_n$ in some Banach algebra $\mathcal{A}$, with $X_n, Y_n \in \mathcal{A}$, we define its **Cauchy product** as

$$\sum_{n=0}^{\infty} X_n \cdot \sum_{n=0}^{\infty} Y_n = \sum_{n=0}^{\infty} Z_n, \tag{2.3}$$

where $Z_n = \sum_{k=0}^{n} X_k Y_{n-k}$.

Considering the case of complex numbers, in [Rud1] this definition is found to be motivated as follows. If we take two power series $\sum_{n=0}^{\infty} a_n z^n$ and $\sum_{n=0}^{\infty} b_m z^n$, multiply them term by term, and collect terms containing the same power of z, we get

$$\begin{aligned}
\sum_{n=0}^{\infty} a_n z^n \cdot \sum_{n=0}^{\infty} b_n z^n &= \left(a_0 + a_1 z + a_2 z^2 + \cdots\right)\left(b_0 + b_1 z + b_2 z^2 + \cdots\right) \\
&= a_0 b_0 + (a_0 b_1 + a_1 b_0)\, z + (a_0 b_2 + a_1 b_1 + a_2 b_0)\, z^2 + \cdots \\
&= c_0 + c_1 z + c_2 z^2 + \cdots .
\end{aligned}$$

Setting $z = 1$, we would arrive (2.3) for the case of complex numbers.

In general, given two series in a normed algebra that are convergent, its Cauchy product may not converge to the product of the series. However, we can give a sufficient condition for this to happen.

Proposition 2.11 (Mertens' Theorem for the Cauchy Product). *If $\sum_{n=0}^{\infty} x_n = X$ and $\sum_{n=0}^{\infty} y_n = Y$ are two convergent series in some normed algebra and at least one of them converges absolutely, then their Cauchy product converges to XY.*

Mertens' theorem can be regarded as the infinite series generalization of the ordinary distributive law of the product with respect to the sum.

We will give the same proof as that of Theorem 3.50 of [Rud1], although there's a difference between Rudin's statement of Mertens' theorem and ours: he states it for complex numbers series while we state it here for an arbitrary normed algebra.

Proof. Call $z_n = \sum_{k=0}^{n} x_k y_{n-k}$ and

$$X_n = \sum_{k=0}^{n} x_k, \quad Y_n = \sum_{k=0}^{n} y_k, \quad Z_n = \sum_{k=0}^{n} z_k, \quad \gamma_n = Y_n - Y.$$

Then

$$\begin{aligned} Z_n &= x_0 y_0 + (x_0 y_1 + x_1 y_0) + \cdots + (x_0 y_n + x_1 y_{n-1} + \cdots + x_n y_0) \\ &= x_0 Y_n + x_1 Y_{n-1} + \cdots + x_n Y_0 \\ &= x_0 (Y + \gamma_n) + x_1 (Y + \gamma_{n-1}) + \cdots + x_n (Y + \gamma_0) \\ &= X_n Y + x_0 \gamma_n + x_1 \gamma_{n-1} + \cdots + x_n \gamma_0. \end{aligned}$$

Put

$$\rho_n = x_0 \gamma_n + x_1 \gamma_{n-1} + \cdots + x_n \gamma_0.$$

We wish to show that $Z_n \to XY$. Since $X_n Y \to XY$, it suffices to show that

$$\lim_{n\to\infty} \rho_n = 0. \tag{2.4}$$

Suppose $\sum_{n=0}^{\infty} x_n$ converges absolutely and put

$$\chi = \sum_{n=0}^{\infty} \|x_n\| .$$

Let $\varepsilon > 0$ be given. Since $\gamma_n \to 0$, we can choose N such that $\|\gamma_n\| \leq \varepsilon$ for $n \geq N$, in which case

$$\begin{aligned} \|\rho_n\| &\leq \|\gamma_0 x_n + \cdots + \gamma_N x_{n-N}\| + \|\gamma_{N+1} x_{n-N-1} + \cdots + \gamma_n x_0\| \\ &\leq \|\gamma_0 x_n + \cdots + \gamma_N x_{n-N}\| + \varepsilon\chi. \end{aligned}$$

Keeping N fixed, and letting $n \to \infty$, we get

$$\limsup_{n\to\infty} \|\rho_n\| \leq \varepsilon\chi,$$

since $x_k \to 0$ as $k \to \infty$. Since ε is arbitrary, (2.4) follows. ◻

Proposition 2.12. *Let X, Y be arbitrary complex matrices. Then we have the following:*

1. $e^0 = I$.
2. $(e^X)^* = e^{X^*}$.
3. e^X *is invertible and* $(e^X)^{-1} = e^{-X}$.
4. $e^{(\alpha+\beta)X} = e^{\alpha X}e^{\beta X}$.
5. *If* $XY = YX$*, then* $e^{X+Y} = e^X e^Y$.
6. *If* C *is invertible, then* $e^{CXC^{-1}} = Ce^X C^{-1}$.

Proof. Point 1 is obvious and point 2 follows from continuity of $A \mapsto A^*$, which allows to take term-by-term adjoints of the series for e^X.

Points 3 and 4 are special cases of point 5. To verify point 5, we simply multiply the two power series term by term, which is permitted by Proposition 2.11 because both series converge absolutely. Multiplying out $e^X e^Y$ and collecting terms where the power of X plus the power of Y equals m, we obtain

$$e^X e^Y = \sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{X^k}{k!}\frac{Y^{m-k}}{(m-k)!} = \sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}X^k Y^{m-k}. \tag{2.5}$$

Now, because (and *only* because) X and Y commute,

$$(X+Y)^m = \sum_{k=0}^{m}\frac{m!}{k!(m-k)!}X^k Y^{m-k},$$

and, thus, (2.5) becomes

$$e^X e^Y = \sum_{m=0}^{\infty}\frac{1}{m!}(X+Y)^m = e^{X+Y}.$$

To prove point 6, simply note that

$$(CXC^{-1})^m = CX^mC^{-1}$$

and, thus by continuity of matrix multiplication, the two sides of point 6 are equal term by term. □

The previous proof is formally the same for a Banach algebra, but we have it here for complex matrices since that is the main focus of the thesis.

Proposition 2.13. *Let X be a $n \times n$ complex matrix. Then e^{tX} is a smooth curve in $M_n(\mathbb{C})$ and*

$$\frac{d}{dt}e^{tX} = Xe^{tX} = e^{tX}X.$$

In particular,

$$\left.\frac{d}{dt}e^{tX}\right|_{t=0} = X.$$

Proof. We will prove that

$$\lim_{\varepsilon \to 0} \left\| \frac{e^{(t+\varepsilon)X} - e^{tX}}{\varepsilon} - Xe^{tX} \right\| = 0.$$

We have

$$\begin{aligned}
\left\| \frac{e^{(t+\varepsilon)X} - e^{tX}}{\varepsilon} - Xe^{tX} \right\| &= \frac{1}{|\varepsilon|} \left\| e^{(t+\varepsilon)X} - e^{tX} - \varepsilon X e^{tX} \right\| \\
&\leq \frac{1}{|\varepsilon|} \left\| e^{tX} \right\| \left\| e^{\varepsilon X} - I - \varepsilon X \right\| \\
&= \frac{1}{|\varepsilon|} \left\| e^{tX} \right\| \left\| \sum_{n=2}^{\infty} \frac{(\varepsilon X)^n}{n!} \right\| \\
&\leq \frac{1}{|\varepsilon|} \left\| e^{tX} \right\| \sum_{n=2}^{\infty} \frac{(|\varepsilon| \, \|X\|)^n}{n!} \\
&= |\varepsilon| \left\| e^{tX} \right\| \|X\|^2 \sum_{n=2}^{\infty} \frac{(|\varepsilon| \, \|X\|)^{n-2}}{n!} \\
&\leq |\varepsilon| \left\| e^{tX} \right\| \|X\|^2 \sum_{n=2}^{\infty} \frac{(|\varepsilon| \, \|X\|)^{n-2}}{(n-2)!} \\
&= |\varepsilon| \left\| e^{tX} \right\| \|X\|^2 \, e^{|\varepsilon| \|X\|} \longrightarrow 0 \quad \text{as} \quad \varepsilon \longrightarrow 0.
\end{aligned}$$

□

2.3 The Matrix Logarithm

Lemma 2.14. *The function*

$$\log z = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m}$$

is defined and analytic in a circle of radius 1 *about* $z = 1$.

For all z *with* $|z - 1| < 1$,

$$e^{\log z} = z.$$

For all u *with* $|u| < \log 2$, *we have* $|e^u - 1| < 1$ *and*

$$\log e^u = u.$$

Proof. The usual logarithm for real, positive numbers satisfies

$$\frac{d}{dx} \log(1 - x) = \frac{-1}{1 - x} = -\left(1 + x + x^2 + \cdots\right)$$

for $|x| < 1$. Integrating term by term and noting that $\log 1 = 0$ gives

$$\log(1-x) = -\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots\right).$$

Taking $z = 1 - x$ (so that $x = 1 - z$), we have

$$\begin{aligned}\log z &= -\left((1-z) + \frac{(1-z)^2}{2} + \frac{(1-z)^3}{3} + \cdots\right) \\ &= \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m}. \qquad (2.6)\end{aligned}$$

The series (2.6) has radius of convergence 1 and defines an holomorphic function on the set $\{|z-1| < 1\}$, which coincides with the usual logarithm for real z in the interval $(0,2)$. Now $\exp(\log z) = z$ for $z \in (0,2)$ and since both sides of this identity are holomorphic in z, the identity continues to hold on the whole set $\{|z-1| < 1\}$.

On the other hand, if $|u| < \log 2$, then

$$|e^u - 1| = \left| u + \frac{u^2}{2!} + \cdots \right| \leq |u| + \frac{|u|^2}{2!} + \cdots = e^{|u|} - 1 < 1.$$

Thus, $\log(\exp u)$ makes sense for all such u. Since $\log(\exp u) = u$ for real u with $|u| < \log 2$, it follows by holomorphicity that $\log(\exp u) = u$ for all complex numbers with $|u| < \log 2$. □

Definition 2.15. The matrix logarithm is defined as

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m}, \quad A \in M_n(\mathbb{C}),$$

whenever the series converges.

By the previous lemma and Proposition 2.7, the matrix logarithm is well-defined and is continuous for $A \in M_n(\mathbb{K})$, $\|A - I\| < 1$, where $\|\cdot\|$ is any matrix norm.

Lemma 2.16. *The set of diagonalizable matrices is dense in $M_n(\mathbb{C})$.*

Proof. Let $X \in M_n(\mathbb{C})$. Every complex square matrix is triangularizable (see for example [Axl], 5.27), so there exists $A \in \mathsf{GL}(n;\mathbb{C})$ such that $X = ATA^{-1}$, where $T = (t_{jk})$ is upper-triangular. Eigenvalues are preserved by matrix conjugation:

$$|X - \lambda I| = |ATA^{-1} - \lambda I| = |A(T - \lambda I)A^{-1}| = |A||T - \lambda I||A^{-1}| = |T - \lambda I|,$$

so the eigenvalues of X are thus the diagonal entries of T. Were they all different and we would know then that X was diagonalizable. But in general X can have repeated eigenvalues.

We want to find then a sequence T_m of diagonalizable matrices such that $T_m \to T$. Define

$$d = \begin{cases} 1 & \text{if } t_{11} = t_{22} = \cdots = t_{nn}, \\ \min\limits_{\substack{t,t' \in \{t_{11}, t_{22}, \ldots, t_{nn}\} \\ t \neq t'}} |t - t'| & \text{otherwise.} \end{cases}$$

For $m \geq 1$, let $T_m = (t_{jk}^{(m)})$ be the $n \times n$ complex matrix defined by

$$t_{jk}^{(m)} = \begin{cases} t_{jk} & j \neq k, \\ t_{jj} + \dfrac{\delta_j}{m} & j = k, \end{cases}$$

where $\delta_j = \frac{d}{q_j+1}$ and $q_j = \#\{l \in \{1, 2, \ldots, j\} : t_{ll} = t_{jj}\}$, that is, q_j is the quantity of t_{ll}'s with $l \leq j$ and $t_{ll} = t_{jj}$.

In other words, T_m is a triangular matrix identical to T except on its diagonal entries. To them, a little quantity has been added accordingly so that now we can assure that $t_{jj}^{(m)} = t_{kk}^{(m)}$ if and only if $j = k$. Since the t_{jj}'s are the eigenvalues of T_m, we have that T_m is diagonalizable.[1] Since matrix conjugation preserves diagonalizability (the matrix $B = ECE^{-1}$ is diagonalizable if and only if C is diagonalizable), AT_mA^{-1} is diagonalizable and by construction $\lim_m AT_mA^{-1} = A(\lim_m T_m)A^{-1} = ATA^{-1} = X$. This proves that for any matrix $X \in M_n(\mathbb{C})$ there exists a sequence $\{X_m\}_{m\geq1}$ of diagonalizable matrices such that $X_m \to X$. In other words, the set of triangularizable matrices is dense in $M_n(\mathbb{C})$. □

Lemma 2.17. *Let $\|\cdot\|$ be any matrix norm in $M_n(\mathbb{K})$. Then for every $X \in M_n(\mathbb{K})$ and every eigenvalue $\lambda \in \mathbb{K}$ of X, we have $|\lambda| \leq \|X\|$.*

Proof. Let $v \in \mathbb{K}^n$ be an eigenvector of $X \in M_n(\mathbb{K})$ of eigenvalue $\lambda \in \mathbb{K}$. Denoting by $(v|0|\cdots|0)$ the $n \times n$ matrix whose first column is v and with the rest of its entries equal to zero, we have $A(v|0|\cdots|0) = \lambda(v|0|\cdots|0)$, so that $\|A(v|0|\cdots|0)\| = |\lambda|\,\|(v|0|\cdots|0)\|$. Since $\|\cdot\|$ is submultiplicative, it follows that

$$|\lambda|\,\|(v|0|\cdots|0)\| \leq \|A\|\,\|(v|0|\cdots|0)\|\,.$$

Since the matrix $(v|0|\cdots|0)$ is not the zero matrix (eigenvectors, by definition, are non-zero), its norm is non-zero and we conclude $|\lambda| \leq \|A\|$. □

Theorem 2.18. *The function*

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A - I)^m}{m} \tag{2.7}$$

[1] In more detail: since T_m has n different eigenvalues, it has n eigenvectors of a different eigenvalue each. Since eigenvectors of different eigenvalues are always linearly independent (see for example [Axl], 5.10), for T_m there exists a set of n linearly independent eigenvectors, so this set must be then also a generating system and thus a basis. That is, T_m has a basis of eigenvectors, i. e., T_m is diagonalizable.

is defined and continuous on the set of $n \times n$ complex matrices A with $\|A - I\| < 1$,

For all A with $\|A - I\| < 1$,

$$e^{\log A} = A.$$

For all X with $\|X\| < \log 2$, $\left\|e^X - I\right\| < 1$ and

$$\log e^X = X.$$

Proof. Convergence and continuity of the series (2.7) for $\|A - I\| < 1$ follows from Proposition 2.7 and Lemma 2.14.

Suppose now that A satisfies $\|A - I\| < 1$. If A is diagonalizable with eigenvalues $z_1, \ldots, z_n$, then we can express A in the form CDC^{-1} with D diagonal, in which case

$$(A - I)^m = C \begin{pmatrix} (z_1 - 1)^m & & 0 \\ & \ddots & \\ 0 & & (z_n - 1)^m \end{pmatrix} C^{-1}.$$

Since $\|A - I\| < 1$, each eigenvalue z_j of A must satisfy $|z_j - 1| < 1$ by Lemma 2.17 (λ is an eigenvalue for A if and only if $\lambda - 1$ is an eigenvalue for $A - I$). Thus,

$$\sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A - I)^m}{m} = C \begin{pmatrix} \log z_1 & & 0 \\ & \ddots & \\ 0 & & \log z_n \end{pmatrix} C^{-1},$$

and by Lemma 2.14,

$$e^{\log A} = C \begin{pmatrix} e^{\log z_1} & & 0 \\ & \ddots & \\ 0 & & e^{\log z_n} \end{pmatrix} C^{-1} = A.$$

By continuity and density of diagonalizable matrices (Lemma 2.16), we get $e^{\log A} = A$ for all A with $\|A - I\| < 1$.

Now, the same argument as in the complex case shows that if $\|X\| < \log 2$, then $\left\|e^X - I\right\| < 1$ and thus $\log e^X$ is defined. The proof that $\log e^X = X$ is then very similar to the proof that $e^{\log A} = A$. □

In particular, since exp always gives back an invertible matrix, it deduces that if $A \in M_n(\mathbb{C})$ is such that $\|A - I\| < 1$, then A is invertible.

Proposition 2.19. *There exists a constant $c > 0$ such that for all $n \times n$ matrices B with $\|B\| < \frac{1}{2}$,*

$$\|\log(I + B) - B\| \leq c \|B\|^2.$$

Proof. Note that

$$\log(I+B)-B=\sum_{m=2}^{\infty}(-1)^{m+1}\frac{B^m}{m}=B^2\sum_{m=2}^{\infty}(-1)^{m+1}\frac{B^{m-2}}{m}$$

so that

$$\|\log(I+B)-B\|\leq\|B\|^2\sum_{m=2}^{\infty}\frac{\left(\frac{1}{2}\right)^{m-2}}{m},$$

which is an estimate of the desired form. □

We may restate the proposition in a more concise way by saying that

$$\log(I+B)=B+O\left(\|B\|^2\right),$$

where $O\left(\|B\|^2\right)$ denotes a quantity of order $\|B\|^2$ (i.e., a quantity that is bounded by a constant times $\|B\|^2$ for all sufficiently small values of $\|B\|$).

2.4 Further Properties of the Exponential

Theorem 2.20 (Lie Product Formula). *For al $X,Y\in M_n(\mathbb{C})$, we have*

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

Proof. If we multiply the power series for $e^{\frac{X}{m}}$ and $e^{\frac{Y}{m}}$ term by term (we can do this thanks to Mertens' theorem, 2.11), all but three terms will involve $1/m^2$ or higher powers of $1/m$. Thus

$$e^{\frac{X}{m}}e^{\frac{Y}{m}}=I+\frac{X}{m}+\frac{Y}{m}+O\left(\frac{1}{m^2}\right).$$

Now, since $e^{\frac{X}{m}}e^{\frac{Y}{m}}\longrightarrow I$ as $m\longrightarrow\infty$, $e^{\frac{X}{m}}e^{\frac{Y}{m}}$ is in the domain of the logarithm for all sufficiently large m. By Proposition 2.19,

$$\begin{aligned}\log\left(e^{\frac{X}{m}}e^{\frac{Y}{m}}\right)&=\log\left(I+\frac{X}{m}+\frac{Y}{m}+O\left(\frac{1}{m^2}\right)\right)\\&=\frac{X}{m}+\frac{Y}{m}+O\left(\left\|\frac{X}{m}+\frac{Y}{m}+O\left(\frac{1}{m^2}\right)\right\|^2\right)\\&=\frac{X}{m}+\frac{Y}{m}+O\left(\frac{1}{m^2}\right),\end{aligned}$$

since

$$\left\|\frac{X}{m}+\frac{Y}{m}+O\left(\frac{1}{m^2}\right)\right\|^2=\frac{1}{m^2}\left\|X+Y+O\left(\frac{1}{m}\right)\right\|^2$$

$$
\begin{aligned}
&\leq \frac{1}{m^2}\left(\|X\| + \|Y\| + \left\|O\left(\frac{1}{m}\right)\right\|\right)^2 \\
&\leq \frac{1}{m^2}\left(\|X\| + \|Y\| + \frac{C}{m}\right)^2 && \text{for some } C > 0 \\
&\leq \frac{1}{m^2}(\|X\| + \|Y\| + C)^2 \\
&= O\left(\frac{1}{m^2}\right).
\end{aligned}
$$

Exponentiating the logarithm then gives

$$e^{\frac{X}{m}} e^{\frac{Y}{m}} = \exp\left(\frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right)$$

and therefore

$$\left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right)^m = \exp\left(X + Y + O\left(\frac{1}{m}\right)\right).$$

Thus, by continuity of the exponential, we conclude that

$$\lim_{m\to\infty}\left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right) = \exp(X+Y),$$

which is the Lie product formula. □

Theorem 2.21. *For any $X \in M_n(\mathbb{C})$, we have*

$$\det(e^X) = e^{\operatorname{trace}(X)}. \tag{2.8}$$

Proof. If X is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$, then e^X is diagonalizable with eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_n}$ (by Proposition 2.12, Point 6). Thus, since the determinant and trace of a matrix equals respectively the product and sum of its eigenvalues,

$$\det(e^X) = e^{\lambda_1}\cdots e^{\lambda_n} = e^{\lambda_1+\cdots+\lambda_n} = e^{\operatorname{trace}(X)}.$$

Since both sides of (2.8) are continuous functions of X, by Lemma 2.16 the identity follows for all $X \in M_n(\mathbb{C})$. □

Definition 2.22. A **one-parameter subgroup** of $\mathsf{GL}(n;\mathbb{C})$ is a group homomorphism $A : (\mathbb{R}, +) \to \mathsf{GL}(n;\mathbb{C})$ which is continuous.

Theorem 2.23 (One-parameter subgroups). *If $A(\cdot)$ is a one-parameter subgroup of $\mathsf{GL}(n;\mathbb{C})$, there exists a unique $n \times n$ complex matrix X such that*

$$A(t) = e^{tX}.$$

To prove the theorem we need the following lemma.

Lemma 2.24. *Fix some $\varepsilon \in (0, \log 2)$. Let $B_{\varepsilon/2}$ be the ball of radius $\varepsilon/2$ around the origin in $M_n(\mathbb{C})$, and let $U = \exp(B_{\varepsilon/2})$. Then every $B \in U$ has a unique square root C in U, given by $C = \exp(\frac{1}{2}\log B)$.*

Proof. By construction, C is a square root of B, since $B_{\varepsilon/2} \underset{\log}{\overset{\exp}{\rightleftarrows}} U$ are inverses of each other, by Theorem 2.18. Also, C is in U, as for $\log B \in B_{\varepsilon/2}$, also $\frac{1}{2}\log B \in B_{\varepsilon/2}$, and hence $C = \exp(\frac{1}{2}\log B) \in U$.

To establish uniqueness, suppose $\tilde{C} \in U$ satisfies $\tilde{C}^2 = B$. Let $Y = \log \tilde{C}$; then $\exp(Y) = \tilde{C}$ and

$$\exp(2Y) = \tilde{C}^2 = B = \exp(\log B).$$

We have that $Y \in B_{\varepsilon/2}$ and, thus, $2Y \in B_{\varepsilon}$, and also that $\log B \in B_{\varepsilon/2} \subset B_{\varepsilon}$. Since, by Theorem 2.18, exp is injective on B_{ε} and $\exp(2Y) = \exp(\log B)$, we must have $2Y = \log B$. Thus $\tilde{C} = \exp(Y) = \exp(\frac{1}{2}\log B) = C$. $\square$

Proof of Theorem 2.23. The uniqueness is immediate, since if there is such an X, then $X = \frac{d}{dt}A(t)\big|_{t=0}$. To prove existence, let U be as in Lemma 2.24, which is an open set of $\mathsf{GL}(n;\mathbb{C})$. Indeed, exp always gives back an invertible matrix, so $U \subset \mathsf{GL}(n;\mathbb{C})$; and secondly, since log maps U onto $B_{\varepsilon/2}$, we have $U = \log^{-1}(B_{\varepsilon/2})$ for log being injective. By the continuity of log, the set U is open in the domain of log, $\{A \in M_n(\mathbb{C}) : \|A - I\| < 1\} \subset \mathsf{GL}(n;\mathbb{C})$, which is an open set of $\mathsf{GL}(n;\mathbb{C})$, so U is open in $\mathsf{GL}(n;\mathbb{C})$.

Since A is a group homomorphism, $A(0) = I$. Since U is an open neighborhood of I, the continuity of A guarantees that there exists $t_0 > 0$ such that $A(t) \in U$ for all t with $|t| \le t_0$. Define

$$X = \frac{1}{t_0}\log(A(t_0)),$$

so that $t_0 X = \log(A(t_0))$. Then $t_0 X \in B_{\varepsilon/2}$ and

$$e^{t_0 X} = A(t_0).$$

Now, $A(t_0/2)$ is again in U and $A(t_0/2)^2 = A(t_0)$. But by Lemma 2.24, $A(t_0)$ has a unique square root in U; namely, $\exp(\frac{1}{2}\log A(t_0)) = \exp(t_0 X/2)$. Thus,

$$A(t_0/2) = \exp(t_0 X/2).$$

Applying this argument repeatedly, we conclude that

$$A(t_0/2^k) = \exp(t_0 X/2^k)$$

for all positive integers k. Then, since A is a group homomorphism, for any integer m we have

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

It follows that $A(t) = \exp(tX)$ for all real numbers t of the form $t = mt_0/2^k$, and the set of such t's is dense in $\mathbb{R}$. Since both $\exp(tX)$ and $A(t)$ are continuous, it follows that $A(t) = \exp(tX)$ for all real numbers t. $\square$

2.5 The Differentiability of a Matrix Power Series

As a conclusion to the chapter, we devote this section to the proof of the differentiability of a matrix power series. The result is stated and proven in Theorem 2.28. We are only interested in real differentiability, as this is what it is needed to study Lie groups from the real manifolds perspective. Even though the results from this section deal with one-variable or several-variables complex functions, the focus on real differentiability will be made patent in their statements.

A matrix power series can be regarded as a special type of several complex variables power series. For that matter, in order to study the differentiability of the former, one must be acquainted before with the differentiability of the latter. Lemma 2.25 gives the differentiability situation for a several complex variables power series.

Right before stating the lemma, we introduce the *multi-index notation* for multivariate power series. This notation eases the reasonings with these series and is therefore ubiquitous in texts from several complex variables theory. From now on, in this section $\mathbb{N} = \{0, 1, 2, \dots\}$ will denote the set of natural numbers with zero. Let $\alpha \in \mathbb{N}^n$ be a vector of non-negative integers. For $z \in \mathbb{C}^n$, the multi-index notation is

$$z^\alpha \overset{\text{def}}{=} z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n}, \quad |z^\alpha| \overset{\text{def}}{=} |z_1|^{\alpha_1} |z_2|^{\alpha_2} \cdots |z_n|^{\alpha_n}, \quad |\alpha| \overset{\text{def}}{=} \alpha_1 + \alpha_2 + \cdots + \alpha_n.$$

Lastly, if (X, d) is some metric space and $x \in X$, for $\varepsilon > 0$ we denote by $B_X(x, \varepsilon) = \{y \in X \mid d(x, y) < \varepsilon\}$ the open ball of radius ε and center x.

Lemma 2.25. *If the complex multivariate power series $f(z) = \sum_{\alpha \in \mathbb{N}^n} c_\alpha z^\alpha$ converges absolutely when $z = w \in \mathbb{C}^n$, then the series converges absolutely for all $z \in \prod_{j=1}^n B_{\mathbb{C}}(0, |w_j|) = B_{\mathbb{C}}(0, |w_1|) \times \cdots \times B_{\mathbb{C}}(0, |w_n|) =: D_w$. Furthermore, in this case and for each j, f is holomorphic with respect to to the variable z_j in all D_w. Therefore, with the canonical identification $\mathbb{C} \cong \mathbb{R}^2$ we have $D_w \subset \mathbb{R}^{2n}$ and that f is infinitely differentiable in D_w as a function of several real variables.*

Before giving the proof of the lemma, we need two auxiliary results. They are the Weierstraß M-test and the term-by-term differentiation theorem of a single complex variable power series.

