An introduction to matrix groups and their applications

Andrew Baker

[14/7/2000]

Department of Mathematics, University of Glasgow, Glasgow G12 8QW, Scotland.
E-mail address: a.baker@maths.gla.ac.uk
URL: http://www.maths.gla.ac.uk/∼ajb

Introduction

These notes are intended to provide a first view of Lie theory accessible to mathematics undergraduates. Although the prerequisites are kept as low level as possible, the material is sophisticated and contains many of the key themes of the mature subject. In order to achieve this we confine ourselves to *matrix groups*, i.e., closed subgroups of general linear groups. One of the main results that we prove shows that every matrix group is in fact a Lie subgroup, the proof being modelled on that in the expository paper of Howe [**5**]. Indeed the latter paper together with the book of Curtis [**4**] played a central part in setting our goals for the course based on these notes.

Of course, the classical Lie groups are easily introduced at undergraduate level, and it is possible to discuss many of their features. The spinor groups are also introduced and through them the rôle of global topology.

In Chapter 1 the *general linear groups* $\mathrm{GL}_n(\Bbbk)$ where $\Bbbk = \mathbb{R}$, the real numbers, or $\Bbbk = \mathbb{C}$, the complex numbers, are introduced and studied as both groups and topological spaces. Matrix groups are defined and a number of standard examples are discussed, including the unimodular groups $\mathrm{SL}_n(\Bbbk)$, orthogonal $\mathrm{O}(n)$ and special orthogonal groups $\mathrm{SO}(n)$, unitary $\mathrm{U}(n)$ and special unitary groups $\mathrm{SU}(n)$, as well as more exotic examples such as Lorentz groups and symplectic groups. The relation of complex to real matrix groups is also studied and finally the exponential map for the general linear groups is introduced.

In Chapter 2 the *Lie algebra* of a matrix group is defined. The special cases of $\mathrm{SU}(2)$ and $\mathrm{SL}_2(\mathbb{C})$ and their relationships with $\mathrm{SO}(3)$ and the Lorentz group are studied in detail.

In Chapter 3 the units in a finite dimensional algebra over $\mathbb{R}$ or $\mathbb{C}$ are studied as a source of matrix groups using the reduced regular representation. The quaternions and more generally the real Clifford algebras are defined and spinor groups constructed and shown to double cover the special orthogonal groups. The quaternionic symplectic groups $\mathrm{Sp}(n)$ are also defined, thus completing the list of compact classical groups and their universal covers.

In Chapter 4 we define the idea of a Lie group and show that all matrix groups are Lie subgroups of general linear groups.

In Chapter 5 we discuss homeogeneous spaces and show how to recognise them as orbits of smooth actions. Then in Chapter 6 we discuss connectivity of Lie groups and use homogeneous spaces to prove that many familiar Lie groups connected.

In Chapter 7 the basic theory of compact connected Lie groups and their maximal tori is studied and the relationship to well known diagonalisation results highlighted.

I would like to thank the Universität Bern for inviting me to visit and teach a course in the spring of 2000; particular thanks go to the students who spotted numerous errors and obscurities and Z. Balogh who helped with the problem classes. Thanks also the mathematicians of Glasgow, especially R. Odoni. Finally, many thanks to the topologists and fellow travellers of Manchester University from whom I learnt much of my mathematics.

Contents

Introduction 3

Chapter 1. Real and complex matrix groups 1
1. Groups of matrices 1
2. Groups of matrices as metric spaces 1
3. Matrix groups 5
4. Some examples of matrix groups 7
5. Complex matrix groups as real matrix groups 10
6. Continuous homomorphisms of matrix groups 11
7. Continuous group actions 12
8. The matrix exponential and logarithm functions 13

Chapter 2. Lie algebras for matrix groups 17
1. Differential equations in matrices 17
2. One parameter subgroups 18
3. Curves, tangent spaces and Lie algebras 19
4. Some Lie algebras of matrix groups 23
5. SO(3) and SU(2) 28
6. $\mathrm{SL}