Proposition 2.26 (Weierstraß M-test). *Suppose that $f_k : X \longrightarrow \mathcal{B}$ is a sequence of functions, where X is a set and $\mathcal{B}$ is a Banach space, and that there is sequence of non-negative numbers $\{M_k\}$ satisfying the conditions*

- *$\|f_k(x)\| \leq M_k$ for all $n \geq 1$ and all $x \in X$, and*
- *$\sum_{k=1}^\infty M_k$ converges.*

Then the series $\sum_{k=1}^\infty f_k(x)$ converges absolutely and uniformly on X.

(We won't be using the uniformity of the series convergence.) A proof can be found, for example, in [Rud1], Theorem 7.10. Although there it is proven for $\mathcal{B} = \mathbb{C}$, the proof is formally identical for a general Banach space.

Lemma 2.27 (Term-by-term power series differentiation). *Let $f(z) = \sum_{n=0} a_n z^n$ be a univariate complex power series with radius of convergence $R > 0$. Then f is differentiable in $B_{\mathbb{C}}(0, R)$ and*

$$f'(z) = \sum_{n=1}^{\infty} a_n n z^{n-1} \tag{2.9}$$

for every $z \in B_{\mathbb{C}}(0, R)$. Furthermore, the power series (2.9) has same radius of convergence R. Therefore f is infinitely differentiable in $B_{\mathbb{C}}(0, R)$.

For a proof, see for example [Rud2], Theorem 10.6. Since f' is also a power series, the previous lemma applies to f' itself and we get that f'' also exists and is written as a power series with same radius of convergence. This way, proceeding iteratively we obtain that $f^{(n)}$ exists in all the domain of f, or in other words, that f is infinitely differentiable.

The absolute convergence part of the following proof is taken from [Boas], slide #4.

Proof of Lemma 2.25. By hypothesis, there is a constant M such that $|c_\alpha w^\alpha| \leq M$ for all $\alpha \in \mathbb{N}^n$. If some w_j is zero, then the statement is vacuously true. Suppose $w_j \neq 0$ for every j. Now, for each j, let $r_j \in (0, |w_j|)$ and pick some $\lambda \in (0, 1)$ such that, for every j, it is $r_j \leq \lambda |w_j|$. In that case, for $z \in B_{\mathbb{C}}(0, r_1) \times \cdots \times B_{\mathbb{C}}(0, r_n)$, and if we call $r = (r_1, \ldots, r_n)$, we have

$$|c_\alpha z^\alpha| \leq |c_\alpha r^\alpha| \leq |c_\alpha \lambda^{|\alpha|} w^\alpha| \leq M \lambda^{|\alpha|}.$$

Now

$$\sum_{\alpha \in \mathbb{N}^n} \lambda^{|\alpha|} = \sum_{\alpha_1=0}^{\infty} \cdots \sum_{\alpha_n=0}^{\infty} \lambda^{\alpha_1} \cdots \lambda^{\alpha_n} = \frac{1}{(1-\lambda)^n}$$

by Fubini–Tonelli theorem for series with general term of several indices.

Therefore $\sum_{\alpha \in \mathbb{N}^n} c_\alpha z^\alpha$ converges absolutely by the Weierstraß M-test. This ends the proof of the first part of the lemma.

Let us now see about the differentiability of f in the domain D_w. We shall see that all partial derivatives $\frac{\partial^m f}{\partial z_j^m}$ of all orders $m \geq 1$ exist in D_w. In that case, since a function of one complex variable that is complex differentiable at some point is real differentiable at that point too (when considering $\mathbb{C} \cong \mathbb{R}^2$), we will obtain that f is real differentiable in $D_w \subset \mathbb{R}^{2n}$ as a function of several real variables.

For the sake of notational clarity, let us suppose $j = n$. The proof for existence of $\frac{\partial^m f}{\partial z_j^m}$ for the other j's will be the same.

Since $f(z) = \sum_{\alpha\in\mathbb{N}^n} c_\alpha z^\alpha$ is absolutely convergent in D_w, it is unconditionally convergent in D, and for that matter, there is no problem in reordering the terms in the sum:

$$f(z) = \sum_{\alpha_n=1}^{\infty} \sum_{\alpha\in\mathbb{N}^{n-1}} c_{(\alpha,\alpha_n)} z^{(\alpha,\alpha_n)} = \sum_{\alpha_1=1}^{\infty} \left(\sum_{\alpha\in\mathbb{N}^{n-1}} c_{(\alpha,\alpha_n)} z_1^{\alpha_1} \cdots z_{n-1}^{\alpha_{n-1}} \right) z_n^{\alpha_n}, \tag{2.10}$$

where we have applied the Fubini-Tonelli theorem for series. Now, for each $z \in D_w$, the series (2.10) has the form of a complex power series in z_n; and furthermore, there exists $\varepsilon > 0$ such that $\{(z_1, \dots, z_{n-1})\} \times B_{\mathbb{C}}(z_n, \varepsilon) \subset D_w$. That is, for fixed $(z_1, \dots, z_{n-1}) \in \pi(D_w)$, where $\pi : z \in \mathbb{C}^n \mapsto (z_1, \dots, z_{n-1})$, there is an open disk of possible values of z_n where (2.10) converges. Thus, by Lemma 2.27, f is infinitely complex differentiable with respect to z_n. □

The following theorem corresponds to Proposition 2.16 of [Hall1]. On his book, Hall states the result just for the matrix exponential, although from close inspection, his proof seems to be applicable to an arbitrary matrix power series. In fact, Hall implicitly admits the differentiability of the matrix logarithm in the proofs of the results of Sect. 3.8 of his book (which correspond to our Sect. 3.8). For this reason, here we state and prove the result in general.

Theorem 2.28. *Every analytic function $f : U \subset \mathbb{C} \to \mathbb{C}$, where U is an open subset of $\mathbb{C}$, induces an infinitely differentiable function[2] $\tilde{f} : \tilde{U} \subset M_n(\mathbb{C}) \to M_n(\mathbb{C})$ defined by the rule:*

$$\begin{aligned} &\text{If } c \in U \text{ and } \varepsilon > 0 \text{ are such that} \quad f(z) = \sum_{m=0}^{\infty} a_m (z-c)^m \quad \text{for } z \in B_{\mathbb{C}}(c,\varepsilon) \subset U, \\ &\text{then} \quad \tilde{f}(Z) = \sum_{m=0}^{\infty} a_m (Z - cI)^m \quad \text{for } Z \in B_{M_n(\mathbb{C})}(cI, \varepsilon). \end{aligned} \tag{2.11}$$

Where the set $\tilde{U}$ is the union of all such $B_{M_n(\mathbb{C})}(cI;\varepsilon)$, and thus it is open.

Proof. By Proposition 2.7, we know that the matrix power series given in (2.11) converges and depends continuously on $Z \in B_{M_n(\mathbb{C})}(cI, \varepsilon)$.

We must ensure that the value for $\tilde{f}$ of (2.11) is independent of the chosen power series representation. That is, we must verify that if $c, d \in U$ and $\varepsilon, \delta > 0$ are such that

$$\begin{aligned} f(z) &= \sum_{m=0}^{\infty} a_m (z-c)^m, \quad z \in B_{\mathbb{C}}(c,\varepsilon) \subset U, \\ f(z) &= \sum_{m=0}^{\infty} b_m (z-d)^m, \quad z \in B_{\mathbb{C}}(d,\delta) \subset U, \end{aligned} \tag{2.12}$$

[2] In the sense of real manifolds. Equivalently, since $M_n(\mathbb{C}) \cong \mathbb{R}^{2n^2}$, all partial derivatives of all orders exist.

then

$$\sum_{m=0}^{\infty} a_m (Z - cI)^m = \sum_{m=0}^{\infty} b_m (Z - dI)^m \quad \text{for all } Z \in B_{M_n(\mathbb{C})}(cI, \varepsilon) \cap B_{M_n(\mathbb{C})}(dI, \delta). \tag{2.13}$$

Let $c, d \in U$ and $\varepsilon, \delta > 0$ be fulfilling (2.12). Let $A \in B_{M_n(\mathbb{C})}(cI, \varepsilon) \cap B_{M_n(\mathbb{C})}(dI, \delta)$ be a diagonalizable matrix and $C \in \mathsf{GL}(n; \mathbb{C})$ be such that $A = CDC^{-1}$, where D is diagonal with diagonal equal to $(d_1, \dots, d_n)$. These are the eigenvalues of A. Now, by the Lemma 2.17, it follows that $|d_j - c| \le \|A - cI\| < \varepsilon$, since $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if $\lambda - c$ is an eigenvalue of $A - cI$. Thus $\sum_{m=0}^{\infty} a_m (d_j - c)^m$ converges and therefore

$$\begin{aligned}
\sum_{m=0}^{\infty} a_m (A - cI)^m &= \sum_{m=0}^{\infty} a_m \underbrace{(CDC^{-1} - cI)^m}_{(C[D-cI]C^{-1})^m = C(D-cI)^m C^{-1}} \\
&= \sum_{m=0}^{\infty} a_m C (D - cI)^m C^{-1} \\
&= \sum_{m=0}^{\infty} a_m C \begin{pmatrix} (d_1 - c)^m & & \\ & \ddots & \\ & & (d_n - c)^m \end{pmatrix} C^{-1} \\
&= C \left[\sum_{m=0}^{\infty} a_m \begin{pmatrix} (d_1 - c)^m & & \\ & \ddots & \\ & & (d_n - c)^m \end{pmatrix} \right] C^{-1} \\
&= C \begin{pmatrix} \sum_{m=0}^{\infty} a_m (d_1 - c)^m & & \\ & \ddots & \\ & & \sum_{m=0}^{\infty} a_m (d_n - c)^m \end{pmatrix} C^{-1} \\
&= C \begin{pmatrix} f(d_1) & & \\ & \ddots & \\ & & f(d_n) \end{pmatrix} C^{-1}.
\end{aligned}$$

By a totally analogous argument, we get

$$\sum_{m=0}^{\infty} b_m (A - dI)^m = C \begin{pmatrix} f(d_1) & & \\ & \ddots & \\ & & f(d_n) \end{pmatrix} C^{-1};$$

and so, by continuity and Lemma 2.16, we get (2.13).

Let's now see about the differentiability of $\tilde{f}$. Let $c \in U$ and pick some $\varepsilon > 0$ such that

$$f(z) = \sum_{m=0}^{\infty} a_m (z - c)^m, \quad z \in B_{\mathbb{C}}(c, \varepsilon) \subset U,$$

for certain coefficients $a_m \in \mathbb{C}$.

We must show that the function

$$\tilde{f}(Z) = \sum_{m=0}^{\infty} a_m (Z - cI)^m$$

is differentiable in $Z \in B_{M_n(\mathbb{C})}(cI, \varepsilon)$. Without loss of generality, we can assume $c = 0$, for if we perform the change of variable $W = Z - cI$, it suffices to show differentiability for

$$g(W) = \tilde{f}(W + cI) = \sum_{m=0}^{\infty} a_m W^m, \quad W \in B_{M_n(\mathbb{C})}(0, \varepsilon),$$

as that would imply differentiability for $\tilde{f}(Z) = g(Z - cI)$ in $B_{M_n(\mathbb{C})}(cI, \varepsilon)$.

For the sake of keeping the same notation, we write then

$$\tilde{f}(Z) = \sum_{m=0}^{\infty} a_m Z^m, \quad Z \in B_{M_n(\mathbb{C})}(0, \varepsilon).$$

Let $j, k \in \{1, \ldots, n\}$ be fixed. We shall show that $\tilde{f}_{jk} : B_{M_n(\mathbb{C})}(0, \varepsilon) \to \mathbb{C}$ given by $\tilde{f}_{jk}(Z) = (\tilde{f}(Z))_{jk}$ is differentiable. To this aim, we will apply Lemma 2.25. We must write the matrix power series

$$\tilde{f}_{jk}(Z) = \sum_{m=0}^{\infty} a_m (Z^m)_{jk}, \quad Z \in B_{M_n(\mathbb{C})}(0, \varepsilon) \tag{2.14}$$

as a several complex variables power series.

The series (2.14) converges absolutely:

$$\sum_{m=0}^{\infty} |a_m| |(Z^m)_{jk}| \leq \sum_{m=0}^{\infty} |a_m| \, \|Z^m\|_\infty \leq \sum_{m=0}^{\infty} |a_m| C \, \|Z^m\| \leq C \sum_{m=0}^{\infty} |a_m| \, \|Z\|^m < +\infty, \tag{2.15}$$

whenever $Z \in B_{M_n(\mathbb{C}), \|\cdot\|}(0, \varepsilon)$, where $\|\cdot\|_\infty$ is the infinity norm and where $C > 0$ is such that $\|X\|_\infty \leq C \, \|X\|$ for all $X \in M_n(\mathbb{C})$ (recall that we always state our results in terms of an arbitrary matrix norm $\|\cdot\|$).

It may seem that at first that (2.15) is the absolute convergence required to apply Lemma 2.25, and it might surprise you that this isn't true. What happens is that two different notions of *absolute convergence* have arisen, and the "absolute convergence" notion of (2.15) is different from the "absolute convergence" notion of Lemma 2.25. These notions are, respectively, those of equations (2.16) and (2.17). We will untangle this problem step-by-step.

Observe first that the term $(Z^m)_{jk}$ is a homogeneous polynomial of degree m in the variables $z_{11}, z_{12}, \dots, z_{nn}$. That is, we can write it as

$$(Z^m)_{jk} = \sum_{(j_1,k_1)\leq\cdots\leq(j_m,k_m)} c^{(jk)}_{j_1k_1\dots j_mk_m} z_{j_1k_1}\cdots z_{j_mk_m},$$

where $c^{(jk)}_{j_1k_1\dots j_mk_m}$ are non-negative integral coefficients and if $(a_1, b_1), (a_2, b_2) \in \mathbb{N}^2$, we write $(a_1, b_1) \leq (a_2, b_2)$ if either $a_1 < a_2$ or $a_1 = a_2$ and $b_1 \leq b_2$. That is, $\leq$ is the lexicographical order in $\mathbb{N}^2$. In other words: we are reading the matrix elements from left to right and from up to down.

With these notations, the absolute convergence of (2.14) is written as

$$\sum_{m=0}^{\infty} |a_m| \left| \sum_{(j_1,k_1)\leq\cdots\leq(j_m,k_m)} c^{(jk)}_{j_1k_1\dots j_mk_m} z_{j_1k_1}\cdots z_{j_mk_m} \right| < +\infty \tag{2.16}$$

for $Z \in B_{M_n(\mathbb{C})}(0, \varepsilon)$. Problem is, from (2.16) it does *not* follow that

$$\sum_{m=0}^{\infty} |a_m| \sum_{(j_1,k_1)\leq\cdots\leq(j_m,k_m)} c^{(jk)}_{j_1k_1\dots j_mk_m} \left|z_{j_1k_1}\cdots z_{j_mk_m}\right| < +\infty, \tag{2.17}$$

which is the actual condition we would need to apply Lemma 2.25.

We will see how to workaround this matter. Let $Z_0 \in B_{M_n(\mathbb{C})}(0, \varepsilon)$ be fixed. We shall see that f is differentiable in Z_0.

Define the matrix $Z_{\text{abs}} = (|z_{0,jk}|)$ to be the matrix of absolute values of the entries of Z_0. Then $Z_{\text{abs}} \in B_{M_n(\mathbb{C})}(0, \varepsilon)$. Let $\mathbb{1}$ be the $n \times n$ matrix whose entries are all equal to 1 and pick some $\delta > 0$ such that $Z' := Z_{\text{abs}} + \delta\mathbb{1} \in B_{M_n(\mathbb{C})}(0, \varepsilon)$ (such a δ always exists for $B_{M_n(\mathbb{C})}(0, \varepsilon)$ is open). This way, it is the case that $Z_0 \in D_{Z'}$, where

$$D_{Z'} = \prod_{j,k=1}^{n} B_{\mathbb{C}}(0, |(Z')_{jk}|) = \prod_{j,k=1}^{n} B_{\mathbb{C}}(0, |z_{0,jk}| + \delta)$$

(this notation is taken from Lemma 2.25). Now, since $\|Z'\| < \varepsilon$ the series $\tilde{f}_{jk}(Z')$ of (2.14) is convergent, and thus, it converges absolutely in the sense of (2.16), just as before. Moreover, because (and only because) the entries of Z' are all positive, it also converges absolutely in the sense of (2.17). Thus, by Lemma 2.25 the complex multivariate series

$$\sum_{m=0}^{\infty} a_m \sum_{(j_1,k_1)\leq\cdots\leq(j_m,k_m)} c^{(jk)}_{j_1k_1\dots j_mk_m} z_{j_1k_1}\cdots z_{j_mk_m}$$

converges in all $D_{Z'}$ and is infinitely differentiable there. So f is infinitely differentiable in $D_{Z'} \cap B_{M_n(\mathbb{C})}(0, \varepsilon)$, an open set which contains Z_0. □

In particular, the theorem applies for the matrix exponential and the matrix logarithm. Thus, the matrix exponential is everywhere C^∞ and the matrix logarithm is C^∞ in $B_{M_n(\mathbb{C})}(I, 1)$.

3 | Lie Algebras

3.1 Definitions and First Examples

We now introduce the "abstract" notion of a Lie algebra. In a coming section we will associate to each matrix Lie group a Lie algebra. It is customary to use lower case Gothic (Fraktur) characters such as $\mathfrak{g}$ and $\mathfrak{h}$ to refer to Lie algebras.

Definition 3.1. A **real or complex Lie algebra** is a real or complex algebra $(\mathfrak{g}, [\cdot,\cdot])$ such that

- $[\cdot,\cdot]$ is anti-symmetric: $[X,Y] = -[Y,X]$ for all $X, Y \in \mathfrak{g}$.
- The **Jacobi identity** holds:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$

  for all $X, Y, Z \in \mathfrak{g}$.

The map $[\cdot,\cdot]$ is referred to as the **Lie bracket** or **bracket** operation on $\mathfrak{g}$. If $(\mathcal{A}, [\cdot,\cdot])$ is a real or complex algebra, then its bilinear form $[\cdot,\cdot]$ is anti-symmetric if and only if it is alternating, i. e., $[X,X] = 0$ for every $X \in \mathcal{A}$. It is clear that anti-symmetry implies alternation. For the converse, if $[\cdot,\cdot]$ is alternating, then $0 = [X+Y, X+Y] = [X,X] + [Y,Y] + [X,Y] + [Y,X] = [X,Y] + [Y,X]$.

Note that the anti-symmetry condition for the Lie bracket implies that if $\mathfrak{g}$ is a Lie algebra, then $\mathfrak{g}$ is commutative if and only if its Lie bracket is trivial, $[X,Y] = 0$ for all $X, Y \in \mathfrak{g}$. In general, the Lie bracket of a Lie algebra is not associative; nevertheless, the Jacobi identity can be viewed as a substitute for associativity.

Example 3.2. Let $\mathfrak{g} = \mathbb{R}^3$ and let $[\cdot,\cdot] : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$[x,y] = x \times y,$$

where $x \times y$ is the cross product (or vector product). Then $\mathfrak{g}$ is a Lie algebra.

Proof. Bilinearity and anti-symmetry are standard properties of the cross product. To verify the Jacobi identity, it suffices (by bilinearity) to verify it when $x = e_j$, $y = e_k$ and $z = e_l$, where e_1, e_2 and e_3 are the standard basis elements for $\mathbb{R}^3$. If either $j = k = l$ or j, k, l are all distinct, then each term in the Jacobi identity is zero. It remains to consider the case in which two of j, k, l are equal and the third is different: we must verify the identity

$$\left[e_j, \left[e_j, e_k\right]\right] + \left[e_j, \left[e_k, e_j\right]\right] + \left[e_k, \left[e_j, e_j\right]\right] = 0. \tag{3.1}$$

The first two terms in (3.1) are negatives of each other and the third is zero. □

Given an associative algebra $\mathcal{A}$, we define its commutator $[\cdot,\cdot] : \mathcal{A} \times \mathcal{A} \longrightarrow \mathcal{A}$ as $[X, Y] = XY - YX$ for $X, Y \in \mathcal{A}$. This way, X and Y commute in $\mathcal{A}$ if and only if $[X, Y] = 0$.

Proposition 3.3. *Every associative algebra is also a Lie algebra with the commutator as the Lie bracket.*

Proof. The bilinearity and anti-symmetry of the bracket are evident. To verify the Jacobi identity, note that each double bracket generates four terms, for a total of 12 terms. It is left to the reader to verify that the product of X, Y and Z in each of the six possible orderings occurs twice, once with a plus sign and once with a minus sign. □

If we look carefully at the proof of the Jacobi identity, we see that the two occurrences of, say, XYZ occur with different groupings, once as $X(YZ)$ and once as $(XY)Z$. Thus, associativity of the algebra $\mathcal{A}$ is essential. For any Lie algebra, the Jacobi identity means that the bracket operation *behaves as if* it were $XY - YX$ in some associative algebra, even if it is not actually defined this way. As a curiosity, every Lie algebra can be embedded in an associative algebra in such a way that the Lie bracket becomes $XY - YX$, it is the *universal enveloping algebra* of a Lie algebra (cf. Sect. 9.3 of [Hall1] for more on this topic).

Of particular interest to us is the case in which $\mathcal{A}$ is the space $M_n(\mathbb{C})$ of $n \times n$ complex matrices.

Definition 3.4. A **subalgebra** of a real or complex Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h}$ of $\mathfrak{g}$ such that $[H_1, H_2] \in \mathfrak{h}$ for all $H_1, H_2 \in \mathfrak{h}$. If $\mathfrak{g}$ is a complex Lie algebra and $\mathfrak{h}$ is a real subspace of $\mathfrak{g}$ which is closed under brackets, then $\mathfrak{h}$ is said to be a **real subalgebra** of $\mathfrak{g}$.

A subalgebra of a Lie algebra is a Lie algebra by its own right, with the restriction of the Lie bracket to the subalgebra in question.

Example 3.5. Let $\mathsf{sl}(n;\mathbb{C})$ be the space $X \in M_n(\mathbb{C})$ for which $\operatorname{trace} X = 0$. Then $\mathsf{sl}(n;\mathbb{C})$ is a Lie subalgebra of $M_n(\mathbb{C})$.

Proof. For any X and Y in $M_n(\mathbb{C})$, we have

$$\operatorname{trace}(XY - YX) = \operatorname{trace}(XY) - \operatorname{trace}(YX) = 0,$$

since trace is invariant under cyclic permutations (Lemma 1.10). □

Definition 3.6. If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, then a linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ is called a **Lie algebra homomorphism** if it preserves Lie brackets; that is, $\phi([X, Y]) = [\phi(X), \phi(Y)]$ for all $X, Y \in \mathfrak{g}$. If, in addition, ϕ is bijective, then ϕ^{-1} is also a Lie algebra homomorphism and ϕ is then called a **Lie algebra isomorphism**. A Lie algebra isomorphism of a Lie algebra with itself is called a **Lie algebra automorphism**.

Indeed, if $\phi : \mathfrak{g} \to \mathfrak{h}$ is a bijective Lie algebra homomorphism, then ϕ^{-1} is linear and a Lie algebra homomorphism: for every $Z, H \in \mathfrak{h}$ there exists unique $X, Y \in \mathfrak{g}$ such that $\phi(X) = Z$, $\phi(Y) = H$ and we have

$$\phi^{-1}([Z, H]) = \phi^{-1}([\phi(X), \phi(Y)]) = \phi^{-1}(\phi([X, Y])) = [X, Y] = [\phi^{-1}(Z), \phi^{-1}(Y)].$$

Definition 3.7. If $\mathfrak{g}$ is a Lie algebra and X is an element of $\mathfrak{g}$, define a linear map $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ by

$$\operatorname{ad}_X(Y) = [X, Y].$$

The map $X \mapsto \operatorname{ad}_X$ is the **adjoint map** or **adjoint representation**.

Although $\operatorname{ad}_X(Y)$ is just $[X, Y]$, the alternative "ad" notation can be useful. For example, instead of writing

$$[X, [X, [X, [X, Y]]]],$$

we can now write

$$(\operatorname{ad}_X)^4(Y).$$

This sort of notation will be essential in chapter 5. We can view ad (that is, the map $X \mapsto \operatorname{ad}_X$) as a linear map of $\mathfrak{g}$ into $\operatorname{End}(\mathfrak{g})$, the space of linear operators on $\mathfrak{g}$. The Jacobi identity is interpretable in terms of the properties of ad_X. In a real or complex algebra $(\mathcal{A}, [\cdot, \cdot])$ a **derivation** is a linear map $D : \mathcal{A} \to \mathcal{A}$ that satisfies Leibniz's law:

$$D([X, Y]) = [D(X), Y] + [X, D(Y)].$$

The concept of a derivation generalizes the product rule $\frac{d(fg)}{dt} = \frac{df}{dt}g + f\frac{dg}{dt}$ for real differentiable functions. With this terminology, it turns out that the Jacobi identity is equivalent to the assertion that ad_X is a derivation in the Lie algebra:

$$\operatorname{ad}_X([Y, Z]) = [\operatorname{ad}_X(Y), Z] + [Y, \operatorname{ad}_X(Z)].$$

Proposition 3.8. *If $\mathfrak{g}$ is a Lie algebra, then*

$$\mathrm{ad}_{[X,Y]} = \mathrm{ad}_X \mathrm{ad}_Y - \mathrm{ad}_Y \mathrm{ad}_X = [\mathrm{ad}_X, \mathrm{ad}_Y];$$

that is, $\mathrm{ad} : \mathfrak{g} \longrightarrow \mathrm{End}(\mathfrak{g})$ *is a Lie algebra homomorphism.*

Proof. Observe that

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

whereas

$$[\mathrm{ad}_X, \mathrm{ad}_Y](Z) = [X,[Y,Z]] - [Y,[X,Z]].$$

Thus, we want to show that

$$[[X,Y],Z] = [X,[Y,Z]] - [Y,[X,Z]],$$

which is equivalent to the Jacobi identity. □

Definition 3.9. If $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are Lie algebras, the **direct sum** of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ is the vector space direct sum of $\mathfrak{g}_1$ and $\mathfrak{g}_2$, with the bracket given by

$$[(X_1, X_2), (Y_1, Y_2)] = ([X_1, Y_1], [X_2, Y_2]). \tag{3.2}$$

If $\mathfrak{g}$ is a Lie algebra and $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are subalgebras, we say that $\mathfrak{g}$ **decomposes as the Lie algebra direct sum** of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ if $\mathfrak{g}$ is the direct sum of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ as vector spaces and $[X_1, X_2] = 0$ for all $X_1 \in \mathfrak{g}_1$ and $X_2 \in \mathfrak{g}_2$.

It is straightforward to verify that the bracket in (3.2) makes $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ into a Lie algebra. If $\mathfrak{g}$ decomposes as a Lie algebra direct sum of subalgebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, it is easy to check that $\mathfrak{g}$ is isomorphic as a Lie algebra to the "abstract" direct sum of $\mathfrak{g}_1$ and $\mathfrak{g}_2$. (This would not be the case without the assumption that every element of $\mathfrak{g}_1$ commutes with every element of $\mathfrak{g}_2$.)

Definition 3.10. Let $\mathfrak{g}$ be a finite-dimensional real or complex Lie algebra, and let $X_1, \ldots, X_N$ be a basis for $\mathfrak{g}$ (as a vector space). Then the unique constants c_{jkl} such that

$$[X_j, X_k] = \sum_{l=1}^{N} c_{jkl} X_l$$

are called the **structure constants** of $\mathfrak{g}$ (with respect to the chosen basis).

Although we will not have much occasion to use them, structure constants do appear frequently in the physics literature. The structure constants satisfy the following two conditions:

$$c_{jkl} + c_{kjl} = 0,$$
$$\sum_n \left(c_{jkn}c_{nlm} + c_{kln}c_{njm} + c_{ljn}c_{nkm}\right) = 0$$

for all j, k, l, m. The first of these conditions comes from the anti-symmetry of the bracket, and the second comes from the Jacobi identity.

3.2 The Lie Algebra of a Matrix Lie Group

In this section, we associate to each matrix Lie group G a Lie algebra $\mathfrak{g}$. Many questions involving a group can be studied by transferring them to the Lie algebra, where we can use tools of linear algebra. We begin by defining $\mathfrak{g}$ as a set, and then proceed to give $\mathfrak{g}$ the structure of a Lie algebra.

Definition 3.11. Let G be a matrix Lie group. The **Lie algebra of** G, denoted $\mathfrak{g}$ or $\operatorname{Lie} G$, is the set of all matrices X such that e^{tX} is in G for all real numbers t.

Equivalently, X is in $\mathfrak{g}$ if and only if the entire one-parameter subgroup (Definition 2.22) generated by X lies in G. Note that merely having e^X in G does not guarantee that X is in $\mathfrak{g}$. Even though G is a subgroup of $\mathsf{GL}(n;\mathbb{C})$ (and not necessarily of $\mathsf{GL}(n;\mathbb{R})$), we do *not* require that e^{tX} be in G for all complex numbers t, but only for all *real* numbers t. We will show in Sect. 3.7 that every matrix Lie group is an embedded submanifold of $\mathsf{GL}(n;\mathbb{C})$. We will then show that $\mathfrak{g}$ is the **tangent space to** G **at the identity**.

We will now establish various basic properties of the Lie algebra $\mathfrak{g}$ of a matrix Lie group G. In particular, we will see that there is a bracket operation on $\mathfrak{g}$ that makes $\mathfrak{g}$ into a Lie algebra in the sense of definition 3.1.

Theorem 3.12. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. If X and Y are elements of $\mathfrak{g}$, the following results hold.*

1. *$AXA^{-1} \in \mathfrak{g}$ for all $A \in G$.*
2. *$sX \in \mathfrak{g}$ for all real numbers s.*
3. *$X + Y \in \mathfrak{g}$.*
4. *$XY - YX \in \mathfrak{g}$.*

It follows from this result and Proposition 3.3 that the Lie algebra of a matrix Lie group is a real Lie algebra, with bracket given by $[X, Y] = XY - YX$.

Proof. For Point 1, we observe that, by Proposition 2.12,

$$e^{t(AXA^{-1})} = Ae^{tX}A^{-1} \in G$$

for all t, showing that AXA^{-1} is in $\mathfrak{g}$. For Point 2, we observe that $e^{t(sX)} = e^{(ts)X}$ which must be in G for all $t \in \mathbb{R}$ if X is in $\mathfrak{g}$. For Point 3 we use the Lie product formula, which says that

$$e^{t(X+Y)} = \lim_{m\to\infty} (e^{tX/m}e^{tY/m})^m.$$

(Theorem 2.20.) Thus, $(e^{tX/m}e^{tY/m})^m$ is in G for all m. Since G is closed, the limit must be again in G. This shows that $X + Y$ is in $\mathfrak{g}$.

Finally, for Point 4, we use the product rule[1] and Proposition 2.13 to compute

$$\begin{aligned}\frac{d}{dt}\left(e^{tX}Ye^{-tX}\right)\bigg|_{t=0} &= (XY)e^0 + \left(e^0Y\right)(-X)\\ &= XY - YX.\end{aligned}$$

Now, by Point 1, $e^{tX}Ye^{-tX}$ is in $\mathfrak{g}$ for all t. Furthermore, by Points 2 and 3, $\mathfrak{g}$ is a real subspace of $M_n(\mathbb{C})$, from which it follows that $\mathfrak{g}$ is a (topologically) closed subset of $M_n(\mathbb{C})$. Thus,

$$XY - YX = \lim_{h\to 0}\frac{e^{hX}Ye^{-hX} - Y}{h}$$

belongs to $\mathfrak{g}$. □

Note that even if the elements of G have complex entries, the Lie algebra $\mathfrak{g}$ of G is not necessarily a complex vector space, since Point 2 holds, in general, only for $s \in \mathbb{R}$. Nevertheless it may happen in certain cases that $\mathfrak{g}$ is a complex vector space.

Definition 3.13. A matrix Lie group G is said to be **complex** if its Lie algebra $\mathfrak{g}$ is a complex subspace of $M_n(\mathbb{C})$, that is, if $iX \in \mathfrak{g}$ for all $X \in \mathfrak{g}$.

Examples of complex groups are $\mathsf{GL}(n;\mathbb{C})$ and $\mathsf{SL}(n;\mathbb{C})$, as the calculations in Sect. 3.3 will show.

Proposition 3.14. *If G is commutative then $\mathfrak{g}$ is commutative.*

We will see in Sect. that if G is path-connected and $\mathfrak{g}$ is commutative, G must be commutative.

Proof. For any two matrices $X, Y \in M_n(\mathbb{C})$, the commutator of X and Y may be computed as

$$[X, Y] = \frac{d}{dt}\left(\frac{d}{ds}e^{tX}e^{sY}e^{-tX}\bigg|_{s=0}\right)\bigg|_{t=0}. \tag{3.3}$$

If G is commutative and X and Y belong to $\mathfrak{g}$, then e^{tX} commutes with e^{sY} and the expression in parentheses on the right hand side of (3.3) is independent of t, so that $[X, Y] = 0$. □

[1]If $A, B : I \subset \mathbb{R} \to M_n(\mathbb{C})$ are differentiable matrix-valued functions, where I is an open interval, the product rule for matrices is $\frac{d}{dt}(A(t)B(t)) = \frac{dA(t)}{dt}B(t) + A(t)\frac{dB(t)}{dt}$. Indeed,

$$\begin{aligned}\frac{d}{dt}(A(t)B(t)) &= \lim_{h\to0}\frac{A(t+h)B(t+h) - A(t)B(t)}{h} = \lim_{h\to0}\frac{A(t+h)B(t+h) - A(t)B(t+h) + A(t)B(t+h) - A(t)B(t)}{h}\\ &= \lim_{h\to0}\left(\frac{A(t+h)-A(t)}{h}B(t+h)\right) + \lim_{h\to0}\left(A(t)\frac{B(t+h)-B(t)}{h}\right) = \frac{dA(t)}{dt}B(t) + A(t)\frac{dB(t)}{dt}.\end{aligned}$$

3.3 Examples

Physicists are accustomed to using the map $t \mapsto e^{itX}$ rather than $t \mapsto e^{tX}$. Thus, the physicists' expressions for the Lie algebras of matrix Lie groups will differ by a factor of i from the expressions we now derive.

Note that if G and H are matrix Lie groups, then $G \cap H$ is also a matrix Lie group and, by logic, we have $\mathrm{Lie}(G \cap H) = \mathrm{Lie}\, G \cap \mathrm{Lie}\, H$.

Proposition 3.15. *The Lie algebra of* $\mathsf{GL}(n;\mathbb{C})$ *is the space* $M_n(\mathbb{C})$ *of all* $n \times n$ *matrices with complex entries. Similarly, the Lie algebra of* $\mathsf{GL}(n;\mathbb{R})$ *is equal to* $M_n(\mathbb{R})$. *The Lie algebra of* $\mathsf{SL}(n;\mathbb{C})$ *consists of all* $n \times n$ *complex matrices with trace zero, and the Lie algebra of* $\mathsf{SL}(n;\mathbb{R})$ *consists of all* $n \times n$ *real matrices with trace zero.*

We denote the Lie algebras of these groups as $\mathsf{gl}(n;\mathbb{C})$, $\mathsf{gl}(n;\mathbb{R})$, $\mathsf{sl}(n;\mathbb{C})$, and $\mathsf{sl}(n;\mathbb{R})$, respectively.

Proof. If $X \in M_n(\mathbb{C})$, then e^{tX} is invertible, so that X belongs to the Lie algebra of $\mathsf{GL}(n;\mathbb{C})$. If $X \in M_n(\mathbb{R})$, then e^{tX} is invertible and real, so X is in the Lie algebra of $\mathsf{GL}(n;\mathbb{R})$. Conversely, if e^{tX} is real for all real t, then $X = d(e^{tX})/dt|_{t=0}$ must also be real. If $X \in M_n(\mathbb{C})$ has trace zero, then by Theorem 2.21, $\det(e^{tX}) = 1$, showing that X is in the Lie algebra of $\mathsf{SL}(n;\mathbb{C})$. Conversely, if $\det(e^{tX}) = e^{t\,\mathrm{trace}(X)} = 1$ for all real t, then

$$\mathrm{trace}(X) = \frac{d}{dt} e^{t\,\mathrm{trace}(X)}\bigg|_{t=0} = 0.$$

Finally, the Lie algebra of $\mathsf{SL}(n;\mathbb{R}) = \mathsf{SL}(n;\mathbb{C}) \cap \mathsf{GL}(n;\mathbb{R})$ equals $\mathsf{sl}(n;\mathbb{C}) \cap \mathsf{gl}(n;\mathbb{R})$, the set of $n \times n$ real matrices with trace zero. ☐

Proposition 3.16. *The Lie algebra of* $\mathsf{U}(n)$ *consists of all complex matrices satisfying* $X^* = -X$ *(that is,* X *is anti-Hermitian) and the Lie algebra of* $\mathsf{SU}(n)$ *consists of all complex matrices satisfying* $X^* = -X$ *and* $\mathrm{trace}(X) = 0$. *The Lie algebra of the orthogonal group* $\mathsf{O}(n)$ *consists of all real matrices* X *satisfying* $X^{tr} = -X$ *and the Lie algebra of* $\mathsf{SO}(n)$ *is the same as that of* $\mathsf{O}(n)$.

The Lie algebras of $\mathsf{U}(n)$ and $\mathsf{SU}(n)$ are denoted $\mathsf{u}(n)$ and $\mathsf{su}(n)$, respectively. The Lie algebra of $\mathsf{SO}(n)$ (which is the same as that of $\mathsf{O}(n)$) is denoted $\mathsf{so}(n)$.

Proof. A matrix U is unitary if and only if $U^* = U^{-1}$. Thus, e^{tX} is unitary if and only if

$$\left(e^{tX}\right)^* = \left(e^{tX}\right)^{-1} = e^{-tX}. \tag{3.4}$$

By Point 2 of Proposition 2.12, $\left(e^{tX}\right)^* = e^{tX^*}$, and so (3.4) becomes

$$e^{tX^*} = e^{-tX}. \tag{3.5}$$

The condition (3.5) holds for all real t if and only if $X^* = -X$ (for the implication to the right, apply $\frac{d}{dt}(\cdot)|_{t=0}$ to (3.5)). Thus, the Lie algebra of $\mathsf{U}(n)$ consists precisely of matrices X such that $X^* = -X$. The Lie algebra of $\mathsf{SU}(n) = \mathsf{U}(n) \cap \mathsf{SL}(n;\mathbb{C})$ equals $\mathsf{u}(n) \cap \mathsf{sl}(n;\mathbb{C})$, the family of anti-Hermitian matrices with trace zero. The Lie algebra of $\mathsf{O}(n) = \mathsf{U}(n) \cap \mathsf{GL}(n;\mathbb{R})$ equals $\mathsf{u}(n) \cap \mathsf{gl}(n;\mathbb{R})$, the family of real anti-symmetric matrices. Observe that a matrix $X \in \mathrm{Lie}(\mathsf{O}(n))$ has trace zero, for X is real and so $X^{\mathrm{tr}} = -X$ implies that the diagonal elements must be zero. Finally, the Lie algebra of $\mathsf{SO}(n) = \mathsf{O}(n) \cap \mathsf{SL}(n;\mathbb{R})$ equals $\mathrm{Lie}(\mathsf{O}(n)) \cap \mathsf{sl}(n;\mathbb{R}) = \mathrm{Lie}(\mathsf{O}(n))$, the set of real anti-symmetric matrices (which therefore must have trace zero). □

Lemma 3.17. *A linear function between Lie algebras preserves Lie brackets if and only if it preserves them over a generating set. Specifically, let $\mathfrak{g}$ and $\mathfrak{h}$ be real or complex Lie algebras, let $\{e_j\}_{j\in I} \subset \mathfrak{g}$ be such that $\mathrm{span}(\{e_j\}_{j\in J}) = \mathfrak{g}$ and suppose that $\phi : \mathfrak{g} \to \mathfrak{h}$ is linear. Then ϕ is a Lie algebra homomorphism if and only if*

$$\phi([e_j, e_k]) = [\phi(e_j), \phi(e_k)], \quad \text{for all } j, k \in I. \tag{3.6}$$

Furthermore, if ϕ is additionally a vector space isomorphism, then ϕ is a Lie algebra isomorphism if and only if (3.6) holds.

Proof. We begin by proving the first part of the lemma.

($\Rightarrow$) It is clear.

($\Leftarrow$) Let $u, v \in \mathfrak{g}$. There exists a finite set $J \subset I$ and scalars a_j, b_j, for $j \in J$, such that

$$u = \sum_{j\in J} a_j e_j, \quad v = \sum_{j\in J} b_j e_j.$$

By bilinearity of the bracket and linearity of ϕ, we have

$$\begin{aligned}
\phi\left([u,v]\right) &= \phi\left(\left[\sum_{j\in J} a_j e_j, \sum_{k\in J} b_k e_k\right]\right) \\
&= \phi\left(\sum_{j,k\in J} a_j b_k [e_j, e_k]\right) \\
&= \sum_{j,k\in J} a_j b_k \phi([e_j, e_k]) \\
&= \sum_{j,k\in J} a_j b_k [\phi(e_j), \phi(e_k)] \\
&= \left[\sum_{j\in J} a_j \phi(e_j), \sum_{k\in J} b_k \phi(e_k)\right] \\
&= \left[\phi\left(\sum_{j\in J} a_j e_j\right), \phi\left(\sum_{k\in J} b_k e_k\right)\right]
\end{aligned}$$

$$= [\phi(u), \phi(v)].$$

For the second part of the lemma, the implication ($\Rightarrow$) is clear again so we prove ($\Leftarrow$). If ϕ is a vector space isomorphism and (3.6) holds, then we have already proven that ϕ is also a Lie algebra homomorphism. But then, by definition, ϕ is also a Lie algebra isomorphism. □

Example 3.18. From Proposition 3.16 it follows that the following elements form a basis for the Lie algebra $\mathsf{su}(2)$:

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

These elements satisfy the commutation relations $[E_1, E_2] = E_3$, $[E_2, E_3] = E_1$, and $[E_3, E_1] = E_2$. In the same manner, from Proposition 3.16 it follows that the following elements form a basis for the Lie algebra $\mathsf{so}(3)$:

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

These elements satisfy the commutation relations $[F_1, F_2] = F_3$, $[F_2, F_3] = F_1$, and $[F_3, F_1] = F_2$.

Note that the listed relations completely determine all commutation relations among, say, E_1, E_2, and E_3, since by the anti-symmetry of the bracket, we must have $[E_1, E_1] = 0$, $[E_2, E_1] = -E_3$, and so on. Since E_1, E_2, and E_3 satisfy the same commutation relations as F_1, F_2, and F_3, by Lemma 3.17 the two Lie algebras are isomorphic.

3.4 Categories and Functors

The text of this section is taken from sections 1.1 and 1.3 of [Rieh]. We have added minor notational and terminological changes and selected only the parts relevant for this thesis.

A general strategy in Mathematics when one is dealing with a difficult object—such as, say, a Lie group—is to associate each instance of this object with a new "simpler" object, easier to study. Usually, the "simpler" object is cleverly chosen so as to retain some kind of information of the difficult object, in a way that through the study of the former one can obtain information of the latter.

In our case, the "simpler" object that we associate each Lie group with is its Lie algebra. After all, vector spaces are easier to study than groups.

The precise mathematical description of the association map $G \mapsto \operatorname{Lie} G$, as Corollary 3.31 asserts, is that of a *functor*. But before seeing that, we need to define the concept of a functor, of course. That is what we will do in this section.

A functor is a certain kind of map between *categories*.

Definition 3.19. A **category** consists of

- a class of **objects** $X, Y, Z, \dots$
- a class of **morphisms** $f, g, h, \dots$

so that:

- Each morphism has specified **domain** and **codomain** objects; the notation $f : X \to Y$ signifies that f is a morphism with domain X and codomain Y.
- Each object has a designated **identity morphism** $1_X : X \to X$.
- For any pair of morphisms f, g with the codomain of f equal to the domain of g, there exists a specified **composite morphism**[2] gf whose domain is equal to the domain of f and whose codomain is equal to the codomain of g, i. e.:

$$f : X \to Y, \quad g : Y \to Z \quad \rightsquigarrow \quad gf : X \to Z.$$

This data is subject to the following two axioms:

- For any $f : X \to Y$, the composites $1_Y f$ and $f 1_X$ are both equal to f.
- For any composable triple of morphisms f, g, h, the composites $h(gf)$ and $(hg)f$ are equal and henceforth denoted by hgf.

$$f : X \to Y, \quad g : Y \to Z, \quad h : Z \to W \quad \rightsquigarrow \quad hgf : X \to W.$$

That is, the composition law is associative and unital with the identity morphisms serving as two-sided identities.

It is traditional to name a category after its objects; typically, the preferred choice of accompanying structure-preserving morphisms is clear. Let's see it with examples.

Example 3.20. Many familiar varieties of mathematical objects assemble into a category.

(i) The category of sets, denoted Set, has the class of all sets as its class of objects and has the functions between sets as its morphisms.
(ii) Top has topological spaces as its objects and continuous functions as its morphisms.
(iii) Group has groups as objects and group homomorphisms as morphisms. The categories Ring of associative and unital rings and ring homomorphisms and Field of fields and fields homomorphisms are defined similarly.

[2] The composite may be written less concisely as $g \circ f$ when this adds typographical clarity.

(iv) LieGrp has Lie groups as objects and Lie group homomorphisms as morphisms. $\mathsf{MtxLieGrp}$ has matrix Lie groups as objects and Lie group homomorphisms between them as morphisms.

(v) If $\mathbb{K}$ is any field, $\mathsf{Vect}_{\mathbb{K}}$ is the category of $\mathbb{K}$-vector spaces and $\mathbb{K}$-vector space homomorphisms. We also denote $\mathsf{FinVect}_{\mathbb{K}}$ the category of finite-dimensional $\mathbb{K}$-vector spaces and $\mathbb{K}$-vector space homomorphisms between them.

(vi) If $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, $\mathsf{LieAlg}_{\mathbb{K}}$ is the category of Lie $\mathbb{K}$-algebras and Lie $\mathbb{K}$-algebra homomorphisms. We also denote $\mathsf{FinLieAlg}_{\mathbb{K}}$ the category of finite-dimensional Lie $\mathbb{K}$-algebras and Lie $\mathbb{K}$-algebra homomorphisms between them.

(vii) Man has differentiable manifolds as objects and differentiable functions as morphisms.

(viii) Categories Set_*, Top_*, and Man_* have sets (respectively, topological spaces and manifolds) with a specified basepoint as objects and base-point preserving (continuous, differentiable) functions as morphisms. That is, objects are pairs (X, x), where X is a set (resp., a topological space or a manifold) and $x \in X$; and morphisms which have domain (X, x) and codomain (Y, y) are (continuous or differentiable) functions $f : X \to Y$ such that $f(x) = y$. The category Set_* (resp., Top_*, Man_*) is called the category of pointed sets (resp., of pointed topological spaces and pointed manifolds).[3]

(ix) Meas has measurable spaces as objects and measurable functions as morphisms.

(x) Poset has partially-ordered sets as objects and order-preserving functions as morphisms.

The previous examples are all instances of so-called "concrete" categories. In these categories, each object of the category has an underlying set and morphisms are really functions between these sets.

However, "abstract" categories are also prevalent:

Example 3.21.

(i) A group G defines a category with a single object. The group elements are its morphisms, each group element representing a distinct endomorphism of the single object, with composition given by multiplication. The identity element $e \in G$ acts as the identity morphism for the unique object in this category. More generally, this construction also works when G is a monoid.

(ii) For a unital ring R, Mat_R is the category whose objects are positive integers and in which the set of morphisms from n to m is the set of $m \times n$ matrices with values in R. Composition is by matrix multiplication

$$n \xrightarrow{A} m, \quad m \xrightarrow{B} k \quad \rightsquigarrow \quad n \xrightarrow{B \cdot A} k$$

with identity matrices serving as the identity morphisms.

[3]In Spanish, "pointed set" is *conjunto punteado.*

(iii) A poset $(P, \leq)$ may be regarded as a category. The elements of P are the objects of the category and there exists a unique morphism $x \to y$ if and only if $x \leq y$. Transitivity of the relation "$\leq$" implies that the required composite morphisms exist. Reflexivity implies that identity morphisms exist. More generally, this same construction can be carried out if $\leq$ is a preorder on P; that is, a binary relation in P which is both reflexive and transitive (if $\leq$ is additionally antisymmetric then it is known as a partial order).

The previous examples illustrate the idea that morphisms in a category are not always functions. For that reason, morphisms are also called **arrows** or **maps**, particularly in the contexts of examples 3.21 and 3.20, respectively.

Nonetheless, in the rest of this thesis the only categories we will ever consider are just a few "concrete" categories: those of examples 3.20 (iv) and (v).

A **subcategory** D of a category C is defined by restricting to a subclass of objects and a subclass of morphisms subject to the requirements that the subcategory D contains the domain and codomain of any morphism in D, the identity morphism of any object in D, and the composite of any composable pair of morphisms in D. For example, there is a subcategory $\mathsf{CRing} \subset \mathsf{Ring}$ of commutative unital rings.

A category provides a context in which to answer the question "When is one thing the same as the other thing?" Almost universally in mathematics, one regards two objects of the same category to be "the same" when they are isomorphic, in a precise categorical sense that we now introduce

Definition 3.22. An **isomorphism** in a category is a morphism $f : X \to Y$ for which there exists a morphism $g : Y \to X$ so that $gf = 1_X$ and $fg = 1_Y$. The objects X and Y are **isomorphic** whenever there exists an isomorphism between X and Y, in which case one writes $X \cong Y$.

An **endomorphism**, i. e., a morphism whose domain equals its codomain, that is an isomorphism is called an **automorphism**.

Example 3.23.

(i) The isomorphisms in Set are precisely the **bijections**.
(ii) The isomorphisms in Group, Ring, Field, or $\mathsf{Vect}_{\mathbb{K}}$ are the bijective homomorphisms.
(iii) The isomorphisms in the category Top are the **homeomorphisms**, and the isomorphisms in Man are the **diffeomorphisms**.

In mathematics, after introducing a new object, the next thing which is usually introduced is the kind of map which allows to relate two instances of the same object. They are the structure-preserving maps of a specific mathematical structure. In abstract algebra, they are the homomorphisms, which relate two instances of the same algebraic structure, such as ring, group or module homomorphisms. In category theory, the map which relates two categories is a *functor*.

Definition 3.24. A **functor**[4] $F : \mathsf{C} \to \mathsf{D}$, between categories C and D, consists of the following data:

- An object $Fc \in \mathsf{D}$, for each object $c \in \mathsf{C}$.
- A morphism $Ff : Fc \to Fc' \in \mathsf{D}$, for each morphism $f : c \to c' \in \mathsf{C}$, so that the domain and codomain of Ff are, respectively, equal to F applied to the domain or codomain of f.

The assignments are required to satisfy the following two **functoriality axioms**:

- For any composable pair f, g in C, $Fg \circ Ff = F(g \circ f)$.
- For each object c in C, $F(1_c) = 1_{Fc}$.

Put concisely, a functor consists of a mapping on objects and a mapping on morphisms that preserves all of the structure of a category, namely domains and codomains, composition, and identities.

The previous definition is actually the definition for a **covariant** functor. In contrast, a **contravariant** functor $F : \mathsf{C} \to \mathsf{D}$ between categories C and D is a thing which satisfies Definition 3.24 except the first functoriality axiom. Instead, and by definition, a contravariant functor *reverses* compositions, so for any composable pair f, g in C, we have $F(g \circ f) = Ff \circ Fg$.

Example 3.25.

(i) There is a covariant endofunctor[5] $P : \mathsf{Set} \to \mathsf{Set}$ that sends a set A to its power set $PA = \{A' \subset A\}$ and a function $f : A \to B$ to the direct-image function $f_* : PA \to PB$ that sends $A' \subset A$ to $f(A') \subset B$.

(ii) There is a contravariant functor $\tilde{P} : \mathsf{Set} \to \mathsf{Set}$ that sends a set A to its power set PA and a function $f : A \to B$ to the inverse-image function $f^{-1} : PB \to PA$ that sends $B' \subset B$ to $f^{-1}(B') \subset A$.

(iii) Each of the categories listed in Example 3.20 has a **forgetful functor**,[6] a general term that is used for any functor that forgets structure, whose codomain is the category of sets. For example, $U : \mathsf{Group} \to \mathsf{Set}$ sends a group to its underlying set and a group homomorphism to its underlying function. The functor $U : \mathsf{Top} \to \mathsf{Set}$ sends a topological space to its set of points. These mappings are functorial because in each instance a morphism in the domain category has an underlying function. Other instance of a forgetful functor may be $U : \mathsf{LieGrp} \to \mathsf{Man}$, which sends a Lie group to its underlying manifold, forgetting the group

[4]In Spanish, the word is *funtor*.

[5]An **endofunctor** is a functor whose domain is equal to its codomain.

[6]In Spanish, the term is *funtor olvidadizo*.

structure in the process. We also could define the analogous forgetful functor $U : \mathsf{LieGrp} \to \mathsf{Group}$.

(iv) There is a functor $(-)^* : \mathsf{Vect}_{\mathbb{K}} \to \mathsf{Vect}_{\mathbb{K}}$ that carries a vector space to its **dual vector space** $V^* = \mathrm{Hom}(V, \mathbb{K})$. This functor is contravariant, with a linear map $\phi : V \to W$ sent to the linear map $\phi^* : W^* \to V^*$ that pre-composes a linear functional $W \xrightarrow{\omega} \mathbb{K}$ with ϕ to obtain a linear functional $V \xrightarrow{\phi} W \xrightarrow{\omega} \mathbb{K}$.

(v) In algebraic topology, the **fundamental group** defines a covariant functor $\pi_1 : \mathsf{Top}_* \to \mathsf{Group}$; a continuous function $f : (X, x) \to (Y, y)$ between pointed spaces induces a group homomorphism $f_* : \pi_1(X, x) \to \pi_1(Y, y)$ and this assignment is functorial: it satisfies the functoriality axioms from Definition 3.24.

(vi) The chain rule expresses the functoriality of the derivative. Let Euclid_* denote the category whose objects are pointed finite-dimensional Euclidean spaces $(\mathbb{R}^n, a)$—or, better, open subsets thereof—and whose morphisms are pointed differentiable functions. The **total derivative** of $f : \mathbb{R}^n \to \mathbb{R}^m$, evaluated at the designated basepoint $a \in \mathbb{R}^n$, gives rise to a the **Jacobian matrix** defining the directional derivatives of f at the point a. If f is given by component functions $f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$, the (i, j)-entry of this matrix is $\frac{\partial}{\partial x_j} f_i(a)$. This defines the action on morphisms of a covariant functor $D : \mathsf{Euclid}_* \to \mathsf{Mat}_{\mathbb{R}}$; on objects, D assigns a pointed Euclidean space its dimension. Given $g : \mathbb{R}^m \to \mathbb{R}^k$ carrying the designated basepoint $f(a) \in \mathbb{R}^m$ to $gf(a) \in \mathbb{R}^k$, the functoriality of D asserts that the product of the Jacobian of f at a with the Jacobian of g at $f(a)$ equals the Jacobian of gf at a. This is the chain rule from multivariable calculus.

(vii) A more sophisticated version of the previous Example 3.25 (vi) comes from differential geometry. For each differentiable map $f : M \to N$ between manifolds M and N, we have the derivative $df_a : T_a M \to T_{f(a)} N$ at a point $a \in M$ between the tangent spaces of M at a and of N at $f(a)$. This defines the action on morphisms of a covariant functor $d : \mathsf{Man}_* \to \mathsf{FinVect}_{\mathbb{R}}$, which goes from the category of pointed differentiable manifolds into the category of real finite dimensional vector spaces; on objects, d assigns the pointed manifold (M, a) its tangent space at a. Functoriality here is then precisely the chain rule in differential geometry, $d(g \circ f)_a = dg_{f(a)} \circ df_a$. If instead of fixing our attention to particular points on the manifolds, we consider the differential defined between the tangent bundles $df : TM \to TN$, we get a covariant endofunctor $T : \mathsf{Man} \to \mathsf{Man}$. It sends the manifold M to its tangent bundle TM and it sends the differentiable function $f : M \to N$ to the differentiable function df.

(viii) A **groupoid** is a category in which every morphism is an isomorphism. By Example 3.21 (i), we can regard every group G as a category with one single object, whose morphisms corresponds to the group elements. In this category, every morphism is invertible, i.e., every morphism is an isomorphism. For that reason, we can proceed backwards and *define* a group to be a groupoid with a one single object. This begs the question: what is a functor between groups? If G and H are groups (i.e., groupoids with one single object) and $f : G \to H$ is a functor, then $f(gg') = f(g)f(g')$ for every two morphisms $g, g' \in G$. That is, f corresponds to a *group homomorphism*.

The following result is arguably the first lemma in category theory.

Lemma 3.26. *Functors preserve isomorphisms.*

Proof. Consider a covariant functor $F : \mathsf{C} \longrightarrow \mathsf{D}$ and an isomorphism $f : x \longrightarrow y$ in C with inverse $g : y \longrightarrow x$. Applying the two functoriality axioms:

$$F(g)F(f) = F(gf) = F(1_x) = 1_{Fx}.$$

Thus, $Fg : Fy \longrightarrow Fx$ is a left inverse to $Ff : Fx \longrightarrow Fy$. Exchanging the roles of f and g shows that Fg is also a right inverse. The proof when F is contravariant is analogous. □

Corollary 3.27. *Functors preserve isomorphic objects: if $F : \mathsf{C} \longrightarrow \mathsf{D}$ is a functor and $x \cong y$ in C, then $Fx \cong Fy$ in D.*

Examples of applications of Corollary 3.27 are:

(i) Any two path-connected homeomorphic topological spaces must have isomorphic fundamental groups (by Example 3.25 (v)).
(ii) Any two diffeomorphic manifolds must have same dimension, for their tangent spaces must be isomorphic (by Example 3.25 (vii)); as well as diffeomorphic tangent bundles.

Another remarkable property of functors is that they transform commutative diagrams into commutative diagrams, as a commutative diagram is just a graphical representation of an equality between compositions of morphisms and functors preserves composition of morphisms.

Converse of Lemma 3.26 is false. If D is a category with one single object x and one single morphism—namely, 1_x—then there is a functor F from any category C into D which sends every object of C to x and every morphism of C to 1_x. If C has a morphism f which is not an isomorphism, then F would send the non-isomorphism f to the isomorphism 1_x. This can also provide a counterexample to the converse of Corollary 3.27.

Definition 3.28. A functor $F : \mathsf{C} \longrightarrow \mathsf{D}$ is said to **reflect isomorphisms** if whenever f is a morphism of C such that Ff is an isomorphism in D then f is an isomorphism. The functor F is said to **create isomorphisms** if whenever x and y are objects in C such that $Fx \cong Fy$ then $x \cong y$. Lastly, the functor F is said to **reflect and create isomorphisms** if it both reflects isomorphisms and creates isomorphisms.

In other words, F reflects isomorphisms if for all morphisms f of C, f is an isomorphism if and only if Ff is an isomorphism; and F creates isomorphisms if for all objects x and y of C, $x \cong y$ if and only if $Fx \cong Fy$.

A functor that creates isomorphisms is one that is “injective on objects up to isomorphism.” Functors that reflect isomorphisms are sometimes also called **conservative functors** in the literature.

3.5 The Lie Functor

The following theorem tells us that a Lie group homomorphism between two Lie groups gives rise in a natural way to a map between the corresponding Lie algebras.

Theorem 3.29. *Let G and H be matrix Lie groups, with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Suppose that $\Phi : G \to H$ is a Lie group homomorphism. Then there exists a unique real-linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ such that*

$$\Phi(e^X) = e^{\phi(X)} \tag{3.7}$$

for all $X \in \mathfrak{g}$. Equivalently, such that the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\phi"] \arrow[d, "\exp"'] & \mathfrak{h} \arrow[d, "\exp"] \\ G \arrow[r, "\Phi"] & H \end{tikzcd}$$

commutes. The map ϕ has the following additional properties:

1. $\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A)^{-1}$, *for all $X \in \mathfrak{g}$, $A \in G$.*
2. $\phi([X, Y]) = [\phi(X), \phi(Y)]$, *for all $X, Y \in \mathfrak{g}$.*
3. $\phi(X) = \frac{d}{dt}\Phi(e^{tX})|_{t=0}$, *for all $X \in \mathfrak{g}$.*

In practice, given a Lie group homomorphism Φ, the way one goes about computing ϕ is by using Property 3. In the language of manifolds, Property 3 says that ϕ is the derivative (or differential) of Φ at the identity. By point 2, $\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. Thus, every Lie group homomorphism gives rise to a Lie algebra homomorphism. In Chapter 5, we will investigate the reverse question: If ϕ is a homomorphism between the Lie algebras of two Lie groups, is there an associated Lie group homomorphism Φ?

Proof. The proof is similar to the proof of Theorem 3.12. Since Φ is a continuous group homomorphism, $\Phi(e^{tX})$ will be a one-parameter subgroup of H, for each $X \in \mathfrak{g}$. Thus, by Theorem 2.23, there is a unique matrix Z such that

$$\Phi(e^{tX}) = e^{tZ} \tag{3.8}$$

for all $t \in \mathbb{R}$. We define $\phi(X) = Z$ and check that ϕ has the required properties. First, by putting $t = 1$ in (3.8), we see that $\Phi(e^X) = e^{\phi(X)}$ for all $X \in \mathfrak{g}$. Next, if $\Phi(e^{tX}) = e^{tZ}$

for all t, then $\Phi(e^{tsX}) = e^{tsZ}$, showing that $\phi(sX) = s\phi(X)$. Using the Lie product formula (Theorem 2.20) and the continuity of Φ, we then compute that

$$\begin{aligned} e^{t\phi(X+Y)} &= \Phi\left(e^{t(X+Y)}\right) \\ &= \Phi\left(\lim_{m\to\infty}\left(e^{tX/m}e^{tY/m}\right)^m\right) \\ &= \lim_{m\to\infty}\Phi\left(\left(e^{tX/m}e^{tY/m}\right)^m\right) \\ &= \lim_{m\to\infty}\left(\Phi\left(e^{tX/m}\right)\Phi\left(e^{tY/m}\right)\right)^m \\ &= \lim_{m\to\infty}\left(e^{t\phi(X)/m}e^{t\phi(Y)/m}\right)^m \\ &= e^{t(\phi(X)+\phi(Y))}. \end{aligned}$$

Differentiating this result at $t = 0$ shows that $\phi(X + Y) = \phi(X) + \phi(Y)$.

We have thus obtained a real-linear map ϕ satisfying (3.7). If there were another real-linear map ϕ' with this property, we would have

$$e^{t\phi(X)} = \Phi(e^{tX}) = e^{t\phi'(X)}$$

for all $t \in \mathbb{R}$. Differentiating this result at $t = 0$ shows that $\phi(X) = \phi'(X)$.

We now verify the remaining claimed properties of ϕ. For any $A \in G$, we have

$$\begin{aligned} e^{t\phi(AXA^{-1})} &= e^{\phi(tAXA^{-1})} \\ &= \phi(e^{tAXA^{-1}}) \\ &= \Phi(Ae^{tX}A^{-1}) \\ &= \Phi(A)\Phi(e^{tX})\Phi(A^{-1}) \\ &= \Phi(A)e^{t\phi(X)}\Phi(A)^{-1}. \end{aligned}$$

Differentiating the identity $e^{t\phi(AXA^{-1})} = \Phi(A)e^{t\phi(X)}\Phi(A)^{-1}$ at $t = 0$ gives Point 1.

Meanwhile, for any X and Y in $\mathfrak{g}$, we have, as in the proof of Theorem 3.12,

$$\begin{aligned} \phi([X,Y]) &= \phi\left(\left.\frac{d}{dt}e^{tX}Ye^{-tX}\right|_{t=0}\right) \\ &= \left.\frac{d}{dt}\phi\left(e^{tX}Ye^{-tX}\right)\right|_{t=0}, \end{aligned}$$

where we have used the fact that a derivative commutes with a linear transformation.[7] Thus,

$$\phi([X,Y]) = \left.\frac{d}{dt}\Phi(e^{tX})\phi(Y)\Phi(e^{-tX})\right|_{t=0}$$

---

[7]That is to say, if $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x_0 \in U$, where U is open, and $\psi : \mathbb{R}^m \to \mathbb{R}^k$ is a linear function, then, by the chain rule $D(\psi f)(x_0) = D\psi(f(x_0))Df(x_0) = \psi Df(x_0)$, for linear functions are differentiable and coincide with its own differential. If $n = 1$ then the previous identity says that $\frac{d}{dx}\psi(f(x))\big|_{x=x_0} = \psi\left(\frac{d}{dx}f(x)\big|_{x=x_0}\right)$.

$$= \frac{d}{dt} e^{t\phi(X)} \phi(Y) e^{-t\phi(X)} \bigg|_{t=0}$$
$$= [\phi(X), \phi(Y)],$$

establishing Point 2. Finally, since $\Phi(e^{tX}) = e^{\phi(tX)} = e^{t\phi(X)}$, we can compute $\phi(X)$ as in Point 3. □

Example 3.30. let $\Phi : \mathsf{SU}(2) \longrightarrow \mathsf{SO}(3)$ be the homomorphism in Proposition 1.11. Then the associated Lie algebra homomorphism $\phi : \mathsf{su}(2) \longrightarrow \mathsf{so}(3)$ satisfies

$$\phi(E_j) = F_j, \quad j = 1, 2, 3,$$

where $\{E_1, E_2, E_3\}$ and $\{F_1, F_2, F_3\}$ are the bases of $\mathsf{su}(2)$ and $\mathsf{so}(3)$, respectively, given in Example 3.18.

Since ϕ maps a basis for $\mathsf{su}(2)$ to a basis for $\mathsf{so}(3)$, by Lemma 3.17 we see that ϕ is a Lie algebra isomorphism, even though Φ is not a Lie group isomorphism (since $\ker \Phi = \{I, -I\}$).

Proof. If X is in $\mathsf{su}(2)$ and Y is in the space V in, then

$$\frac{d}{dt} \Phi(e^{tX}) Y \bigg|_{t=0} = \frac{d}{dt} e^{tX} Y e^{-tX} \bigg|_{t=0} = [X, Y].$$

Thus, $\phi(X)$ is the linear map of $V \cong \mathbb{R}^3$ to itself given by $Y \mapsto [X, Y]$. If, say, $X = E_1$, then direct computation shows that

$$\left[E_1, \begin{pmatrix} x_1 & x_2 + ix_3 \\ x_2 - ix_3 & -x_1 \end{pmatrix} \right] = \begin{pmatrix} x_1' & x_2' + ix_3' \\ x_2' - ix_3' & -x_1' \end{pmatrix},$$

where $(x_1', x_2', x_3') = (0, -x_3, x_2)$. Since

$$\begin{pmatrix} 0 \\ -x_3 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{3.9}$$

we conclude that $\phi(E_1)$ is the 3×3 matrix appearing on the right-hand side of (3.9), which is precisely F_1. The computation of $\phi(E_2)$ and $\phi(E_3)$ is similar and is left to the reader. □

Corollary 3.31. *There is a covariant functor, called the* ***Lie functor****,*

$$\mathrm{Lie} : \mathsf{MtxLieGrp} \longrightarrow \mathsf{FinLieAlg}_{\mathbb{R}}$$

from the category of matrix Lie groups and Lie group homomorphisms to the category of finite-dimensional real Lie algebras and Lie algebra homomorphisms. The functor sends each matrix Lie group G to its Lie algebra $\mathfrak{g}$, and it sends each Lie group homomorphism $\Phi : G \longrightarrow H$ between matrix Lie groups to the induced Lie algebra homomorphism $\phi : \mathfrak{g} \longrightarrow \mathfrak{h}$ from Theorem 3.29.

Proof. The only thing we need to check is the functoriality axioms. Firstly, since $\mathrm{id}_G(e^X) = e^X = e^{\mathrm{id}_{\mathfrak{g}}(X)}$, it is clear that $\mathrm{Lie}(\mathrm{id}_G) = \mathrm{id}_{\mathfrak{g}}$ for each matrix Lie group G. Let $\Phi : H \longrightarrow K$, $\Psi : G \longrightarrow H$ be Lie group homomorphisms between the matrix Lie groups G, H, K. Write $\Lambda = \Phi \circ \Psi$ and denote ϕ, ψ, λ the Lie algebra homomorphisms induced by Φ, Ψ, Λ, respectively. For any $X \in \mathfrak{g}$,

$$\Lambda(e^{tX}) = \Phi(\Psi(e^{tX})) = \Phi(e^{t\psi(X)}) = e^{t\phi(\psi(X))}.$$

Hence, $\lambda(X) = \phi(\psi(X))$. The functor is thus covariant. □

Corollary 3.32. *Any two isomorphic matrix Lie groups have isomorphic Lie algebras.*

Proof. Application of Lemma 3.26. □

Example 3.18 gives a counterexample for the converse: $\mathsf{su}(2)$ and $\mathsf{so}(3)$ are isomorphic although $\mathsf{SU}(2)$ and $\mathsf{SO}(3)$ are not, as $\mathsf{SU}(2)$ is simply connected and $\mathsf{SO}(3)$ is not (see respectively the comments following proof of Proposition 1.9 and the last paragraph after the proof of Corollary 1.12). In other words, the Lie functor doesn't create isomorphisms. This functor neither reflects isomorphisms: In Example 3.30, even though $\mathrm{Lie}\,\Phi = \phi : \mathsf{su}(2) \longrightarrow \mathsf{so}(3)$ is an isomorphism, $\Phi : \mathsf{SU}(2) \longrightarrow \mathsf{SO}(3)$ is not, for $\ker \Phi = \{I, -I\}$.

However, in Sect. 5.2 we will see that we can restrict the Lie functor to a subcategory of MtxLieGrp so that the restricted functor will reflect and create isomorphisms. Namely, the the subcategory of simply connected matrix Lie groups.

Proposition 3.33. *If $\Phi : G \longrightarrow H$ is a Lie group homomorphism and $\mathrm{Lie}\,\Phi = \phi : \mathfrak{g} \longrightarrow \mathfrak{h}$ is the associated Lie algebra homomorphism, then the kernel of Φ is a closed, normal subgroup of G and the Lie algebra of the kernel is given by*

$$\mathrm{Lie}(\ker \Phi) = \ker(\mathrm{Lie}\,\Phi).$$

Proof. Since Φ is continuous, $\ker \Phi$ is closed, so it is a matrix Lie group. If $X \in \ker \phi$, then

$$\Phi(e^{tX}) = e^{t\phi(X)} = I,$$

for all $t \in \mathbb{R}$, showing that X is in the Lie algebra of $\ker \Phi$. In the other direction, if e^{tX} lies in $\ker \Phi$ for all $t \in \mathbb{R}$, then

$$e^{t\phi(X)} = \Phi(e^{tX}) = I$$

for all t. Differentiating this relation with respect to t at $t = 0$ gives that $\phi(X) = 0$, showing that $X \in \ker \phi$. □

Definition 3.34 (The Adjoint map). Let G be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Then for each $A \in G$, define a linear map $\mathrm{Ad}_A : \mathfrak{g} \longrightarrow \mathfrak{g}$ by the formula

$$\mathrm{Ad}_A(X) = AXA^{-1}.$$

Proposition 3.35. *Let G be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Let $\mathsf{GL}(\mathfrak{g})$ denote the group of all invertible linear transformations of $\mathfrak{g}$. Then the map $A \mapsto \mathrm{Ad}_A$ is a homomorphism of G into $\mathsf{GL}(\mathfrak{g})$. Furthermore, for each $A \in G$, Ad_A satisfies $\mathrm{Ad}_A([X,Y]) = [\mathrm{Ad}_A(X), \mathrm{Ad}_A(Y)]$ for all $X, Y \in \mathfrak{g}$.*

Proof. Easy. Note that Point of Theorem 3.12 guarantees that $\mathrm{Ad}_A(X)$ is actually in $\mathfrak{g}$ for all $X \in \mathfrak{g}$. □

That is, the map $A \mapsto \mathrm{Ad}_A$ is really a homomorphism of G into the group of Lie algebra automorphisms of $\mathfrak{g}$. By $\mathsf{GL}(\mathfrak{g})$ we denote the vector space automorphisms of $\mathfrak{g}$, which do not necessarily preserve the bracket of $\mathfrak{g}$.

Since $\mathfrak{g}$ is a real vector space with some dimension k, $\mathsf{GL}(\mathfrak{g})$ is essentially the same as $\mathsf{GL}(k;\mathbb{R})$). Thus, we will regard $\mathsf{GL}(\mathfrak{g})$ as a matrix Lie group.[8] The map $\mathrm{Ad} : G \to \mathsf{GL}(\mathfrak{g})$ is continuous. Let $\{X_1, \dots, X_k\}$ be a basis of $\mathfrak{g}$ and let $f : \mathfrak{g} \to \mathbb{R}^k$ be the linear isomorphism given by $f(X_i) = e_i$, where e_i is the i-th vector of the canonical basis of $\mathbb{R}^k$. Then, for each $A \in G$ the matrix of $\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$ with respect to this basis is

$$(f\,(\mathrm{Ad}_A(X_1)) \mid \cdots \mid f(\mathrm{Ad}_A(X_1))) = \left(f(AX_1A^{-1}) \mid \cdots \mid f(AX_kA^{-1})\right).$$

This matrix depends continuously on A, for it is component-wise continuous in A. Thus, Ad is a Lie group homomorphism.

The Lie functor sends $\mathrm{Ad} : G \to \mathsf{GL}(\mathfrak{g})$ to the Lie algebra homomorphism

$$\begin{aligned} \mathrm{ad} : \mathfrak{g} &\to \mathfrak{gl}(\mathfrak{g}) \\ X &\mapsto \mathrm{ad}_X, \end{aligned}$$

with the property that

$$e^{\mathrm{ad}_X} = \mathrm{Ad}_{e^X}\,.$$

Equivalently, the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\mathrm{ad}"] \arrow[d, "\exp"'] & \mathfrak{gl}(\mathfrak{g}) \arrow[d, "\exp"] \\ G \arrow[r, "\mathrm{Ad}"] & \mathsf{GL}(\mathfrak{g}) \end{tikzcd}$$

commutes. Here, $\mathfrak{gl}(\mathfrak{g}) = \mathrm{Lie}(\mathsf{GL}(\mathfrak{g}))$ is the space of all linear maps from $\mathfrak{g}$ to itself, with bracket equal to the commutator of operators.[9]

[8]Choosing a basis in $\mathfrak{g}$ allows to identify $\mathsf{GL}(\mathfrak{g})$ with $\mathsf{GL}(k;\mathbb{R})$. The set $\mathsf{GL}(\mathfrak{g})$ can be endowed with a canonical differentiable structure that makes any such bijection $\mathsf{GL}(\mathfrak{g}) \to \mathsf{GL}(k;\mathbb{R})$ become a diffeomorphism. See comments before Definition 4.2 regarding the differential structure in $\mathsf{GL}(V)$, with V any finite-dimensional vector space.

[9]Similarly as before, the space $\mathfrak{gl}(\mathfrak{g})$ can be endowed with a canonical differentiable structure that makes any linear isomorphism $\mathsf{GL}(\mathfrak{g}) \cong \mathsf{GL}(k;\mathbb{R})$ become a diffeomorphism. As before, see comments before Definition 4.2 as well.

Proposition 3.36. *Let G be a matrix Lie group, let $\mathfrak{g}$ be its Lie algebra, and let* $\mathrm{Ad} : G \longrightarrow \mathsf{GL}(\mathfrak{g})$ *be as in Proposition 3.35. Let* $\mathrm{ad} = \mathrm{Lie}\,\mathrm{Ad} : \mathfrak{g} \longrightarrow \mathsf{gl}(\mathfrak{g})$ *be the associated Lie algebra map. Then for all $X, Y \in \mathfrak{g}$*

$$\mathrm{ad}_X(Y) = [X, Y]. \tag{3.10}$$

The proposition shows that our usage of the notation ad_X in this section is consistent with that in Definition 3.7.

Proof. By Point 3 of Theorem 3.29, ad can be computed as follows:

$$\mathrm{ad}_X = \frac{d}{dt}\,\mathrm{Ad}_{e^{tX}}\bigg|_{t=0}.$$

If V is a finite-dimensional real vector space of dimension n, the space $\mathrm{End}(V)$ of linear endomorphisms in V can be identified with $\mathbb{R}^{n^2}$. The application map

$$\begin{aligned} \mathrm{app} : \mathrm{End}(V) \times V &\longrightarrow V \\ (A, v) &\longmapsto Av \end{aligned}$$

is continuous and thus, if $A : (-\varepsilon, 0) \cup (0, \varepsilon) \longrightarrow \mathrm{End}\,V$ is such that $\lim_{t\to 0} A(t)$ exists, we have

$$\mathrm{app}\left(\lim_{t\to 0} A(t), v\right) = \lim_{t\to 0} \mathrm{app}(A(t), v).$$

Hence,

$$\begin{aligned} \mathrm{ad}_X(Y) &= \frac{d}{dt}\,\mathrm{Ad}_{e^{tX}}\bigg|_{t=0}(Y) \\ &= \frac{d}{dt}\,\mathrm{Ad}_{e^{tX}}(Y)\bigg|_{t=0} \\ &= \frac{d}{dt} e^{tX} Y e^{-tX}\bigg|_{t=0} \\ &= [X, Y], \end{aligned}$$

as claimed. □

We have proved, as a consequence of Theorem 3.29 and Proposition 3.36, the following result.

Proposition 3.37. *For any X in $M_n(\mathbb{C})$, let* $\mathrm{ad}_X : M_n(\mathbb{C}) \longrightarrow M_n(\mathbb{C})$ *be given by* $\mathrm{ad}_X\, Y = [X, Y]$. *Then for any Y in $M_n(\mathbb{C})$, we have*

$$e^X Y e^{-X} = \mathrm{Ad}_{e^X}(Y) = e^{ad_X}(Y),$$

where

$$e^{ad_X}(Y) = Y + [X, Y] + \frac{1}{2}[X, [X, Y]] + \cdots.$$

3.6 The Complexification of a Real Lie Algebra

In studying the representations of a matrix Lie group G (as we will do in later chapters), it is often useful to pass to the Lie algebra $\mathfrak{g}$ of G, which is, in general, only a real Lie algebra. It is often useful to pass to an associated complex Lie algebra, called the *complexification*[10] of $\mathfrak{g}$.

Definition 3.38. If V is a finite-dimensional real vector space, then the **complexification** of V, denoted $V_{\mathbb{C}}$, is the space of formal linear combinations

$$v_1 + iv_2,$$

with $v_1, v_2 \in V$. This becomes a real vector space in the obvious way and becomes a complex vector space if we define

$$i(v_1 + iv_2) = -v_2 + iv_1.$$

We could more pedantically define $V_{\mathbb{C}}$ to be the space of ordered pairs (v_1, v_2) with $v_1, v_2 \in V$, but this is notationally cumbersome. It is straightforward to verify that the above definition really makes $V_{\mathbb{C}}$ into a complex vector space. We will regard V as a real subspace of $V_{\mathbb{C}}$ in the obvious way.

Proposition 3.39. *Let $\mathfrak{g}$ be a finite-dimensional real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ its complexification. Then the bracket operation on $\mathfrak{g}$ has a unique extension to $\mathfrak{g}_{\mathbb{C}}$ that makes $\mathfrak{g}_{\mathbb{C}}$ into a complex Lie algebra. The complex Lie algebra $\mathfrak{g}_{\mathbb{C}}$ is called the* ***complexification*** *of the real Lie algebra $\mathfrak{g}$.*

Proof. The uniqueness of the extension is obvious, since if the bracket operation on $\mathfrak{g}_{\mathbb{C}}$ is to be bilinear, then it must be given by

$$[X_1 + iX_2, Y_1 + iY_2] = ([X_1, Y_1] - [X_2, Y_2]) + i\left([X_1, Y_2] + [X_2, Y_1]\right). \tag{3.11}$$

To show existence, we must now check that (3.11) is really bilinear and antisymmetric and that it satisfies the Jacobi identity. It is clear that (3.11) is *real* bilinear, and antisymmetric. The antisymmetry means that if (3.11) is complex linear in the first factor, it is complex linear in the second factor. Thus, we need only show that

$$[i(X_1 + iX_2), Y_1 + iY_2] = i[X_1 + iX_2, Y_1 + iY_2]. \tag{3.12}$$

The left-hand side of (3.12) is

$$[-X_2 + iX_1, Y_1 + iY_2] = (-[X_2, Y_1] - [X_1, Y_2]) + i([X_1, Y_1] - [X_2, Y_2]),$$

whereas the right-hand side of (3.12) is

$$\begin{aligned} &i\{([X_1, Y_1] - [X_2, Y_2]) + i([X_2, Y_1] + [X_1, Y_2])\} \\ &\quad = (-[X_2, Y_1] - [X_1, Y_2]) + i([X_1, Y_1] - [X_2, Y_2]) \end{aligned}$$

[10] In Spanish, *complejificación*.

and, indeed, these expressions are equal. It remains to check the Jacobi identity. Of course, the Jacobi identity holds if X, Y, and Z are in $\mathfrak{g}$. Furthermore, for all $X, Y, Z \in \mathfrak{g}_{\mathbb{C}}$, the expression

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]]$$

is complex-linear in X with Y and Z fixed. Thus, the Jacobi identity continues to hold if X is in $\mathfrak{g}_{\mathbb{C}}$ and Y and Z are in $\mathfrak{g}$. The same argument then shows that the Jacobi identity holds when X and Y are in $\mathfrak{g}_{\mathbb{C}}$ and Z is in $\mathfrak{g}$. Applying this argument one more time establishes the Jacobi identity for $\mathfrak{g}_{\mathbb{C}}$ in general. □

Proposition 3.40. *Suppose that $\mathfrak{g} \subset M_n(\mathbb{C})$ is a real Lie algebra and that for all nonzero X in $\mathfrak{g}$, the element iX is* not *in $\mathfrak{g}$. Then the "abstract" complexification $\mathfrak{g}_{\mathbb{C}}$ of $\mathfrak{g}$ from Proposition 3.39 is isomorphic to the complex Lie subalgebra of $M_n(\mathbb{C})$ of matrices that can be expressed in the form $X + iY$ with X and Y in $\mathfrak{g}$.*

Proof. The complex subspace of $M_n(\mathbb{C})$ of matrices of the form $X+iY$, with $X, Y \in \mathfrak{g}$, is indeed a subalgebra of $M_n(\mathbb{C})$ since expression (3.11) is also valid when $X_1, X_2, Y_1, Y_2 \in \mathfrak{g} \subset M_n(\mathbb{C})$.

Consider now the map from $\mathfrak{g}_{\mathbb{C}}$ into $M_n(\mathbb{C})$ sending the formal linear combination $X + iY$ to the linear combination $X + iY$ of matrices. This map is easily seen to be a complex Lie algebra homomorphism. If $\mathfrak{g}$ satisfies the assumption in the statement of the proposition, this map is also injective and thus an isomorphism of $\mathfrak{g}_{\mathbb{C}}$ with $\mathfrak{g} + i\mathfrak{g} \subset M_n(\mathbb{C})$. □

Using the proposition, we easily obtain the following list of isomorphisms:

$$\begin{aligned}
\mathsf{gl}(n;\mathbb{R})_{\mathbb{C}} &\cong \mathsf{gl}(n;\mathbb{C}),\\
\mathsf{u}(n)_{\mathbb{C}} &\cong \mathsf{gl}(n;\mathbb{C}),\\
\mathsf{su}(n)_{\mathbb{C}} &\cong \mathsf{sl}(n;\mathbb{C}),\\
\mathsf{sl}(n;\mathbb{R})_{\mathbb{C}} &\cong \mathsf{sl}(n;\mathbb{C}),\\
\mathsf{so}(n)_{\mathbb{C}} &\cong \mathsf{so}(n;\mathbb{C}).
\end{aligned}$$

Let us now verify two examples, those of the complexifications of $\mathsf{u}(n)$ and $\mathsf{su}(n)$. For the first one, if $X^* = -X$, then $(iX)^* = iX$. Thus, X and X^* cannot be both in $\mathsf{u}(n)$ unless X is zero. Furthermore, every X in $M_n(\mathbb{C})$ can be expressed as $X = X_1 + iX_2$, where $X_1 = (X - X^*)/2$ and $X_2 = (X + X^*)/(2i)$ are both in $\mathsf{u}(n)$. This shows that $\mathsf{u}(n)_{\mathbb{C}} \cong \mathsf{gl}(n;\mathbb{C})$. The analogous argument also shows that every matrix of $\mathsf{sl}(n;\mathbb{C})$ can be written in the form X_1+iX_2, with $X_1, X_2 \in \mathsf{su}(n)$. Conversely, if $X_1, X_2 \in \mathsf{su}(n)$, then X_1+iX_2 has trace zero and is thus in $\mathsf{sl}(n;\mathbb{C})$. This shows that $\mathsf{su}(n)_{\mathbb{C}} \cong \mathsf{sl}(n;\mathbb{C})$.

Although both $\mathsf{su}(2)_{\mathbb{C}}$ and $\mathsf{sl}(2;\mathbb{R})_{\mathbb{C}}$ are isomorphic to $\mathsf{sl}(2;\mathbb{C})$, the Lie algebra $\mathsf{su}(2)$ is *not* isomorphic to $\mathsf{sl}(2;\mathbb{R})$. This is because $\mathsf{sl}(2;\mathbb{R})$ has two-dimensional subalgebras, whereas $\mathsf{su}(2)$ has not. Certainly, the matrices

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are both in $\mathfrak{sl}(2;\mathbb{R})$, and they generate a two-dimensional subalgebra, as $[H, X] = 2X$. On the other hand, if we consider the basis $\{E_1, E_2, E_3\}$ of $\mathfrak{su}(2)$ given in Example 3.18, then the linear isomorphism $\psi : \mathfrak{su}(2) \longrightarrow \mathbb{R}^3$ given by $\psi(E_j) = e_j$, where e_j is the j-th vector from the canonical basis of $\mathbb{R}^3$, is also a Lie algebra isomorphism $\mathfrak{su}(2) \cong (\mathbb{R}^3, \times)$ by Lemma 3.17. Here $(\mathbb{R}^3, \times)$ is the Lie algebra of Example 3.2. And indeed in $(\mathbb{R}^3, \times)$ there are not two-dimensional subalgebras, since by the right-hand rule any two linearly independent vectors of $\mathbb{R}^3$ must have a perpendicular cross product, this way being the whole $\mathbb{R}^3$ the Lie subalgebra that they two generate.

Proposition 3.41 (Universal property of the complexification of a Lie algebra). *Let $\mathfrak{g}$ be a real Lie algebra, $\mathfrak{g}_{\mathbb{C}}$ it complexification, and $\mathfrak{h}$ an arbitrary complex Lie algebra. Then every real Lie algebra homomorphism of $\mathfrak{g}$ into $\mathfrak{h}$ extends uniquely to a complex Lie algebra homomorphism of $\mathfrak{g}_{\mathbb{C}}$ into $\mathfrak{h}$.*

$$\begin{tikzcd} \mathfrak{g} \arrow[r] \arrow[d, hook] & \mathfrak{h} \\ \mathfrak{g}_{\mathbb{C}} \arrow[ur, dashed] & \end{tikzcd}$$

Proof. The unique extension is given by $\pi(X+iY) = \pi(X)+i\pi(Y)$ for all $X, Y \in \mathfrak{g}$. It is easy to check that this map is, indeed, a homomorphism of complex Lie algebras. □

3.7 The Exponential Map

Definition 3.42. If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, then the **exponential map** for G is the map

$$\exp : \mathfrak{g} \longrightarrow G.$$

That is to say, the exponential map for G is the matrix exponential restricted to the Lie algebra $\mathfrak{g}$ of G. It can be shown that every matrix in $\mathsf{GL}(n;\mathbb{C})$ is the exponential of some $n \times n$ matrix. Nevertheless, if $G \subset \mathsf{GL}(n;\mathbb{C})$ is a closed subgroup, there may exist A in G such that there is no X *in the Lie algebra* $\mathfrak{g}$ *of* G with $\exp X = A$.

Example 3.43. There does not exist a matrix $X \in \mathfrak{sl}(2;\mathbb{C})$ with

$$e^X = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}, \tag{3.13}$$

even though the matrix on the right-hand side of (3.13) is in $\mathsf{SL}(2;\mathbb{C})$.

Proof. If $X \in \mathfrak{sl}(2;\mathbb{C})$ has distinct eigenvalues, then X is diagonalizable and e^X will also be diagonalizable (by Point 6 of Proposition 2.12), unlike matrix on the right-hand side of (3.13). If $X \in \mathfrak{sl}(2;\mathbb{C})$ has a repeated eigenvalue, this eigenvalue must be 0 or the trace of X would not be zero. Thus, there is a nonzero vector v with $Xv = 0$, from which it follows that

$$e^X v = \left(\sum_{m=0}^{\infty} \frac{1}{m!} X^m\right) v = \sum_{m=0}^{\infty} \frac{1}{m!} X^m v = v + 0 + 0 + \cdots = v,$$

where we have used the continuity of the application map, as in proof of Proposition 3.36. We conclude that e^X has 1 as eigenvalue, unlike the matrix on the right-hand side of (3.13). $\square$

We see, then, that the exponential map for a matrix Lie group G does not necessarily map $\mathfrak{g}$ onto G. Furthermore, the exponential map may not be injective on $\mathfrak{g}$, as may be seen, for example, from the case $\mathfrak{g} = \mathsf{su}(2)$: the matrices

$$\begin{pmatrix} 2k\pi i & 0 \\ 0 & -2k\pi i \end{pmatrix}$$

are in $\mathsf{su}(2)$ and exp sends them all to the 2×2 identity matrix. Nevertheless, the exponential map provides a crucial mechanism for passing information between the group and the Lie algebra. Indeed, we will see that the exponential map is *locally* bijective (Corollary 3.47), a result that will be essential later.

Theorem 3.44. *For* $0 < \varepsilon < \log 2$, *let* $U_\varepsilon = B_{M_n(\mathbb{C})}(0, \varepsilon) = \{X \in M_n(\mathbb{C}) : \|X\| < \varepsilon\}$ *and let* $V_\varepsilon = \exp(U_\varepsilon)$. *Suppose* $G \subset \mathsf{GL}(n; \mathbb{C})$ *is a matrix Lie group with Lie algebra* $\mathfrak{g}$. *Then there exists* $\varepsilon \in (0, \log 2)$ *such that for all* $A \in V_\varepsilon$, A *is in* G *if and only if* $\log A$ *is in* $\mathfrak{g}$.

The condition $\varepsilon < \log 2$ guarantees (Theorem 2.18) that for all $X \in U_\varepsilon$, $\log(e^X)$ is defined and equal to X. Note that if $\log A$ is in $\mathfrak{g}$, then $A = e^{\log A}$ is in G. Thus, the content of the theorem is that for some ε, having A in $V_\varepsilon \cap G$ implies that $\log A$ must be in $\mathfrak{g}$.

We begin with a lemma.

Lemma 3.45. *Suppose* B_m *are elements of* G *and that* $B_m \to I$. *Let* $Y_m = \log B_m$, *which is defined for sufficiently large* m. *Suppose that* Y_m *is nonzero for all* m *and that* $Y_m / \|Y_m\| \to Y \in M_n(\mathbb{C})$. *Then* Y *is in* $\mathfrak{g}$.

Proof. Since $B_m \to I$, we have $\|Y_m\| \to 0$. Thus, we can find integers k_m such that $k_m \|Y_m\| \to t$.[11] Then, by the continuity of the exponential, we have

$$e^{k_m Y_m} = \exp\left[(k_m \|Y_m\|) \frac{Y_m}{\|Y_m\|}\right] \to e^{tY}.$$

[11]For real x, denote by $\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\}$ the floor function of x and by $\{x\} = x - \lfloor x \rfloor$ the mantissa of x, which is always $0 \leq \{x\} < 1$. Then, if x_m are non-zero real numbers with $x_m \to 0$, we have $x_m \left\lfloor \frac{t}{x_m} \right\rfloor = x_m \left(\frac{t}{x_m} - \left\{\frac{t}{x_m}\right\}\right) = t - x_m \left\{\frac{t}{x_m}\right\} \to t$ and $k_m = \left\lfloor \frac{t}{x_m} \right\rfloor$ are the desired integers.

However,
$$e^{k_m Y_m} = (e^{Y_m})^{k_m} = (B_m)^{k_m} \in G$$
and G is closed, and we conclude that $e^{tY} \in G$. This shows that $Y \in \mathfrak{g}$. □

Proof of Theorem 3.44. Let us think of $M_n(\mathbb{C})$ as $\mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$ and let D denote the orthogonal complement of $\mathfrak{g}$ with respect to the usual inner product on $\mathbb{R}^{2n^2}$. Consider the map $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ given by
$$\Phi(Z) = e^X e^Y,$$
where $Z = X + Y$ with $X \in \mathfrak{g}$ and $Y \in D$. By Theorem 2.28, the exponential is continuously differentiable, and thus the map Φ is also continuously differentiable. We may compute then that
$$\left.\frac{d}{dt}\Phi(tX, 0)\right|_{t=0} = X,$$
$$\left.\frac{d}{dt}\Phi(0, tY)\right|_{t=0} = Y.$$
This calculation shows that the derivative of Φ at the point $0 \in \mathbb{R}^{2n^2}$ is the identity. (Recall that the derivative at a point of a function from $\mathbb{R}^{2n^2}$ to itself is a linear map of $\mathbb{R}^{2n^2}$ to itself.) Since the derivative of Φ at the origin is invertible, the inverse function theorem says that Φ has a continuous local inverse, defined in a neighborhood of I.

We need to prove that for some ε, if $A \in V_\varepsilon \cap G$, then $\log A \in \mathfrak{g}$. If this were not the case, we could find a sequence A_m in G such that $A_m \to I$ as $m \to \infty$ and such that for all m, $\log A_m \notin \mathfrak{g}$.[12] Using the local inverse of the map Φ, we can write A_m (for all sufficiently large m) as
$$A_m = e^{X_m} e^{Y_m}, \quad X_m \in \mathfrak{g},\ Y_m \in D,$$
with X_m and Y_m tending to zero as m tends to infinity. We must have $Y_m \neq 0$, since otherwise we would have $\log A_m = X_m \in \mathfrak{g}$. Since e^{X_m} and A_m are in G, we see that
$$B_m := e^{-X_m} A_m = e^{Y_m}$$
is in G.

Since the unit sphere in D is compact, we can choose a subsequence of the Y_m's (still called Y_m) so that $Y_m/\|Y_m\|$ converges to some $Y \in D$, with $\|Y\| = 1$. Then, by the lemma, $Y \in \mathfrak{g}$. This is a contradiction, because D is the orthogonal complement of $\mathfrak{g}$. Thus, there must be some ε such that $\log A \in \mathfrak{g}$ for all A in $V_\varepsilon \cap G$. □

[12] Indeed, for each $m \geq \frac{1}{\log 2}$ and setting $\varepsilon = \frac{1}{m}$, we could take some $A_m \in V_{1/m} \cap G$ with $\log A_m \neq \mathfrak{g}$. For any $0 < \varepsilon < \log 2$, the maps $U_\varepsilon \underset{\log}{\overset{\exp}{\rightleftarrows}} V_\varepsilon$ are inverse homeomorphisms of each other, and V_ε is open, since log is continuous and injective and we thus have $V_\varepsilon = \log^{-1}(U_\varepsilon)$, where U_ε is open. Hence, since $\{U_{1/m}\}_{m \geq 1/\log 2}$ is a local basis at 0, we conclude that $\{V_{1/m}\}_{m \geq 1/\log 2}$ is a local basis at I. Thus, by construction, $A_m \to I$.

3.8 Consequences of Theorem 3.44

In this section, we derive several consequences of the main result of the last section, Theorem 3.44.

Corollary 3.46 (Closed-subgroup theorem for matrix Lie groups). *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$ and let k be the dimension of $\mathfrak{g}$ as a real vector space. Then G is a smooth embedded submanifold of $M_n(\mathbb{C})$ of dimension k and hence a Lie group.*

It follows from the corollary that G is locally path connected: every point in G has a neighborhood U that is homeomorphic to a ball in $\mathbb{R}^k$ and hence path connected. Thus, the connected components of G coincide with its path connected components (see, for example, Theorem 25.5 of [Mun]) and, in particular, G is connected if and only if it is path connected.

Proof. Let ε be such that Theorem 3.44 holds. Then for any $A_0 \in G$, consider the neighborhood A_0V_ε of A_0 in $M_n(\mathbb{C})$. Note that $A \in A_0V_\varepsilon$ if and only if $A_0^{-1}A \in V_\varepsilon$. Define a local coordinate system on A_0V_ε by writing each $A \in A_0V_\varepsilon$ as $A = A_0e^X$, for $X \in U_\varepsilon \subset M_n(\mathbb{C})$. More precisely: the maps

$$\begin{aligned} A_0V_\varepsilon &\leftrightarrow U_\varepsilon \\ A &\mapsto \log(A_0^{-1}A) \\ A_0e^X &\mapsfrom X \end{aligned}$$

are inverse diffeomorphisms of each other, by Theorem 2.28. For $A \in A_0V_\varepsilon$, it follows from Theorem 3.44 that $A \in G$ if and only if $X = \log(A_0^{-1}A) \in \mathfrak{g}$. Thus, in this local coordinate system defined near A_0, the group G looks like the subspace $\mathfrak{g}$ of $M_n(\mathbb{C})$. Since we can find such local coordinates near any point A_0 in G, we conclude that G is an embedded submanifold of $M_n(\mathbb{C})$.

Now, the operation of matrix multiplication is clearly smooth. Furthermore, the map $A \mapsto A^{-1}$ is also smooth in $\mathsf{GL}(n;\mathbb{C})$ (Proposition 1.3). The restrictions of these maps to G are then also smooth, showing that G is a Lie group. □

It follows that the exponential map of a matrix Lie group is smooth.

Corollary 3.47. *If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, there exists an open neighborhood U of 0 in $\mathfrak{g}$ and an open neighborhood V of I in G such that the exponential map takes U diffeomorphically onto V. Furthermore, the inverse of this map is the matrix logarithm restricted to V.*

Proof. Let ε be such that Theorem 3.44 holds and set $U = U_\varepsilon \cap \mathfrak{g}$ and $V = V_\varepsilon \cap G$. Theorem 3.44 implies that exp takes U onto V. Observe that V_ε is open, for the reason given in footnote 12 of page 62, so V is open in G. Furthermore, exp is a diffeomorphism of U onto V, since by Theorem 2.28 it is differentiable and there is a differentiable inverse map, namely, the restriction of the matrix logarithm to V. $\square$

In other words, log and exp are local diffeomorphisms, respectively, at $I \in G$ and at $0 \in \mathfrak{g}$ between the group and the Lie algebra.

Corollary 3.48. *Suppose $G \subset \mathsf{GL}(n;\mathbb{C})$ is a matrix Lie group with Lie algebra $\mathfrak{g}$. Then a matrix X is in $\mathfrak{g}$ if and only if there exists a smooth curve γ in $M_n(\mathbb{C})$ with $\gamma(t) \in G$ for all t and such that $\gamma(0) = I$ and $\left.\frac{d\gamma}{dt}\right|_{t=0} = X$. Thus, $\mathfrak{g}$ is the **tangent space at the identity** to G.*

Proof. If X is in $\mathfrak{g}$, then we may take $\gamma(t) = e^{tX}$ and then $\gamma(0) = I$ and $\left.\frac{d\gamma}{dt}\right|_{t=0} = X$. In the other direction, suppose $\gamma(t)$ is a smooth curve in G with $\gamma(0) = I$. For all sufficiently small t, we can write $\gamma(t) = e^{\delta(t)}$, where $\delta(t) = \log(\gamma(t))$ is in G and $\delta(0) = 0$. Since exp is infinitely differentiable, Proposition 2.13 tells us that the differential of exp at the zero matrix equals the identity, and thus, by the chain rule, we have

$$\gamma'(0) = D\exp(\delta(0))\delta'(0) = \delta'(0).$$

Since $\delta(t)$ belongs to $\mathfrak{g}$ for all sufficiently small t, we conclude (as in the proof of Theorem 3.12) that $\delta'(0) = \gamma'(0)$ belongs to $\mathfrak{g}$. $\square$

It follows that if G is matrix Lie group, then e^{tX} is in G for all $t \in (-\varepsilon, \varepsilon)$ if and only if e^{tX} is in G for all real t.

Corollary 3.49. *If G is a connected matrix Lie group, every element A of G can be written in the form*

$$A = e^{X_1}e^{X_2}\cdots e^{X_m} \tag{3.14}$$

for some $X_1, X_2, \ldots, X_m \in \mathfrak{g}$.

Even if G is connected, it is in general not possible to write every $A \in G$ as a single exponential, $A = \exp X$, with $X \in \mathfrak{g}$. (See Example 3.43.) We begin with a simple analytic lemma.

Lemma 3.50. *Suppose $A : [a,b] \to \mathsf{GL}(n;\mathbb{C})$ is a continuous map. Then for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $s, t \in [a,b]$ satisfy $|a - b| < \delta$, then*

$$\left\|A(s)A(t)^{-1} - I\right\| < \varepsilon.$$

Proof. We note that

$$\begin{aligned} \left\|A(s)A(t)^{-1} - I\right\| &= \left\|(A(s) - A(t))A(t)^{-1}\right\| \\ &\le \|A(s) - A(t)\| \left\|A(t)^{-1}\right\|. \end{aligned} \tag{3.15}$$

Since $[a,b]$ is compact and the map $t \mapsto \left\|A(t)^{-1}\right\|$ is continuous, there is a constant C such that $\left\|A(t)^{-1}\right\| \le C$ for all $t \in [a,b]$. Furthermore, since $[a,b]$ is compact, Theorem 4.19 in [Rud1] in tells us that the map A is actually *uniformly* continuous on $[a,b]$. Thus, for any $\varepsilon > 0$, there exists $\delta > 0$ such that when $|s-t| < \delta$, we have $\|A(s) - A(t)\| < \varepsilon/C$. Thus, in light of (3.15), we have the desired δ. ◻

Proof of Corollary 3.49. Let V be a neighborhood of I in G such that V is contained in the image of the exponential map of G (such a V exists due to Corollary 3.47). For any $A \in G$, choose a continuous path $A : [0,1] \to G$ with $A(0) = I$ and $A(1) = A$. By Lemma 3.50, we can find some $\delta > 0$ such that if $|s-t| < \delta$, then $A(s)A(t)^{-1} \in V$. Now divide $[0,1]$ into m pieces of equal length, where $1/m < \delta$. Then for $j = 1, \ldots, m$, we see that $A((j-1)/m)^{-1}A(j/m)$ belongs to V, so that

$$A((j-1)/m)^{-1}A(j/m) = e^{X_j}$$

for some elements $X_1, \ldots, X_m$ of $\mathfrak{g}$. Thus,

$$\begin{aligned} A &= A(0)^{-1}A(1) \\ &= A(0)^{-1}A(1/m)A(1/m)^{-1}A(2/m) \cdots A((m-1)/m)^{-1}A(1) \\ &= e^{X_1}e^{X_2} \cdots e^{X_m}, \end{aligned}$$

as claimed. ◻

Given a category C and objects x, y of C, it is traditional to write

$$\mathsf{C}(x,y) \quad \text{or} \quad \operatorname{Hom}(x,y)$$

for the class of morphisms from x to y.

Definition 3.51. Let C and D be categories. A functor $F : \mathsf{C} \to \mathsf{D}$ is said to be **faithful**[13] if for each two objects $x, y \in \mathsf{C}$, the induced map $\mathsf{C}(x,y) \xrightarrow{F} \mathsf{D}(Fx, Fy)$ is injective. If $\mathcal{O}$ is a subclass of the class of objects of C, the functor F is said to be **faithful for morphisms departing** $\mathcal{O}$ if for each two objects x of $\mathcal{O}$ and y of C the map $\mathsf{C}(x,y) \xrightarrow{F} \mathsf{D}(Fx, Fy)$ is injective.

Corollary 3.52. *The Lie functor is faithful for morphisms departing connected matrix Lie groups.*

Proof. Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and assume that G is connected. Suppose Φ_1 and Φ_2 are Lie group homomorphisms of G into H and let ϕ_1 and ϕ_2 be the associated Lie algebra homomorphisms of $\mathfrak{g}$ into $\mathfrak{h}$. We assume $\phi_1 = \phi_2$ and show that then $\Phi_1 = \Phi_2$.

Since G is connected, Corollary 3.49 tells us that every element of G can be written as $e^{X_1}e^{X_2} \cdots e^{X_m}$, with $X_j \in \mathfrak{g}$. Thus,

$$\begin{aligned} \Phi_1(e^{X_1} \cdots e^{X_m}) &= \Phi_1(e^{X_1}) \cdots \Phi_1(e^{X_m}) \\ &= e^{\phi_1(X_1)} \cdots e^{\phi_1(X_m)} \\ &= e^{\phi_2(X_1)} \cdots e^{\phi_2(X_m)} \\ &= \Phi_2(e^{X_1} \cdots e^{X_m}), \end{aligned}$$

[13] In Spanish, the word is *funtor fiel*.

as claimed. □

Corollary 3.53. *Every continuous homomorphism between two matrix Lie groups is smooth.*

Proof. Let U and V be open neighborhoods, respectively, of 0 in $\mathfrak{g}$ and of I in G such that $U \underset{\log}{\overset{\exp}{\rightleftarrows}} V$ are inverse diffeomorphisms of each other. If $g \in G$, then gV is open, for the right-multiplication map in a Lie group is a homeomorphism (the inverse of the continuous map $h \mapsto gh$ is the continuous map $h \mapsto g^{-1}h$). Observe that $h \in gV \Leftrightarrow g^{-1}h \in V$ and consider the coordinate map

$$\begin{aligned} f : gV &\to U \\ h &\mapsto \log(g^{-1}h) \end{aligned}$$

from the open neighborhood gV of G and the open set U of the vector space $\mathfrak{g}$. It is indeed a coordinate map, for it is a diffeomorphism: its inverse is $f^{-1} : X \in U \mapsto ge^X \in gV$ and both f and f^{-1} are smooth, since exp and log are and so is right-multiplication in a Lie group.

Let $\Phi : G \to G'$ be a Lie group homomorphisms between matrix Lie groups G and G'. Then Φ, read with coordinates given by f, is

$$\Phi(f^{-1}(X)) = \Phi(ge^X) = \Phi(g)\Phi(e^X) = \Phi(g)e^{\phi(X)},$$

which is a smooth function of $X \in U$, since right-multiplication and the exponential map of a Lie group are both smooth. So we have that $\Phi \circ f^{-1}$ is smooth in U. Since f is also smooth, the composition $(\Phi \circ f^{-1}) \circ f = \Phi|_{gV}$ is smooth as well. This shows that for each $g \in G$, the map Φ is smooth in an open neighborhood of g; or, equivalently, that Φ is smooth in the whole G. □

Corollary 3.54. *If G is a connected matrix Lie group and the Lie algebra $\mathfrak{g}$ of G is commutative, then G is commutative.*

This result is a partial converse to Proposition 3.14.

Proof. Since $\mathfrak{g}$ is commutative, any two elements of G, when written as in Corollary 3.49, will commute. □

4 | Basic Representation Theory

4.1 Representations and Actions

If M is a manifold, X is a set and $f : M \to X$ is a bijection, the function f induces both a topology and a differentiable structure in X in a natural way, so that f is a diffeomorphism. Furthermore, this differentiable structure is independent of diffeomorphisms in M in the following sense:

Lemma 4.1. *If $\Psi : M \to M$ is a diffeomorphism, then the induced differentiable structures on X by f and by $f \circ \Psi$ are the same.*

Proof. Let X_1 and X_2 be the manifold with underlying set equal to X and with topology and differentiable structure equal to that induced by f and $f \circ \Psi$, respectively. The two topologies on X are the same one: since both $f : M \to X_1$ and $f \circ \Psi : M \to X_2$ are homeomorphisms, we have

$$\begin{aligned} U \subset X_1 \text{ is open} &\Leftrightarrow f^{-1}(U) \subset M \text{ is open,} \\ &\Leftrightarrow \Psi^{-1}(f^{-1}(U)) \subset M \text{ is open,} \\ &\Leftrightarrow \underbrace{(f \circ \Psi)(\Psi^{-1}(f^{-1}(U)))}_{U} \subset X_2 \text{ is open.} \end{aligned}$$

Similarly, the two differentiable structures in X are the same one. Let $U \subset X$ be open in X_1 (equivalently, in X_2) and $G \subset \mathbb{R}^n$ be open. Let U_1 and U_2 be the manifold with underlying set equal to U and with differential structure equal to the restriction to U of that of X_1 and X_2, respectively. Then, since both $f : M \to X_1$ and $f \circ \Psi : M \to X_2$ are diffeomorphisms, we have

$$\begin{aligned} \varphi : U_1 \to G \text{ is a diffeomorphism} &\Leftrightarrow \varphi \circ f : f^{-1}(U) \to G && \text{is a diffeom.} \\ &\Leftrightarrow \varphi \circ f \circ \Psi : \Psi^{-1}(f^{-1}(U)) \to G && \text{is a diffeom.} \\ &\Leftrightarrow \underbrace{\varphi \circ f \circ \Psi \circ (f \circ \Psi)^{-1}}_{\varphi} : U_2 \to G && \text{is a diffeom.} \end{aligned} \tag{4.1}$$

The situation is depicted in the following commutative diagram:

$$\begin{tikzcd} U_2 \arrow[r, "(f\circ\Psi)^{-1}"] \arrow[drrr, bend right] & \Psi^{-1}((f^{-1}(U)) \arrow[r, "\Psi"] & f^{-1}(U) \arrow[r, "f"] & U_1 \arrow[d] \\ & & & G \end{tikzcd}$$

Equivalence (4.1) means that coordinate maps of X_1 are the same ones as those of X_2. That is, $X_1 = X_2$ have same differential structure. □

If V is a finite-dimensional real or complex space, let $\mathsf{GL}(V)$ denote the group of invertible linear transformations of V. If we choose a basis for V, we can identify $\mathsf{GL}(V)$ with $\mathsf{GL}(n;\mathbb{R})$ or $\mathsf{GL}(n;\mathbb{C})$. Any such identification gives rise to a differentiable structure on $\mathsf{GL}(V)$, and from Lemma 4.1 it follows that the differential structure is independent of the choice of basis. With this discussion in mind, we think of $\mathsf{GL}(V)$ as a matrix Lie group. Similarly, we let $\mathsf{gl}(V) = \mathrm{End}(V)$ denote the space of all linear operators from V to itself, which forms a Lie algebra under the bracket $[X, Y] = XY - YX$. Choosing a basis in V allows to identify the space $\mathsf{gl}(V)$ with $\mathsf{gl}(n;\mathbb{R})$ or $\mathsf{gl}(n;\mathbb{C})$, and as before, this identification induces a differentiable structure on $\mathsf{gl}(V)$ which is seen to be independent of the choice of basis by Lemma 4.1.

Definition 4.2. Let G be a matrix Lie group. A **complex representation** of G is a Lie group homomorphism

$$\Pi : G \to \mathsf{GL}(V),$$

where V is a finite-dimensional complex vector space (with $\dim(V) \geq 1$). A **real representation of** G is a Lie group homomorphism Π of G into $\mathsf{GL}(V)$, where V is a finite-dimensional real vector space.

If $\mathfrak{g}$ is a real or complex Lie algebra, then a **complex representation** of $\mathfrak{g}$ is a Lie algebra homomorphism π of $\mathfrak{g}$ into $\mathsf{gl}(V)$, where V is a finite-dimensional complex vector space. If $\mathfrak{g}$ is a real Lie algebra, then a **real representation** of $\mathfrak{g}$ is a Lie algebra homomorphism π of $\mathfrak{g}$ into $\mathsf{gl}(V)$.

If Π or π is an injective homomorphism, the representation is called **faithful**.

If in any of the previous cases the involved vector space V is finite-dimensional, the representation is called **finite-dimensional**.

All representations of Lie groups and Lie algebras that we shall consider will be finite-dimensional. After all, matrices are the main object of interest in this thesis.

Whenever one wants to insist on the implicated vector space, a representation of a Lie group or Lie algebra is sometimes also defined as a pair (V, Π), where V and Π are like in Definition 4.2. The analogous notation (V, π) for a representation of a Lie algebra may be used as well. Also, if it is the field of scalars of the vector space what one wants to highlight, one speaks of a $\mathbb{K}$**-representation** of a Lie group or Lie algebra in a vector space V, where $\mathbb{K}$ is the field of scalars of V.

A representation can be understood as a *linear action* of a group or Lie algebra on a vector space (since, say, to every $g \in G$, there associated an operator $\Pi(g)$, which acts on the space V). If the homomorphism $\Pi : G \to \mathsf{GL}(V)$ is fixed, we will occasionally use the notation

$$g \cdot v \tag{4.2}$$

in place of $\Pi(g)v$. We will often use terminology such as "Let Π be a representation of G acting on the space V."

If a representation Π is a faithful representation of a matrix Lie group G, then $\{\Pi(A) \mid A \in G\}$ is a group of matrices that is isomorphic to the original group G. Thus, Π allows us to *represent* G as a group of matrices. This is the motivation for the term "representation." (Of course, we still call Π a representation even if it is not faithful.) Despite the origin of the term, the goal of representation theory is *not* simply to represent a group as a group of matrices. After all, the groups we study in this thesis are already matrix Lie groups! Rather, the goal is to determine (up to isomorphism) *all* the ways a fixed group can act as a group of matrices.

Linear actions of groups on vector spaces arise naturally in many branches of both mathematics and physics. A typical example would be a linear differential equation in three-dimensional space which has rotational symmetry, such as the equation that describe the energy states of a hydrogen atom in quantum mechanics. Since the equation is rotationally invariant, the space of solutions is invariant under rotations and thus constitutes a representation of the rotation group $\mathsf{SO}(3)$. The representation theory of $\mathsf{SO}(3)$ (or of its Lie algebra) is very helpful in narrowing down what the space of solutions can be. See, for example, Chapter 18 in [Hall2].

There is a precise way to understand what do we mean by an "action" of a group when considering representations of the group.

Definition 4.3. Let G be a group and let X be a set. A **group action** of G on the set X is a map

$$\begin{aligned} \bullet : G \times X &\to X \\ (g, x) &\mapsto g \bullet x \end{aligned}$$

that satisfies the **group action axioms**:

$$\begin{aligned} &\text{Identity:} && e \bullet x = x, \\ &\text{Compatibility:} && g \bullet (h \bullet x) = (gh) \bullet x \end{aligned}$$

for all $x \in X$ and $g, h \in G$. If X is a vector space, then the group action is **linear** if

$$\begin{aligned} g \bullet (u + v) &= g \bullet u + g \bullet v \\ g(\lambda v) &= \lambda(g \bullet v) \end{aligned}$$

for all $g \in G$, $u, v \in X$ and every scalar λ. That is, if for each $g \in G$, the map $v \mapsto g \cdot v$ is linear.

If G is a Lie group and X is a manifold, a group action of G on X is called a **Lie group action** if the action $\cdot : G \times X \to X$ is smooth.

It is customary to say things like "the group G acts on the set X" to refer to an action $G \times X \to X$. The compatibility axiom makes possible to use the notation gx for a group action, instead of $g \cdot x$. This is because compatibility enables writing expressions like ghx, which can be interpreted unambiguously (where $g, h \in G$ and $x \in X$), and so, the compatibility axiom can be regarded as an "associative law" of the action.

A group action is the way in which symmetry on a object is mathematically disclosed, it is the way to explicit the symmetry of the object. There is this idea that mathematical groups encode all the possible ways in which things can be symmetric; and in this picture, the group actions are deemed to be the way to express the symmetry of a particular object through this group coding. For that reason, there are some mathematicians that argue that group actions are the *raison d'étre* of groups themselves.

Given an action of G on a set X, it follows from the action axioms that, for each $g \in G$, the map $x \mapsto g \cdot x$ is a bijection, since its inverse is $x \mapsto g^{-1} \cdot x$. Furthermore, if $\operatorname{Sym} X$ is the **symmetric group of** X (that is, the group of bijections $X \to X$), then every group action $\cdot : G \times X \to X$ induces a map

$$\begin{aligned} \Pi_{\cdot} : G &\to \operatorname{Sym} X \\ g &\mapsto \Pi_{\cdot}(g) = (x \mapsto g \cdot x) \end{aligned}$$

which is seen to be a group homomorphism, by the compatibility axiom. A **permutation representation** of a group G on a set X is defined to be a group homomorphism $G \to \operatorname{Sym} X$. The previous observation is then stated as "every group action on a set induces a permutation representation of the group on the same set." If $X = V$ is a vector space and the action is linear, then the induced permutation representation is really $G \to \mathsf{GL}(V) \subset \operatorname{Sym} V$, and this is a **representation** of G on the vector space X. If G is Lie group, the space V is a finite-dimensional $\mathbb{K}$-space, with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$,[1] and the linear action is a Lie group linear action, then the induced homomorphism $G \to \mathsf{GL}(V)$ is smooth and thus a Lie group $\mathbb{K}$-representation: let $\{v_1, \ldots, v_n\}$ be a basis in V and let $f : V \to \mathbb{K}^n$ be the linear isomorphism given by $f(v_j) = e_j$, where e_j is the j-th vector from the canonical basis of $\mathbb{K}^n$. Then, for each $g \in G$, the matrix of the linear automorphism $v \mapsto g \cdot v$ of V with respect to our basis is

$$(f(g \cdot e_1) \mid \cdots \mid f(g \cdot e_n)). \tag{4.3}$$

[1]We can endow any real or complex finite-dimensional V with a canonical differentiable structure by means of choosing a a basis on V and after identifying $V \cong \mathbb{K}^n$ (where $\mathbb{K}$ is the field of scalars of V). By Lemma 4.1, the differential structure is independent of the choice of basis.

Since f is smooth and so is the group action, each entry of this matrix depends smoothly on g and thus the induced homomorphism $G \longrightarrow \mathsf{GL}(V)$ is smooth.

Conversely, every permutation representation of a group on some set induces a group action on the same set. If $\Pi : G \longrightarrow \operatorname{Sym} X$ is a permutation representation of G on X, then the map

$$\begin{aligned} \bullet_\Pi : G \times X &\longrightarrow X \\ (g, x) &\longmapsto g \bullet_\Pi x = \Pi(g)x \end{aligned}$$

is shown to be a group action of G on X. If $X = V$ is a vector space and the permutation representation is actually a representation $\Pi : G \longrightarrow \mathsf{GL}(V)$, then the induced group action on V is linear. If, in addition, G is a Lie group, the space V is finite-dimensional real or complex and $\Pi : G \longrightarrow \mathsf{GL}(V)$ is a Lie group representation, then it follows from the fact that the application map is differentiable (see proof of Proposition 3.49) that the induced linear action $\bullet_\Pi$ is smooth and thus a Lie group linear action.

Proposition 4.4. *Let G be a group and X be a set. There is a one-to-one correspondence*

$$\begin{aligned} \{\text{Group actions of } G \text{ on } X\} &\leftrightarrow \{\text{Permutation representations of } G \text{ on } X\} \\ \bullet &\longmapsto \Pi. \\ \bullet_\Pi &\longleftarrow\!\shortmid\ \Pi \end{aligned} \tag{4.4}$$

Furthermore, if $X = V$ is a vector space, then the maps of (4.4) induce a one-to-one correspondence

$$\{\text{Linear group actions of } G \text{ on } V\} \leftrightarrow \{\text{Representations of } G \text{ on } V\}, \tag{4.5}$$

and if G is a Lie group and V is a is finite-dimensional $\mathbb{K}$-vector space, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, then the maps of (4.4) induce a one-to-one correspondence

$$\{\text{Lie group } \mathbb{K}\text{-linear actions of } G \text{ on } V\} \leftrightarrow \{\text{Lie group } \mathbb{K}\text{-representations of } G \text{ on } V\}. \tag{4.6}$$

Proof. The previous discussion shows that the maps that go from left to right and the maps that go from right to left in (4.4), (4.5) and (4.6) are well-defined.

The only thing left to be proved is that these maps are bijective, and for that it suffices to show bijectivity in (4.4). It must be shown that if $* : G \times X \longrightarrow X$ is a group action then $\bullet_{\Pi_*} = *$, and that if $\Sigma : G \longrightarrow \operatorname{Sym} X$ is a permutation representation, then $\Pi_{\bullet_\Sigma} = \Sigma$. Since this is only a matter of untangling the definitions of each induced map, the proof can be left safely to the reader. □

Group actions and group permutation representations are just two sides of the same coin, and it makes no difference to work with one of the first type or of the second type. Only context can say when it is more appropriate to either write things with the "action notation" or with the representation one. When some action or representation is fixed, mathematicians tend to change without further notice between the action and the representation at their own convenience, because this is allowed by Proposition 4.4.

Corollary 4.5. *Every linear group action of a matrix Lie group on a finite-dimensional real or complex vector space that is continuous is also smooth, and thus a Lie group action.*

Proof. Let G be a matrix Lie group, V be a finite-dimensional real or complex vector space and suppose that $\bullet : G \times V \longrightarrow V$ is a continuous linear action of G on V. Then, expressing the matrix of $\Pi_\bullet(g)$ in some basis of V like we did in (4.3), we get that the associated representation $\Pi_\bullet : G \longrightarrow \mathsf{GL}(V)$ is a continuous group homomorphism between matrix Lie groups. Thus, by Proposition 3.53, we have that $\Pi_\bullet$ is smooth. This means that $\Pi_\bullet$ is on the right-hand side of (4.6), and by the one-to-one correspondence, this implies that the action $\bullet$ was on the left-hand side of (4.6) the whole time. □

There also a way to understand representations of a Lie algebra as *Lie algebra actions* on a vector space.

Definition 4.6. Let $\mathfrak{g}$ be a Lie algebra over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and let V be a vector space over the same field $\mathbb{K}$. A **Lie algebra action** of $\mathfrak{g}$ on the vector space V is a map

$$\begin{aligned} \bullet : \mathfrak{g} \times V &\longrightarrow V \\ (X, v) &\longmapsto X \bullet v \end{aligned}$$

that satisfies the **Lie algebra action axioms**:

Linearity:	$X \bullet (u + v) = X \bullet u + X \bullet v, \quad X \bullet (\lambda v) = \lambda(X \bullet v)$
Compatibility of the bracket:	$[X, Y] \bullet v = X \bullet (Y \bullet v) - Y \bullet (X \bullet v)$

for every $u, v \in V,\ X, Y \in \mathfrak{g}$ and $\lambda \in \mathbb{K}$.

We will say that the Lie algebra action is **complex** (resp., **real**) if V is a complex (resp., real) vector space. Additionally, we allow the case in which the Lie algebra is real and the vector space V is complex. The axiom of linearity must be then fulfilled for all complex scalars λ and we will speak of a **Lie algebra action of a real Lie algebra on a complex vector space**.

If $\mathfrak{g}$ is a Lie algebra and V is a vector space, then every Lie algebra action $\bullet : \mathfrak{g} \times V \longrightarrow V$ of $\mathfrak{g}$ on V induces a Lie algebra representation

$$\begin{aligned} \pi_\bullet : \mathfrak{g} &\longrightarrow \mathsf{gl}(V) \\ X &\longmapsto \pi_\bullet(X) = (v \mapsto X \bullet v) \end{aligned}$$

The property of bracket preservation by the map $\pi_\bullet$ is precisely the compatibility axiom.

Conversely, if we start with a Lie algebra representation $\pi : \mathfrak{g} \to \mathsf{gl}(V)$ of $\mathfrak{g}$ on the vector space V, then the representation induces a Lie algebra action

$$\begin{aligned} \bullet_\pi : \mathfrak{g} \times V &\to V \\ (X, v) &\mapsto X \bullet_\pi v = \pi(X)v. \end{aligned}$$

Similarly as before with the case of groups actions and group representations, we have the following correspondence.

Proposition 4.7. *Let $\mathfrak{g}$ be a complex or real Lie algebra and let V be a complex or real vector space, but complex if $\mathfrak{g}$ is complex. There is a one-to-one correspondence*

$$\begin{aligned} \{\text{Lie algebra actions of } \mathfrak{g} \text{ on } V\} &\leftrightarrow \{\text{Lie algebra representations of } \mathfrak{g} \text{ in } V\} \\ \bullet &\mapsto \pi_\bullet \\ \bullet_\pi &\leftarrow\!\shortmid \pi \end{aligned}$$

Proof. By the previous discussion, the map from left to right and the map from right to left is well-defined.

Analogously as with the case of groups, the only thing that must be proven is that if $* : \mathfrak{g} \times V \to V$ is a Lie algebra action, then $\bullet_{\pi_*} = *$, and that if $\sigma : \mathfrak{g} \to \mathsf{gl}(V)$ is a Lie algebra representation, then $\pi_{\bullet_\sigma} = \sigma$. □

4.2 Properties of Representations

Definition 4.8. Let Π be a finite-dimensional real or complex representation of a matrix Lie group G, acting on a space V. A subspace W of V is called **invariant** if $\Pi(A)w \in W$ for all $w \in W$ and all $A \in G$. An invariant subspace W is called **nontrivial** if $W \neq \{0\}$ and $W \neq V$. A representation with no nontrivial invariant subspaces is called **irreducible**.

The terms **invariant**, **nontrivial**, and **irreducible** are defined analogously for representations of Lie algebras.

Even if $\mathfrak{g}$ is a Lie algebra, we will consider mainly *complex* representations of $\mathfrak{g}$. It should be emphasized that if we are speaking about complex representations of a real Lie algebra $\mathfrak{g}$ acting on a complex space V, an invariant subspace W is required to be a complex subspace of V.

Definition 4.9. Let G be a matrix Lie group, let Π be a representation of G acting on the space V, and let Σ be a representation of G acting on the space W. A linear map $\phi : V \to W$ is called an **intertwining map** of representations[2] if

$$\phi(\Pi(A)v) = \Sigma(A)\phi(v)$$

[2]In Spanish, one might say *aplicación entrelazadora de representaciones.*

for all $A \in G$ and all $v \in V$. Equivalently, the following diagram commutes for all $A \in G$:

$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[d, "\Pi(A)"'] & W \arrow[d, "\Sigma(A)"] \\ V \arrow[r, "\phi"] & W \end{tikzcd}$$

The analogous property defines an intertwining maps of representations of a Lie algebra.

If ϕ is an intertwining map of representations and, in addition, ϕ is invertible, then ϕ is said to be an **isomorphism** of representations. If there exists an isomorphism between V and W, then the representations are said to be **isomorphic**.

Whenever the representation maps are needed to be made explicit, the notation $\phi : (V, \Pi) \to (W, \Sigma)$ will be used for an intertwining map ϕ between representations (V, Π) and (W, Σ) of G. It is not difficult to check that the inverse function of an isomorphism of representations is also an intertwining map.

If we use the action notation, the defining property of an intertwining map may be written as

$$\phi(A \cdot v) = A \cdot \phi(v)$$

for all $A \in G$ and $v \in V$. That is to say, ϕ should commute with the action of G. A typical problem in representation theory is to determine, up to isomorphism, all of the irreducible representations of a particular group or Lie algebra. In Sect. 4.4 we will determine all the finite-dimensional complex irreducible representations of the Lie algebra $\mathsf{sl}(2;\mathbb{C})$.

If $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, given a matrix Lie group G we can define $\mathsf{Rep}_{\mathbb{K}}(G)$ the category of representations of G in vector spaces over $\mathbb{K}$, whose objects are pairs (V, Π) which are representations of G in vector spaces over the field $\mathbb{K}$ and whose morphisms are intertwining maps of representations of G between these representations, $(V, \Pi) \to (W, \Sigma)$. Analogously, given a real or complex Lie algebra $\mathfrak{g}$, we can define $\mathsf{Rep}_{\mathbb{C}}(\mathfrak{g})$ the category of complex representations of $\mathfrak{g}$. Similarly, if $\mathfrak{g}$ is a real Lie algebra, we can define $\mathsf{Rep}_{\mathbb{R}}(\mathfrak{g})$ the category of real representations of $\mathfrak{g}$. If we restrict to finite-dimensional representations, we will write $\mathsf{FinRep}_{\mathbb{K}}(G) \subset \mathsf{Rep}_{\mathbb{K}}(G)$ for the subcategory of finite-dimensional $\mathbb{K}$-representations of G. And $\mathsf{FinRep}_{\mathbb{K}}(\mathfrak{g}) \subset \mathsf{Rep}_{\mathbb{K}}(\mathfrak{g})$ for the subcategory of finite-dimensional $\mathbb{K}$-representations of $\mathfrak{g}$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ if $\mathfrak{g}$ is real and $\mathbb{K} = \mathbb{C}$ if $\mathfrak{g}$ is complex).

The previous categories are well defined, since the identity map is an intertwining map and since the composition of intertwining maps gives an intertwining maps. For example, for the case of groups, if (V, Π), (W, Σ) and (U, Γ) are representations of the matrix Lie group G and $\phi : (V, \Pi) \to (W, \Sigma)$ and $\psi : (W, \Sigma) \to (U, \Gamma)$ are intertwining maps, then, for all $A \in G$, we have that each square subdiagram of

$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[d, "\Pi(A)"'] & W \arrow[r, "\psi"] \arrow[d, "\Sigma(A)"'] & U \arrow[d, "\Gamma(A)"] \\ V \arrow[r, "\phi"'] & W \arrow[r, "\psi"'] & U \end{tikzcd}$$

commutes, and from that, commutativity of the whole diagram follows. That is, $\psi \circ \phi : (V, \Pi) \to (U, \Gamma)$ is an intertwining map.

After identifying $\mathsf{GL}(V)$ with $\mathsf{GL}(n; \mathbb{R})$ or $\mathsf{GL}(n; \mathbb{C})$, Theorem 3.29 has the following consequence.

Proposition 4.10. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$ and let Π be a (finite-dimensional real or complex) representation of G, acting on the space V. Then there is a unique representation π of $\mathfrak{g}$ acting on the same space such that*

$$\Pi(e^X) = e^{\pi(X)}$$

for all $X \in \mathfrak{g}$. The representation π can be computed as

$$\pi(X) = \frac{d}{dt}\Pi(e^{tX})\bigg|_{t=0}$$

and satisfies

$$\pi\left(AXA^{-1}\right) = \Pi(A)\pi(X)\Pi(A)^{-1}$$

for all $X \in \mathfrak{g}$ and all $A \in G$.

The induced representation is precisely $\pi = \operatorname{Lie} \Pi$. Given a matrix Lie group G with Lie algebra $\mathfrak{g}$, we may ask whether *every* representation π of $\mathfrak{g}$ comes from a representation Π of G. As it turns out, this is not true in general, but it is true if G is simply connected. See Sect. 4.5 for examples of this phenomenon and Sect. 5.2 for the general result.

Proposition 4.11. *Any intertwining map $\phi : (V, \Pi) \to (W, \Sigma)$ between representations (V, Π) and (W, Σ) of a matrix Lie group G is also an intertwining map $\phi : (V, \pi) \to (W, \sigma)$ between the induced representations (V, π) and (W, σ) of the Lie algebra $\mathfrak{g}$ (where $\pi = \operatorname{Lie}\Pi$ and $\sigma = \operatorname{Lie}\Sigma$).*

Proof. For all $A \in G$, we have $\phi \circ \Pi(A) = \Sigma(A) \circ \phi$, and in particular, $\phi \circ \Pi(e^{tX}) = \Sigma(e^{tX}) \circ \phi$ for all $X \in \mathfrak{g}$. In that case, by the product rule for linear operators (see footnote 1 in page 42), we have

$$\begin{aligned} \frac{d}{dt}[\phi \circ \Pi(e^{tX})]\bigg|_{t=0} &= \frac{d}{dt}[\Sigma(e^{tX}) \circ \phi]\bigg|_{t=0} \\ \phi \circ \left(\frac{d}{dt}\Pi(e^{tX})\bigg|_{t=0}\right) &= \left(\frac{d}{dt}\Sigma(e^{tX})\bigg|_{t=0}\right) \circ \phi \end{aligned}$$

$$\phi \circ \pi(X) = \sigma(X) \circ \phi,$$

as claimed. □

Proposition 4.12. *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. There is a covariant functor*

$$\mathsf{FinRep}_{\mathbb{K}}(G) \to \mathsf{FinRep}_{\mathbb{K}}(\mathfrak{g})$$

from the category of finite-dimensional $\mathbb{K}$-representations of a matrix Lie group G to the category of finite-dimensional $\mathbb{K}$-representations of its Lie algebra $\mathfrak{g}$. This functor sends each $\mathbb{K}$-representation (V, Π) of G to the representation (V, π) of $\mathfrak{g}$, where $\pi = \operatorname{Lie} \Pi$, and sends each intertwining map of representations $\phi : (V, \Pi) \to (W, \Sigma)$ of G to the intertwining map $\phi : (V, \pi) \to (W, \sigma)$ itself, now between the induced representations (V, π) and (W, σ) of the Lie algebra $\mathfrak{g}$ of G (where $\pi = \operatorname{Lie} \Pi$ and $\sigma = \operatorname{Lie} \Sigma$).

Furthermore, this functor is faithful.

Proof. Functoriality axioms follow at once from the fact that, set-theoretically speaking, the intertwining map $\phi : (V, \Pi) \to (W, \Sigma)$, as a function, is the same as the intertwining map $\phi : (V, \pi) \to (W, \sigma)$, so that if $\psi : (W, \Sigma) \to (U, \Gamma)$ is another intertwining map of representations of G, then, as linear maps between vector spaces, $\phi : V \to W$, $\psi : W \to U$ and $\psi \circ \phi : V \to U$ are always the same functions.

The functor is thus faithful by construction. □

By construction, this functor is faithful (recall Definition 3.51) and it reflects isomorphisms (recall Definition 3.28). If in addition G is connected, the functor gets another interesting property, namely, that of *fullness*.

Definition 4.13. Let C and D be categories. A functor $F : \mathsf{C} \to \mathsf{D}$ is said to be **full**[3] if for each two objects $x, y \in \mathsf{C}$, the induced map $\mathsf{C}(x, y) \xrightarrow{F} \mathsf{D}(Fx, Fy)$ is surjective. A functor which is full and faithful is said to be **fully faithful**.[4] Similarly as in Definition 3.51, if O is a subclass of the class of objects of C, the functor F is said to be **full for morphisms departing O** (resp., **fully faithful for morphisms departing O**) if for each two objects x of O and y of C the map $\mathsf{C}(x, y) \xrightarrow{F} \mathsf{D}(Fx, Fy)$ is surjective (resp., bijective).

If C and D are categories, we say that D is a **full subcategory** of C if D is a subcategory of C and the inclusion functor $\mathsf{D} \to \mathsf{C}$ is full. That is, D contains all possible morphisms between its objects. A faithful functor that is also injective on objects is called an **embedding** and identifies the domain category as a subcategory of the codomain. A fully faithful functor that is also injective on objects is called an **full embedding** of the domain category into the codomain category.

[3] In Spanish, the word is *funtor pleno*.

[4] In Spanish, *funtor plenamente fiel*.

The next lemma gives a remarkable property of fully faithful functors, a property that we will further use in Chapter 5. There, on Sect. 5.2, we will show that the Lie functor restricted to the full subcategory of simply connected matrix Lie group is fully faithful.

Recall Definition 3.28 for the concept of a functor that reflects and creates isomorphisms.

Lemma 4.14. *Fully faithful functors reflect and create isomorphisms.*

Proof. Let $F : \mathsf{C} \to \mathsf{D}$ be a fully faithful functor between categories C and D. Let first see that F reflects isomorphisms. Suppose $f : x \to y$ is a morphism in C such that Ff is an isomorphism. We now show that f is an isomorphism. There exists $g' \in \mathsf{D}(Fy, Fx)$ such that

$$\begin{aligned} Ff \circ g' &= 1_{Fy} \\ g' \circ Ff &= 1_{Fx}. \end{aligned} \tag{4.7}$$

Since F is full, there exists $h \in \mathsf{C}(y, x)$ such that $Fg = g'$. Equations (4.7) mean that

$$\begin{aligned} F(f \circ g) &= Ff \circ Fg = 1_{Fy} = F(1_y) \\ F(g \circ f) &= Fh \circ Ff = 1_{Fx} = F(1_x), \end{aligned} \tag{4.8}$$

but since F is faithful, this implies that $fg = 1_y$ and $gf = 1_x$. That is, f is an isomorphism.

Next we show that F create isomorphisms. Suppose x and y are objects in C such that Fx and Fy are isomorphic. Let's see that then x and y are isomorphic as well. There exist $f' \in \mathsf{D}(Fx, Fy)$ and $g' \in \mathsf{D}(Fy, Fx)$ such that

$$\begin{aligned} f'g' &= 1_{Fy} \\ g'f' &= 1_{Fx}. \end{aligned} \tag{4.9}$$

Since F is full, we can pick morphisms $f \in \mathsf{D}(x, y)$ and $g \in \mathsf{D}(x, y)$ such that $Ff = f'$ and $Fg = g'$. For that reason, equations (4.9) now read as does equations (4.8). In consequence, since F is faithful, we have that $fg = 1_y$ and $gf = 1_x$. That is, $x \cong y$. $\square$

Proposition 4.15. *If G is a connected matrix Lie group, then the functor*

$$\mathsf{FinRep}_{\mathbb{K}}(G) \to \mathsf{FinRep}_{\mathbb{K}}(\mathfrak{g})$$

of Proposition 4.12 is a full embedding.

This proposition says that the study of representations of a connected matrix Lie group is a particular case of the study of representations of its Lie algebra.

Proof. By construction, the functor is faithful. It is also injective on objects: If (V, Π) and (W, Σ) are representations of G such that $(V, \operatorname{Lie}\Pi) = (W, \operatorname{Lie}\Sigma)$, then $V = W$ and so, by Proposition 3.52, also $\Pi = \Sigma$ (if G is connected). To show fullness, let (V, Π)

and (W, Σ) be representations of G and write (V, π) and (W, σ) for the corresponding induced representations of the Lie algebra $\mathfrak{g}$ of G. Let $\phi : (V, \pi) \to (W, \sigma)$ be an intertwining map or representations. We must show that there exists some intertwining map $(V, \Pi) \to (W, \Sigma)$ whose image under the functor of Proposition 4.12 is ϕ. Due to the manner in which the functor is defined, the only possibility for such an intertwining map $(V, \Pi) \to (W, \Sigma)$ is ϕ itself. For that matter, we shall prove that $\phi \circ \Pi(A) = \Sigma(A) \circ \phi$ for all $A \in G$.

Let A be in G. Corollary 3.49 says that there exist $X_1, \ldots, X_m \in \mathfrak{g}$ such that $A = e^{X_1} \cdots e^{X_m}$, so we have

$$\begin{aligned}
\phi \circ \Pi(A) &= \phi \circ \Pi(e^{X_1} \cdots e^{X_m}) \\
&= \phi \Pi(e^{X_1}) \cdots \Pi(e^{X_m}) \\
&= \phi e^{\pi(X_1)} \cdots e^{\pi(X_m)}.
\end{aligned}$$

For any, $X \in \mathfrak{g}$, we know that $\phi\pi(X) = \sigma(X)\phi$. This identity implies that $\phi\pi(X)^k = \sigma(X)^k \phi$ for each k, and hence that $\phi e^{\pi(X)} = e^{\sigma(X)}\phi$, since, by continuity of the product of linear operators, we have

$$\begin{aligned}
\phi e^{\pi(X)} &= \phi \sum_{k=0}^{\infty} \frac{\pi(X)^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{\phi\pi(X)^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{\sigma(X)^k \phi}{k!} \\
&= \left(\sum_{k=0}^{\infty} \frac{\sigma(X)^k}{k!} \right) \phi \\
&= e^{\sigma(X)} \phi.
\end{aligned}$$

Thus, swapping each $\phi e^{\pi(X_j)} = e^{\sigma(X_j)}\phi$ one by one,

$$\begin{aligned}
\phi \Pi(A) &= \phi e^{\pi(X_1)} \cdots e^{\pi(X_m)} \\
&= e^{\sigma(X_1)} \cdots e^{\sigma(X_m)} \phi \\
&= \Sigma(e^{X_1}) \cdots \Sigma(e^{X_m}) \phi \\
&= \Sigma(e^{X_1} \cdots e^{X_m}) \phi \\
&= \Sigma(A) \phi,
\end{aligned}$$

as we wanted to show. □

Proposition 4.16.

1. *Let G be a connected matrix Lie group with Lie algebra $\mathfrak{g}$. Let Π be a representation of G and π be the associated representation of $\mathfrak{g}$. Then π irreducible implies Π irreducible, and the converse also holds if G is connected.*

2. *Let G be a matrix Lie group, let Π_1 and Π_2 be representations of G, and let π_1 and π_2 be the associated Lie algebra representations. Then Π_1 and Π_2 isomorphic implies π_1 and π_2 isomorphic, and the converse also holds if G is connected.*

Proof. For Point 1, suppose first that π is irreducible. We want to show that Π is irreducible. So, let W be a subspace of V that is invariant under $\Pi(A)$ for all $A \in G$. We want to show that W is either $\{0\}$ or V. In particular, W is invariant under $\Pi(e^{tX})$ for all $X \in \mathfrak{g}$ and, hence, under

$$\pi(X) = \frac{d}{dt}\Pi(e^{tX})\bigg|_{t=0}.$$

Indeed, as in proof of Proposition 3.36, for $v \in W$ we have

$$\begin{aligned}\pi(X)v &= \frac{d}{dt}\Pi(e^{tX})\bigg|_{t=0} v \\ &= \frac{d}{dt}\underbrace{\Pi(e^{tX})v}_{\in W}\bigg|_{t=0}.\end{aligned}$$

and the previous expression is in W, since we are taking a limit inside the topologically closed space W.

Since π is irreducible and W is invariant under each $\pi(X)$, W must be either $\{0\}$ or V. This shows that Π is irreducible.

In the other direction, suppose that G is connected and assume that Π is irreducible and that W is an invariant subspace for π. Let $A \in G$. Corollary 3.49 tells us that A can be written as $A = e^{X_1} \cdots e^{X_m}$ for some $X_1, \ldots, X_m$ in $\mathfrak{g}$. Since W is invariant under $\pi(X_j)$ it will also be invariant under $\exp(\pi(X_j)) = I + \pi(X_j) + \pi(X_j)^2/2 + \cdots$, by the continuity of the application map (see proof of Proposition 3.49) and due to the fact that W is topologically closed. Hence, W will be also invariant under

$$\begin{aligned}\Pi(A) &= \Pi(e^{X_1} \cdots e^{X_m}) = \Pi(e^{X_1}) \cdots \Pi(e^{X_m}) \\ &= e^{\pi(X_1)} \cdots e^{\pi(X_m)}\end{aligned}$$

Thus, since Π is irreducible, W is $\{0\}$ or V, and we conclude that π is irreducible. This establishes Point 1 of the proposition.

For Point 2, the first part follows from Corollary 3.27 and Proposition 4.12, so we have that in general $\Pi_1 \cong \Pi_2$ implies $\pi_1 \cong \pi_2$. If, in addition, G is connected, the functor of Proposition 4.12 is fully faithful (by Proposition 4.15) and thus it *creates* isomorphisms, by Lemma 4.14. That is, we have that $\pi_1 \cong \pi_2$ implies $\Pi_1 \cong \Pi_2$ for G connected. □

Proposition 4.17. *Let $\mathfrak{g}$ be a real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ its complexification. Then every finite-dimensional complex representation π of $\mathfrak{g}$ has a unique extension to a complex-linear representation of $\mathfrak{g}_{\mathbb{C}}$, also denoted π. Furthermore, π is irreducible as a representation of $\mathfrak{g}_{\mathbb{C}}$ if and only if it is irreducible as a representation of $\mathfrak{g}$.*

Of course, the extension of π is given by $\pi(X+iY) = \pi(X)+i\pi(Y)$ for all $X, Y \in \mathfrak{g}$.

Proof. The existence and uniqueness of the extension follows from the universal property of the complexification of a Lie algebra (Proposition 3.41). The claim about irreducibility holds because a complex subspace W of V is invariant under $\pi(X)$ for all $X \in \mathfrak{g}$ if and only if it is invariant under $\pi(X+iY) = \pi(X)+i\pi(Y)$ for all $X, Y \in \mathfrak{g}$. Thus, the representation of $\mathfrak{g}$ and its extension to $\mathfrak{g}_{\mathbb{C}}$ have precisely the same invariant subspaces. □

4.3 Examples of Representations

A matrix Lie group G is, by definition, a subset of some $\mathsf{GL}(n;\mathbb{C})$. The inclusion map of G into $\mathsf{GL}(n;\mathbb{C})$ (i.e., the map $\Pi(A) = A$) is a representation of G, called the **standard representation** of G. If G happens to be contained in $\mathsf{GL}(n;\mathbb{R}) \subset \mathsf{GL}(n;\mathbb{C})$, then we can also think of the standard representation as a real representation. Thus, for example, the standard representation of $\mathsf{SO}(3)$ is the one in which $\mathsf{SO}(3)$ acts in the usual way on $\mathbb{R}^3$ and the standard representation of $\mathsf{SU}(2)$ is the one in which $\mathsf{SU}(2)$ acts on $\mathbb{C}^2$ in the usual way. Similarly, if $\mathfrak{g} \subset M_n(\mathbb{C})$ is a Lie algebra of matrices, the map $\pi(X) = X$ is called the **standard representation** of $\mathfrak{g}$.

Consider the one-dimensional vector space $\mathbb{C}$. For any matrix Lie group G, we can define the **trivial representation**, $\Pi : G \to \mathsf{GL}(1;\mathbb{C})$, by the formula

$$\Pi(A) = I$$

for all $A \in G$. Of course, this is an irreducible representation, since $\mathbb{C}$ has no nontrivial subspaces, let alone nontrivial invariant subspaces. If $\mathfrak{g}$ is a Lie algebra, we can also define the **trivial representation** of $\mathfrak{g}$, $\pi : \mathfrak{g} \to \mathfrak{gl}(1;\mathbb{C})$, by

$$\pi(X) = 0$$

for all $X \in \mathfrak{g}$. This is an irreducible representation. Observe that both the Lie functor and the functor of Proposition 4.12 send the trivial representation of G, on the first case as a morphism and the second one as an object, to the trivial representation of $\operatorname{Lie} G$.

Recall the adjoint map of a group or Lie algebra, described in definitions 3.34 and 3.7.

Definition 4.18. If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, the **adjoint representation** of G is the map $\mathrm{Ad} : G \to \mathsf{GL}(\mathfrak{g})$ given by $A \mapsto \mathrm{Ad}_A$. Similarly, the **adjoint representation** of a finite-dimensional Lie algebra $\mathfrak{g}$ is the map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ given by $X \mapsto \mathrm{ad}_X$.

If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, then by Proposition 3.36, the Lie algebra representation associated to the adjoint representation of G is the adjoint representation of $\mathfrak{g}$.

Example 4.19. Let V_m denote the space of homogeneous polynomials of degree m in two complex variables. For each $U \in \mathsf{SU}(2)$, define a linear transformation $\Pi_m(U)$ on the space V_m by the formula

$$[\Pi_m(U)f](z) = f\left(U^{-1}z\right), \quad f \in V_m. \tag{4.10}$$

Then Π_m is a representation of $\mathsf{SU}(2)$.

Elements of V_m have the form

$$f(z_1, z_2) = a_0 z_1^m + a_1 z_1^{m-1} z_2 + a_2 z_1^{m-2} z_2^2 + \cdots + a_m z_2^m \tag{4.11}$$

with $z_1, z_2 \in \mathbb{C}$ and the a_j's arbitrary complex constants, from which we see that $\dim V_m = m + 1$. Explicitly, if f is as in (4.11), then

$$[\Pi_m(U)f](z_1, z_2) = \sum_{k=0}^{m} a_k \left(U_{11}^{-1} z_1 + U_{12}^{-1} z_2\right)^{m-k} \left(U_{21}^{-1} z_1 + U_{22}^{-1} z_2\right)^k .$$

By expanding out the right-hand side of this formula, we see that $\Pi_m(U)f$ is again a homogeneous polynomial of degree m. Thus, $\Pi_m(U)$ actually maps V_m into V_m.

To see that Π_m is actually a representation, compute that

$$\begin{aligned} \Pi_m(U_1)[\Pi_m(U_2)f](z) &= [\Pi_m(U_2)f]\left(U_1^{-1}z\right) = f\left(U_2^{-1}U_1^{-1}z\right) \\ &= \Pi_m(U_1U_2)f(z). \end{aligned}$$

The inverse on the right-hand side of (4.10) is necessary in order to make Π_m a representation. We will see in Proposition 4.20 that each Π_m is irreducible and we will see in Sect. 4.4 that every finite-dimensional irreducible representation of $\mathsf{SU}(2)$ is isomorphic to one (and only one) of the Π_m's. (Of course, no two of the Π_m's are isomorphic, since they do not even have the same dimension.)

The associated representation π_m of $\mathsf{su}(2)$ can be computed as

$$(\pi_m(X)f)(z) = \left.\frac{d}{dt} f(e^{-tX}z)\right|_{t=0}$$

Now, let $z(t) = (z_1(t), z_2(t))$ be the curve in $\mathbb{C}^2$ defined as $z(t) = e^{-tX}z$. By the chain rule, we have

$$\pi_m(X)f = \left.\frac{\partial f}{\partial z_1}\frac{dz_1}{dt}\right|_{t=0} + \left.\frac{\partial f}{\partial z_2}\frac{dz_2}{dt}\right|_{t=0}$$

Since $\left.\dfrac{dz}{dt}\right|_{t=0} = -Xz$, we obtain

$$\pi_m(X)f = -\frac{\partial f}{\partial z_1}(X_{11}z_1 + X_{12}z_2) - \frac{\partial f}{\partial z_2}(X_{21}z_1 + X_{22}z_2). \tag{4.12}$$

We may then take the unique complex-linear extension of π to $\mathfrak{sl}(2;\mathbb{C}) \cong \mathfrak{su}(2)_{\mathbb{C}}$, as in Proposition 3.41. This extension is given by the same formula, but with $X \in \mathfrak{sl}(2;\mathbb{C})$.

If X, Y, and H are the following basis elements for $\mathfrak{sl}(2;\mathbb{C})$:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{4.13}$$

then applying formula (4.12) gives

$$\begin{aligned} \pi_m(H) &= -z_1 \frac{\partial}{\partial z_1} + z_2 \frac{\partial}{\partial z_2} \\ \pi_m(X) &= -z_2 \frac{\partial}{\partial z_1} \\ \pi_m(Y) &= -z_1 \frac{\partial}{\partial z_2}. \end{aligned}$$

Applying these operators to a basis elements $z_1^{m-k} z_2^k$ for V_m gives

$$\begin{aligned} \pi_m(H)\left(z_1^{m-k} z_2^k\right) &= (-m+2k) z_1^{m-k} z_2^k, && 0 \le k \le m, \\ \pi_m(X)\left(z_1^{m-k} z_2^k\right) &= -(m-k) z_1^{m-k-1} z_2^{k+1}, && 0 \le k < m, \\ \pi_m(Y)\left(z_1^{m-k} z_2^k\right) &= -k z_1^{m-k+1} z_2^{k-1} && 0 < k \le m, \end{aligned} \tag{4.14}$$

and $\pi_m(X)(z_2^m) = \pi_m(Y)(z_1^m) = 0$. Thus, $z_1^{m-k} z_2^k$ is an eigenvector for $\pi_m(H)$ with eigenvalue $-m+2k$, while $\pi_m(X)$ and $\pi_m(Y)$ have the effect of shifting the exponent k of z_2 up or down by one. Note that since $\pi_m(X)$ increases the value of k, this operator increases the eigenvalue of $\pi_m(H)$ by 2, whereas $\pi_m(Y)$ decreases the eigenvalue of $\pi_m(H)$ by 2.

Proposition 4.20. *For each $m \ge 0$, the representation π_m is irreducible.*

Proof. It suffices to show that every nonzero invariant subspace of V_m is equal to V_m. So, let W be such a space and let w be a nonzero element of W. Then w can be written in the form

$$w = a_0 z_1^m + a_1 z_1^{m-1} z_2 + a_2 z_1^{m-2} z_2^2 + \cdots + a_m z_2^m$$

with at least one of the a_k's being nonzero. Let k_0 be the smallest value of k for which $a_k \neq 0$ and consider

$$\pi_m(X)^{m-k_0} w.$$

Since each application of $\pi_m(X)$ raises the power of z_2 by 1, $\pi_m(X)^{m-k_0}$ will kill all the terms in w except $a_{k_0} z_1^{m-k_0} z_2^{k_0}$ term. On the other hand, since $\pi_m(z_1^{m-k} z_2^k)$ is zero only if $k = m$, we see that $\pi_m(X)^{m-k_0} w$ is a nonzero multiple of z_2^m. since W is assumed invariant, W must contain this multiple of z_2^m and so also z_2^m itself. Now, for $0 \le k \le m$, it follows from (4.14) that $\pi_m(Y)^k z_2^m$ is a nonzero multiple of $z_1^k z_2^{m-k}$. Therefore, W must also contain $z_1^k z_2^{m-k}$ for all $0 \le k \le m$. Since these elements form a basis for V_m, we see that $W = V_m$, as desired. □

4.4 Representations of $\mathsf{sl}(2;\mathbb{C})$

In this section, we will compute (up to isomorphism) all the finite-dimensional irreducible complex representations of the Lie algebra $\mathsf{sl}(2;\mathbb{C})$. This computation is important for several reasons. First, $\mathsf{sl}(2;\mathbb{C})$ is the complexification of $\mathsf{su}(2)$, which in turn is isomorphic to $\mathsf{so}(3)$, and the representations of $\mathsf{so}(3)$ are of physical significance. Indeed, the computation we will perform in the proof of Theorem 4.21 is found in every standard textbook on quantum mechanics, under the title "angular momentum." Second, the representation theory of $\mathsf{su}(2)$ is an illuminating example of how one uses commutation relations to determine the representations of a Lie algebra.

We use the basis (4.13) of $\mathsf{sl}(2;\mathbb{C})$, which have commutation relations

$$\begin{aligned}[][H,X] &= 2X\\ [H,Y] &= -2Y\\ [X,Y] &= H.\end{aligned}$$

If V is a finite-dimensional complex vector space and A, B, and C are operators on V satisfying $[A,B] = 2B$, $[A,C] = -2C$, and $[B,C] = A$, then by Lemma 3.17, the unique linear map $\pi : \mathsf{sl}(2;\mathbb{C}) \to \mathsf{gl}(V)$ satisfying

$$\pi(H) = A, \quad \pi(X) = B, \quad \pi(Y) = C$$

will be a representation of $\mathsf{sl}(2;\mathbb{C})$.

Theorem 4.21. *For each integer $m \geq 0$, there is an irreducible complex representation of $\mathsf{sl}(2;\mathbb{C})$ with dimension $m+1$. Any two irreducible complex representations of $\mathsf{sl}(2;\mathbb{C})$ with the same dimension are isomorphic. If π is an irreducible complex representation of $\mathsf{sl}(2;\mathbb{C})$ with dimension $m+1$, then π is isomorphic to the representation π_m described in Example 4.19.*

Our goal is to show that any finite-dimensional irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ "looks like" one of the representations π_m coming from Example 4.19. In that example, the space V_m is spanned by eigenvectors for $\pi_m(H)$ and the operators $\pi_m(X)$ and $\pi_m(Y)$ act by shifting the eigenvalues up or down in increments of 2. We now introduce a simple but crucial lemma that allows us to develop a similar structure in an arbitrary irreducible representation of $\mathsf{sl}(2;\mathbb{C})$.

Lemma 4.22. *Let $\pi : \mathsf{sl}(2;\mathbb{C}) \to \mathsf{gl}(V)$ be a representation of $\mathsf{sl}(2;\mathbb{C})$ acting on a complex vector space V and let u be an eigenvector of $\pi(H)$ with eigenvalue $\alpha \in \mathbb{C}$. Then we have*

$$\pi(H)\pi(X)u = (\alpha+2)\pi(X)u$$

Thus, either $\pi(X)u = 0$ or $\pi(X)u$ is an eigenvector for $\pi(H)$ with eigenvalue $\alpha+2$. Similarly,

$$\pi(H)\pi(Y)u = (\alpha-2)\pi(Y)u$$

so that either $\pi(Y)u = 0$ or $\pi(Y)u$ is an eigenvector for $\pi(H)$ with eigenvalue $\alpha - 2$.

Proof. We know that $[\pi(H), \pi(X)] = \pi([H, X]) = 2\pi(X)$. Thus

$$\begin{aligned}\pi(H)\pi(X)u &= \pi(X)\pi(H)u + 2\pi(X)u \\ &= \pi(X)(\alpha u) + 2\pi(X)u \\ &= (\alpha + 2)\pi(X)u.\end{aligned}$$

The argument with $\pi(X)$ replaced by $\pi(Y)$ is similar. □

Proof of Theorem 4.21. Let π be an irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ acting on a finite-dimensional complex vector space V. Our strategy is to diagonalize the operator $\pi(H)$. Since we area working over $\mathbb{C}$, the operator $\pi(H)$ must have at least one eigenvector. Let u be an eigenvector for $\pi(H)$ with eigenvalue α. Applying Lemma 4.22 repeatedly, we see that

$$\pi(H)\pi(X)^k u = (\alpha + 2k)\pi(X)^k u.$$

Since an operator on a finite-dimensional space can have only finitely many eigenvalues, the $\pi(X)^k u$'s cannot all be nonzero. Thus, there is some $N \geq 0$ such that

$$\pi(X)^N u \neq 0$$

but

$$\pi(X)^{N+1} u = 0.$$

If we set $u_0 = \pi(X)^N u$ and $\lambda = \alpha + 2N$, then,

$$\pi(H)u_0 = \lambda u_0, \tag{4.15}$$

$$\pi(X)u_0 = 0. \tag{4.16}$$

Let us define

$$u_k = \pi(Y)^k u_0$$

for $k \geq 0$. By Lemma 4.22, we have

$$\pi(H)u_k = (\lambda - 2k)u_k.$$

Now, it is easily verified by induction that

$$\pi(X)u_k = k[\lambda - (k-1)]u_{k-1}, \quad k \geq 1. \tag{4.17}$$

Furthermore, since $\pi(H)$ can have only finitely many eigenvalues, the u_k's cannot all be nonzero. There must, therefore, be a non-negative integer m such that, for all $k \leq m$,

$$u_k = \pi(Y)^k u_0 \neq 0$$

but

$$u_{m+1} = \pi(Y)^{m+1} u_0 = 0.$$

If $u_{m+1} = 0$, then $\pi(X)u_{m+1} = 0$ and so, by (4.17),

$$0 = \pi(X)u_{m+1} = (m+1)(\lambda - m)u_m.$$

Since u_m and $m+1$ are nonzero, we must have $\lambda - m = 0$. Thus, λ must coincide with the non-negative integer m.

Thus, for every irreducible representation (π, V), there exists an integer $m \geq 0$ and nonzero vectors $u_0, \ldots, u_m$ such that

$$\begin{aligned} \pi(H)u_k &= (m-2k)u_k \\ \pi(Y)u_k &= \begin{cases} u_{k+1} & \text{if } k < m \\ 0 & \text{if } k = m \end{cases} \\ \pi(X)u_k &= \begin{cases} k(m-(k-1))u_{k-1} & \text{if } k > 0 \\ 0 & \text{if } k = 0. \end{cases} \end{aligned} \tag{4.18}$$

The vectors must be linearly independent, since they are eigenvectors of $\pi(H)$ with distinct eigenvalues (see, for example, 5.10 of [Axl]). Moreover, the $(m+1)$-dimensional span of $u_0, \ldots, u_m$ is explicitly invariant under $\pi(H)$, $\pi(X)$, and $\pi(Y)$ and, hence, under $\pi(Z)$ for all $Z \in \mathsf{sl}(2;\mathbb{C})$. Since π is irreducible, this space must be all of V.

We have shown that every irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ is of the form (4.18). Conversely, if we *define* $\pi(H)$, $\pi(X)$, and $\pi(Y)$ by (4.18) (where the u_k's are basis elements for some $(m+1)$-dimensional vector space), it is not hard to check that operators defined as in (4.18) really do satisfy the $\mathsf{sl}(2;\mathbb{C})$ commutation relations. Furthermore, we may prove irreducibility of this representation in the same way as in the proof of Proposition 4.20. In any case, Proposition 4.20 already showed that $\mathsf{sl}(2;\mathbb{C})$ has an $(m+1)$-dimensional irreducible representation, namely, π_m.

The preceding analysis shows that every irreducible representation of dimension $m+1$ must have the form in (4.18), which shows that any two such representations are isomorphic. In particular, the $(m+1)$-dimensional representation π_m described in Example 4.19 must be isomorphic to (4.18).

This completes the proof of Theorem 4.21. □

In particular, Theorem 4.21 says that every irreducible complex representation of $\mathsf{sl}(2;\mathbb{C})$ comes from a representation of $\mathsf{su}(2)$. By Proposition 4.17, this supposes the classification of the finite-dimensional complex representations of $\mathsf{su}(2)$.

4.5 Group Versus Lie Algebra Representations

We know from Chapter 3 (Theorem 3.29) that every Lie group homomorphism gives rise to a Lie algebra homomorphism. In particular, every representation of a matrix Lie group gives rise to a representation of the associated Lie algebra. In Chapter 5, we will give a partial converse to this result: If G is a *simply connected* matrix Lie group with Lie algebra $\mathfrak{g}$, then every representation of $\mathfrak{g}$ comes from a representation of G. (See Theorem 5.1.) Thus, for a simply connected matrix Lie group G, the Lie functor induces a one-to-one correspondence between the representations of G and the representations of $\mathfrak{g}$. The precise statement of this result is given in Corollary 5.3.

It is instructive to see how this general theory works out in the case of $\mathsf{SU}(2)$ (which is simply connected) and $\mathsf{SO}(3)$ (which is not). For every irreducible complex representation π of $\mathsf{su}(2)$, the complex-linear extension of π to $\mathsf{sl}(2;\mathbb{C})$ must be isomorphic (Theorem 4.21) to one of the representations π_m described in Example 4.19. Since those representations are constructed from representations of the group $\mathsf{SU}(2)$, we can see directly (without appealing to Theorem 5.1) that every irreducible complex representation of $\mathsf{su}(2)$ comes from a representation of $\mathsf{SU}(2)$. (Where we are using the fact that a representation of a real Lie algebra is irreducible if and only if the complex-linear extended representation to its complexification is irreducible, Proposition 4.17.) Since $\mathsf{SU}(2) \cong S^3$ is connected, by Point 1 of Proposition 4.16, this classifies all finite-dimensional complex irreducible representations of $\mathsf{SU}(2)$.

Now, by Example 3.18, there is a Lie algebra isomorphism $\phi : \mathsf{su}(2) \longrightarrow \mathsf{so}(3)$ such that $\phi(E_j) = F_j$, $j = 1, 2, 3$, where $\{E_1, E_2, E_3\}$ and $\{F_1, F_2, F_3\}$ are the bases listed in the example. Thus, the irreducible complex representations of $\mathsf{so}(3)$ are precisely of the form $\sigma_m = \pi_m \circ \phi^{-1}$. We wish to determine, for a particular m, whether or not there is a representation Σ_m of the group $\mathsf{SO}(3)$ such that $\Sigma_m(e^X) = e^{\sigma_m(X)}$ for all $X \in \mathsf{so}(3)$. Since $\mathsf{SO}(3) \cong \mathbb{RP}^3$ is connected, by Point 1 of Proposition 4.16, all such Σ_m are *all* the finite-dimensional irreducible complex representations of $\mathsf{SO}(3)$. This supposes the classification of such representations of $\mathsf{SO}(3)$.

Proposition 4.23. *Let $\sigma_m = \pi_m \circ \phi^{-1}$ be an irreducible complex representation of the Lie algebra $\mathsf{so}(3)$ ($m \geq 0$). If m is even, there is a representation Σ_m of the group $\mathsf{SO}(3)$ such that $\Sigma_m(e^X) = e^{\sigma_m(X)}$ for all $X \in \mathsf{so}(3)$ (that is, such that $\operatorname{Lie}\Sigma_m = \sigma_m$). If m is odd, there is no such representation of $\mathsf{SO}(3)$.*

Note that the condition that m be even is equivalent to the condition that $\dim V_m = m + 1$ be odd. Thus, it is the odd-dimensional representations of the Lie algebra $\mathsf{so}(3)$ which come from group representations. In the physics literature, the representations of $\mathsf{su}(2) \cong \mathsf{so}(3)$ are labeled by the parameter $\ell = m/2$. In terms of this notation, a representation of $\mathsf{so}(3)$ comes from a representation of $\mathsf{SO}(3)$ if and only if ℓ is an integer. The representations with ℓ an integer are called "integer spin"; the others are called "half-integer spin."

In order to prove Proposition 4.23, we will need Lemma 4.27, which is just a mixture of lemmas 4.24 and 4.25.

Recall the review of quotient topology after Proposition 1.11. If $f : A \to B$ is a function between sets A and B, the **fibers** of f are the sets $f^{-1}(b) \subset A$, where $b \in B$.

Lemma 4.24 (Universal property of quotient spaces). *Let X be a topological space and let (X', p) be a quotient space of X. That is, X' is a topological space and $p : X \to X'$ is a quotient map (i.e., a surjective and strongly continuous map). Every continuous map $f : X \to Y$ which preserves the fibers of p (that is, f sends each fiber of p to a unique point in Y) factors through p by a unique continuous map $\hat{f} : X' \to Y$, $\hat{f} \circ p = f$. That is, there exists a unique continuous map $\hat{f}$ that makes the following diagram commute:*

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "p"'] & Y \\ X' \arrow[ur, dashed, "\hat{f}"'] & \end{tikzcd}$$

For a proof, see Theorem 22.2 of [Mun]. The condition "f preserves fibers of p" is the same as condition "f respects relation $\sim_p$," which means that $x_1 \sim_p x_2$ implies $f(x_1) = f(x_2)$. Lemma 4.24 is just the universal property of the quotient topology in disguise. On the statement, if we replace X' by $X/\sim$, where $\sim$ is some equivalence relation on X, and $p : X \to X'$ by the natural projection $X \to X/\sim$, we end up with the u.p. of quotient topology. In this dictionary, $\sim$ would correspond to $\sim_p$.

Lemma 4.25 (Universal property of group epimorphisms). *Let $\Phi : G \to G'$ be a group epimorphism between groups G and G'. Any group homomorphism $\Psi : G \to H$ which kills $\ker \Phi$ (that is, such that Ψ sends $\ker \Phi$ to $e \in H$) factors through Φ by a unique group homomorphism $\hat{\Psi} : G' \to H$, $\hat{\Psi} \circ \Phi = \Phi$. That is, there exists a unique group homomorphism $\hat{\Psi}$ that makes the following diagram commute:*

$$\begin{tikzcd} G \arrow[r, "\Psi"] \arrow[d, "\Phi"'] & H \\ G' \arrow[ur, dashed, "\hat{\Psi}"'] & \end{tikzcd}$$

Observe that the condition that Ψ kills $\ker \Phi$ is equivalent to the condition that Ψ preserves fibers of Φ. By elementary group theory, the cosets of $\ker \Phi$ in G are precisely the equivalence classes of $\sim_\Phi$ (see footnote 5 on p. 10 for the definition of $\sim_\Phi$), which in turn coincide with the fibers of Φ. Succinctly, $\Phi(x) = \Phi(y)$ if and only if $x \in y \ker \Phi$. Indeed, suppose Ψ were to kill $\ker \Phi$. Then if $x, y \in G$ were in the same fiber, we would have $\Phi(x) = \Phi(y)$, so $x \in y \ker \Phi$ and then $x = yz$ for some $z \in \ker \Phi$. Hence, $\Psi(x) = \Psi(yz) = \Psi(y)\Psi(z) = \Psi(y)$. That is, the whole fiber of x is mapped by Ψ to a same point. Conversely, if we now suppose that Ψ preserves fibers of Φ, in particular it would preserve $\Phi^{-1}(e_H) = \ker\Phi$, so that $\Psi(\ker\Phi)$ would be then a singleton, namely, $\{e_H\}$. That is, Ψ kills $\ker\Phi$.

Proof. Uniqueness is immediate, for if such an $\hat{\Psi}$ existed, it would satisfy

$$\Psi(x) = \hat{\Psi}(\Phi(x)), \quad x \in G. \tag{4.19}$$

That is the images of $\hat{\Psi}$ at all point of G' are determined by Φ and Ψ, since Φ is surjective.

For existence, we *define* $\hat{\Psi}$ as in (4.19). That is, for each $x' \in G'$ and because Φ is surjective, we can pick some $x \in \Phi^{-1}(x')$ and define $\hat{\Psi}(x') = \Psi(x)$. This map is well-defined, for the fibers of Φ are preserved by Ψ, and by construction, it fulfills $\hat{\Psi} \circ \Phi = \Psi$. Furthermore, it is a group homomorphism: If x' and y' be in G' and let $x \in \Phi^{-1}(x')$ and $y \in \Phi^{-1}(y')$, then

$$\begin{aligned}\hat{\Psi}(x'y') &= \hat{\Psi}(\Phi(x)\Phi(y)) = \hat{\Psi}(\Phi(xy)) = \Psi(xy)\\ &= \Psi(x)\Psi(y) = \hat{\Psi}(\Phi(x))\hat{\Psi}(\Phi(y)) = \hat{\Psi}(x')\hat{\Psi}(y').\end{aligned}$$

□

Definition 4.26. A **topological group** is a set G which is both a group and a topological space and such that these two structures satisfy a compatibility condition: the group operation $G \times G \longrightarrow G$ and the group inverse element map $(\cdot)^{-1} : G \longrightarrow G$ are both continuous.

A **topological group homomorphism** is a group homomorphism between topological groups that is also continuous.

Every Lie group is a topological group. Every Lie group homomorphism is a topological group homomorphism. In other words, there is a forgetful functor $\mathsf{LieGrp} \longrightarrow \mathsf{TopGrp}$ from the category of Lie groups to the category of topological groups.

Lemma 4.27 (Universal property of topological quotient groups). *Let G be a group and let (G', p) be a topological quotient group of G. That is, G' is a topological group and $p : G \longrightarrow G'$ is a group homomorphism and a quotient map. Any topological group homomorphism $\Psi : G \longrightarrow H$ which kills* $\ker\Phi$ *factors through Φ by a unique topological group homomorphism $\hat{\Psi} : G' \longrightarrow H$, $\hat{\Psi} \circ \Phi = \Phi$. That is, there exists a unique topological group homomorphism $\hat{\Psi}$ that makes the following diagram commute:*

$$\begin{tikzcd} G \arrow[r, "\Psi"] \arrow[d, "\Phi"'] & H \\ G' \arrow[ur, dashed, "\hat{\Psi}"'] & \end{tikzcd}$$

Proof. The proof is just the conjunction of lemmas 4.24 and 4.25. □

Proof of Proposition 4.23. We will be using notation of Example 3.18, so $\phi : \mathfrak{su}(2) \longrightarrow \mathfrak{so}(3)$ will be a Lie algebra homomorphism such that $\phi(E_j) = F_j$, $j = 1, 2, 3$, where $\{E_1, E_2, E_3\}$ and $\{F_1, F_2, F_3\}$ are the bases listed in the example. Like we did in Example 2.9, we may compute that

$$e^{2\pi F_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\pi & -\sin 2\pi \\ 0 & \sin 2\pi & \cos 2\pi \end{pmatrix} = I. \tag{4.20}$$

Meanwhile, $\sigma_m(F_1) = \pi_m\left(\phi^{-1}(F_1)\right) = \pi_m(E_1)$, with $E_1 = iH/2$, where, as usual, H is the diagonal matrix with diagonal entries $(1, -1)$. We know that there is a basis $u_0, u_1, \ldots, u_m$ for V_m such that u_k is an eigenvector for $\pi_m(H)$ with eigenvalue $m - 2j$. This means that u_j is an also eigenvector for $\sigma_m(F_1) = i\pi_m(H)/2$, with eigenvalue $i(m - 2j)/2$. For that reason, in the basis $\{u_j\}$, we have

$$\sigma_m(F_1) = \begin{pmatrix} \frac{i}{2}m & & & \\ & \frac{i}{2}(m-2) & & \\ & & \ddots & \\ & & & \frac{i}{2}(-m) \end{pmatrix}. \tag{4.21}$$

Suppose first that m were odd and that there exists a representation Σ_m of $\mathsf{SO}(3)$ with $\operatorname{Lie}\Sigma_m = \sigma_m$. Then $m - 2j$ would be an odd integer and, thus, $e^{2\pi\sigma_m(F_1)}$ would have eigenvalues $e^{2\pi i(m-2j)/2} = -1$ in the basis $\{u_j\}$, showing that $e^{2\pi\sigma_m(F_1)} = -I$. We have achieved a contradiction, since by (4.20),

$$I = \Sigma_m(I) = \Sigma_m(e^{2\pi F_1}) = e^{2\pi\sigma_m(F_1)} = -I.$$

Therefore, if m is odd, it cannot exist a representation Σ_m of $\mathsf{SO}(3)$ that yields *the* $(m + 1)$-dimensional representation σ_m of $\mathfrak{so}(3)$, $\operatorname{Lie}\Sigma_m = \sigma_m$.

Suppose now that m were even and let's look for a representation Σ_m of $\mathsf{SO}(3)$ with $\Sigma_m = \operatorname{Lie}\sigma_m$. Recall from Example 3.30 that the Lie algebra isomorphism ϕ comes from the surjective group homomorphism Φ in Proposition 1.11, where $\ker\Phi = \{I, -I\}$. Let Π_m be the representation of $\mathsf{SU}(2)$ in Example 4.19. Now, $e^{2\pi E_1} = -I$, and, thus,

$$\Pi_m(-I) = \Pi_m(e^{2\pi E_1}) = e^{\pi_m(2\pi E_1)}.$$

By (4.21) and for $\pi_m(E_1) = \sigma_m(F_1)$, the matrix $e^{\pi_m(2\pi E_1)}$ is diagonal in the basis $\{u_j\}$, so since m is even, its eigenvalues are $e^{2\pi i(m-2j)/2} = 1$, showing that $\Pi(-I) = e^{\pi_m(2\pi E_1)} = I$. This means that Π_m kills $\ker\Phi$. The map Φ is a quotient map (this was reasoned in proof of Corollary 1.12). Thus, by Lemma 4.27, there exists a continuous group homomorphism $\Sigma_m : \mathsf{SO}(3) \longrightarrow \mathsf{GL}(V_m)$ such that $\Pi_m = \Sigma_m \circ \Phi$. Equivalently, such that the following diagram commutes:

$$\begin{tikzcd} \mathsf{SU}(2) \arrow[r, "\Pi_m"] \arrow[d, "\Phi"'] & \mathsf{GL}(V_m) \\ \mathsf{SO}(3) \arrow[ur, dashed, "\Sigma_m"'] & \end{tikzcd}$$

(Recall that continuity implies C^∞-differentiability for group homomorphisms between matrix Lie groups, Corollary 3.53.)

By functoriality, $\pi_m = \operatorname{Lie}\Sigma_m \circ \phi$, so that $\operatorname{Lie}\Sigma_m = \pi_m \circ \phi^{-1} = \sigma_m$, showing that Σ_m is the desired representation of $\mathsf{SO}(3)$. ☐

5 | The Lie Group–Lie Algebra Correspondence

5.1 The "Hard" Questions

Consider three elementary results from the preceding chapters of this thesis: (1) Every matrix Lie group G has a Lie algebra $\mathfrak{g}$. (2) A continuous group homomorphism Φ between matrix Lie groups G and H gives rise to a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$. (3) If G and H are matrix Lie groups and H is a subgroup of G, then the Lie algebra $\mathfrak{h}$ of H is a subalgebra of the Lie algebra $\mathfrak{g}$ of G. Observe that (1) and (2) are condensed in the assertion "there exists a functor $\mathsf{MtxLieGrp} \to \mathsf{LieAlg}_{\mathbb{R}}$." Each of these results goes in the "easy" direction, from a group notion to an associated Lie algebra notion: in the direction of the Lie functor. In this chapter, we attempt to go in the "hard" direction, from the Lie algebra to the Lie group: in the opposite direction of the Lie functor. We will address three questions relating to the preceding three theorems.

- **Question 1**: Is every finite-dimensional, real Lie algebra the Lie algebra of some matrix Lie group? In other words, is the Lie functor essentially surjective?
- **Question 2**: Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. Does there exist a Lie group homomorphism $\Phi : G \to H$ such that $\operatorname{Lie}\Phi = \phi$? In other words, is the Lie functor full?
- **Question 3**: If G is a matrix Lie group with Lie algebra $\mathfrak{g}$ and $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, is there a matrix Lie group $H \subset G$ whose Lie algebra is $\mathfrak{h}$?

The answer to Question 1 is yes; see Sect. 5.4. The answer to Question 2 is, in general, no, but yes if G is simply connected; see Sect. 5.2. The answer to Question 3 is no, in general, but is yes if we allow H to be a "connected Lie subgroup" that is not necessarily closed; see Sect. 5.3.

Throughout this chapter, we will be stating the theorems that answers the questions while skipping their proofs. The interested reader will be redirected to the corresponding proofs in [Hall1]. Instead, we will focus on the consequences of the theorems and its corollaries. Our point of view will be categorical, and it is the style in which we will be formulating the theorems and the rest of results.

The tool that is needed for proving these profound results is the Baker–Campbell–Hausdorff formula, also known as the BCH formula, for short. This formula expresses $\log(e^X e^Y)$, where X and Y are sufficiently small $n \times n$ matrices, in Lie-algebraic terms, that is, in terms of iterated commutators involving X and Y. The formula implies that all information about the product operation on a matrix Lie group, at least near the identity, is encoded in the Lie algebra. In its series form the formula may be stated the following way: there is a neighborhood of the zero matrix in $\mathfrak{gl}(n;\mathbb{C})$ such that for all matrices X and Y in this neighborhood,

$$\log(e^X e^Y) = X + Y + \tfrac{1}{2}[X,Y] + \tfrac{1}{12}[X,[X,Y]] - \tfrac{1}{12}[Y,[X,Y]] + \cdots ,$$

where the "$\cdots$" refers to additional terms involving iterated brackets of X and Y.

To see a precise statement of the formula and its proof, see sections 5.3, 5.4 and 5.5 of [Hall1].

5.2 Group Versus Lie Algebra Homomorphisms

The Lie functor assigns, to each Lie group homomorphism $\Phi : G \to H$ between matrix Lie groups, the induced Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$ between their Lie algebras. We now state a partial converse of this result, and study its consequences in categorical terms. Recall Definition 4.13 for the concept of a fully faithful functor for morphisms departing a subclass of objects of the domain category.

Theorem 5.1. *The Lie functor is fully faithful for morphisms departing simply connected matrix Lie groups.*

Since simply connected topological spaces are in particular connected spaces, Corollary 3.52 tells us that the Lie functor is faithful for morphisms departing simply connected matrix Lie groups. That makes half of Theorem 5.1. The fullness part is exactly what Theorem 5.6 of [Hall1] states: if $\mathfrak{g}$ and $\mathfrak{h}$ are the Lie algebras of some matrix Lie groups G and H, where G is simply connected, and $\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, then there exists a Lie group homomorphism $\Phi : G \to H$ with $\operatorname{Lie}\Phi = \phi$. This last result is a profound one and we redirect to Hall's book for the proof. The proof idea is the following: at first, one starts with the Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$ and uses the BCH formula to construct a *local homomorphism* from G into H. Then one extends the local homomorphism to a global homomorphism—i.e., to an ordinary Lie group homomorphism—and exploits the simple connectedness of G to prove that the global homomorphism is well-defined. Lastly, one checks that the Lie functor sends this global homomorphism $G \to H$ to the original Lie algebra homomorphism ϕ.

Theorem 5.1 has lots of applications. In particular, it implies that the restriction[1] of the Lie functor to the full subcategory of simply connected matrix Lie groups, denoted $\mathsf{MtxLieGrp}_{\text{simpl}}$, is a fully faithful functor,

$$\mathrm{Lie} : \mathsf{MtxLieGrp}_{\text{simpl}} \to \mathsf{FinLieAlg}_{\mathbb{R}}. \tag{5.1}$$

We deduce two results, the first of which is a partial converse of Corollary 3.32.

Corollary 5.2. *If G and H and simply connected Lie groups with isomorphic Lie algebras, then $G \cong H$.*

Proof. Since the functor (5.1) is fully faithful, it creates isomorphisms by Lemma 4.14. □

Let C and D be categories. We say that C and D are **isomorphic**, and we write $\mathsf{C} \cong \mathsf{D}$, if there exists a pair of covariant functors $F : \mathsf{C} \to \mathsf{D}$ and $G : \mathsf{D} \to \mathsf{C}$ such that $GF = 1_{\mathsf{C}}$ and $FG = 1_{\mathsf{D}}$, where 1_{C} and 1_{D} are the identity functors on C and D, respectively. In that case, F and G will be called **isomorphisms of categories**. It is not difficult to check that a functor is an isomorphism of categories if and only if it is fully faithful and both injective and surjective on objects.

Corollary 5.3. *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. If G is simply connected, then the functor*

$$\mathsf{FinRep}_{\mathbb{K}}(G) \to \mathsf{FinRep}_{\mathbb{K}}(\mathfrak{g}) \tag{5.2}$$

of Proposition 4.12 is an isomorphism of categories.

Proof. Since simply connected implies connected, by Proposition 4.15, the functor (5.2) is a full embedding. It is left to show surjectivity on objects. Let $\pi : \mathfrak{g} \to \mathsf{gl}(V)$ be a representation of the Lie algebra $\mathfrak{g}$ on some finite-dimensional vector space V. The map π is a Lie algebra homomorphism and, hence, since G is simply connected, by Theorem 5.1, we have that there exists a Lie group homomorphism $\Pi : G \to \mathsf{GL}(V)$ such that $\mathrm{Lie}\,\Pi = \pi$. That is, the functor (5.2) sends the representation (V, Π) of G to the representation (V, π) of $\mathfrak{g}$. □

This corollary says that to study the finite-dimensional real (resp., complex) representations of a simply connected matrix Lie group is *the same* as to study the finite-dimensional real (resp., complex) representations of its Lie algebra.

[1] In general, the restriction of a functor $F : \mathsf{C} \to \mathsf{D}$ to *any* subcategory $\mathsf{E} \subset \mathsf{C}$ is a functor as well, $F|_{\mathsf{E}} : \mathsf{E} \to \mathsf{D}$.

5.3 Subgroups and Subalgebras

In this section, we address Question 3 from Sect. 5.1: If G is a matrix Lie group with Lie algebra $\mathfrak{g}$ and $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, does there exist a matrix Lie group $H \subset G$ whose Lie algebra is H?

The answer to Question 3, as stated, is no. Suppose, for example, that $G = \mathsf{GL}(2; \mathbb{C})$ and

$$\mathfrak{h} = \left\{ \left. \begin{pmatrix} it & 0 \\ 0 & ita \end{pmatrix} \right| t \in \mathbb{R} \right\}, \tag{5.3}$$

where a is irrational. If there is going to be a matrix Lie group H with Lie algebra $\mathfrak{h}$, then H would have to contain the closure of the group

$$H_0 = \left\{ \left. \begin{pmatrix} e^{it} & 0 \\ 0 & e^{ita} \end{pmatrix} \right| t \in \mathbb{R} \right\}, \tag{5.4}$$

which is (Exercise 10 in Chapter 1 of [Hall1]) the group

$$H_1 = \left\{ \left. \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\phi} \end{pmatrix} \right| \theta, \phi \in \mathbb{R} \right\}.$$

But the Lie algebra of H would have to contain the Lie algebra of H_1, which is two dimensional![2]

Fortunately, all is not lost. We can still get a group H for each subalgebra $\mathfrak{h}$ if we weaken the condition that H be a matrix Lie group. In the above example, the subgroup we want is H_0, despite the fact that H_0 is not closed.

Definition 5.4. If H is any subgroup of $\mathsf{GL}(n; \mathbb{C})$, the **Lie algebra** $\mathfrak{h}$ of H is the set of all matrices X such that

$$e^{tX} \in H$$

for all real t.

It is possible to prove that for *any* subgroup H of $\mathsf{GL}(n; \mathbb{C})$, the Lie algebra $\mathfrak{h}$ of H is actually a Lie algebra, that is, a real vector space—possibly zero-dimensional—and closed under brackets. (See Proposition 1 and Corollary 7 in Chapter 2 of [Ross].) This result is not, however, directly relevant this section.

Definition 5.5. If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, then $H \subset G$ is a **connected Lie subgroup** of G if the following conditions are satisfied:

[2] The two dimensionality can be seen using the fact that, since H_1 is a matrix Lie group, Corollary 3.47 says that there is a neighborhood of the identity in H_1 which is mapped by log to a neighborhood of zero in $\mathfrak{h}_1 = \mathrm{Lie}\, H_1$. Thus, there exists a neighborhood of zero in $\mathfrak{h}_1$ in which all matrices are of the form $\begin{pmatrix} i\theta & 0 \\ 0 & i\phi \end{pmatrix}$, with $(\theta, \phi) \in U \subset \mathbb{R}^2$, where U is some open neighborhood of $(0, 0)$.

1. H is a subgroup of G.
2. The Lie algebra $\mathfrak{h}$ of H is a Lie subalgebra of $\mathfrak{g}$.
3. Every element of H can be written in the form $e^{X_1}e^{X_2}\cdots e^{X_m}$, with $X_1, X_2, \ldots, X_m \in \mathfrak{h}$.

Note that any group H as in the definition is path connected, since each element of H can be connected to the identity in H by a path of the form

$$t \mapsto e^{(1-t)X_1}e^{(1-t)X_2}\cdots e^{(1-t)X_m}.$$

Theorem 5.6. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$ and let $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{g}$. Then there exists a unique connected Lie subgroup H of G with Lie algebra $\mathfrak{h}$. Namely, $H = \{e^{X_1}e^{X_2}\cdots e^{X_m} \mid X_1, X_2, \ldots, X_m \in \mathfrak{h}\}$.*

Observe that, in any case, $H = \{e^{X_1}e^{X_2}\cdots e^{X_m} \mid X_1, X_2, \ldots, X_m \in \mathfrak{h}\}$ is always a subgroup of $G \subset \mathsf{GL}(n;\mathbb{C})$. For that reason, the only thing that must be proven in Theorem 5.6 is that the Lie algebra of H is $\mathfrak{h}$. For a proof, see Theorem 5.20 of [Hall1]; this proof uses the BCH formula. From the theorem we deduce that if $\mathfrak{h}$ is the subalgebra of $\mathfrak{gl}(2;\mathbb{C})$ in (5.3), then the connected Lie subgroup H is the group H_0 in (5.3), which is not closed.

If a connected Lie subgroup H of $\mathsf{GL}(n;\mathbb{C})$ is not closed, the topology H inherits from $\mathsf{GL}(n;\mathbb{C})$ may be pathological, e.g., not locally connected. Nevertheless, we can give H a new topology that is much nicer.

Theorem 5.7. *Let H be a connected Lie subgroup of $\mathsf{GL}(n;\mathbb{C})$ with Lie algebra $\mathfrak{h}$. Then H can be given the structure of a smooth manifold in such a way that the group operations on H are smooth and the inclusion map of H into $\mathsf{GL}(n;\mathbb{C})$ is smooth.*

Thus, every connected Lie subgroup of $\mathsf{GL}(n;\mathbb{C})$ can be made into a Lie group. In the case of the group H_0 in (5.3), the new topology on H_0 is obtained by identifying H_0 with $\mathbb{R}$ by means of the parameter t in the definition of H_0.

For a proof of Theorem 5.7, see Theorem 5.23 of [Hall1]. The proof of Hall's book is not a completely detailed one, but using the smooth manifold chart lemma (Lemma 1.35 of [Lee]) one can give the details left in the proof.

5.4 Lie's Third Theorem

Lie's third theorem (in its modern, global form) says that for every finite-dimensional, real Lie algebra $\mathfrak{g}$, there exists a Lie group G with Lie algebra $\mathfrak{g}$. We will take this G to be a connected Lie subgroup of $\mathsf{GL}(n;\mathbb{C})$.

Theorem 5.8. *If $\mathfrak{g}$ is any finite-dimensional, real Lie algebra, there exists a connected Lie subgroup G of $\mathsf{GL}(n;\mathbb{C})$ whose Lie algebra is isomorphic to $\mathfrak{g}$.*

Our proof assumes Ado's theorem, which asserts that every finite-dimensional real or complex Lie algebra is isomorphic to an algebra of matrices. (See, for example, Theorem 3.17.17 in [Var].)

Proof. By Ado's theorem, we may identify $\mathfrak{g}$ with a real subalgebra of $\mathsf{gl}(n;\mathbb{C})$. Then, by Theorem 5.6, there is a connected Lie subgroup of $\mathsf{GL}(n;\mathbb{C})$ with Lie algebra $\mathfrak{g}$. □

It is actually possible to choose the subgroup G in Theorem 5.8 to be closed. Indeed, according to Theorem 9 on p. 105 of [Got], if a connected Lie group G can be embedded into some $\mathsf{GL}(n;\mathbb{C})$ as a connected Lie subgroup, then G can be embedded into some other $\mathsf{GL}(n';\mathbb{C})$ as a closed subgroup. Assuming this result, we may reach the following corollary.

Corollary 5.9. *The Lie functor is essentially surjective.*

Bibliography

[Axl] S. Axler. *Linear Algebra Done Right*. 3rd ed. Undergraduate Texts in Mathematics. Springer International Publishing, 2015.

[Boas] Harold P. Boas. *Math 650-600: Several Complex Variables*. Class Handout, Texas A&M University. Jan. 2005. URL: https://www.math.tamu.edu/~boas/courses/650-2005a/jan20.pdf.

[Got] M. Gotô. "Faithful Representations of Lie Groups II". In: *Nagoya Mathematical Journal* 1 (1950), pp. 91–107.

[Hall1] B. C. Hall. *Lie groups, Lie algebras, and Representations: An Elementary Introduction*. Vol. 222. Springer, 2015.

[Hall2] B. C. Hall. *Quantum Theory for Mathematicians*. Graduate Texts in Mathematics. Springer New York, 2013.

[Lee] J. M. Lee. *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer, 2003.

[Mun] J. Munkres. *Topology*. 2nd ed. Prentice Hall, 2000.

[Rieh] E. Riehl. *Category Theory in Context*. Aurora: Dover Modern Math Originals. Dover Publications, 2017. A free (and legal!) digital version of the book is available in the author's webpage: https://math.jhu.edu/~eriehl/context.pdf.

[Ross] W. Rossmann. *Lie Groups: An Introduction Through Linear Groups*. Oxford graduate texts in mathematics. Oxford University Press, 2006.

[Rud1] W. Rudin. *Principles of Mathematical Analysis, Third Edition*. 3rd. McGraw-Hill Science Engineering Math, 1976.

[Rud2] W. Rudin. *Real and complex analysis*. 3rd. McGraw-Hill, 1987.

[Var] V. S. Varadarajan. *Lie groups, lie algebras and their representations*. Graduate texts in mathematics. Springer-Verlag, 1974.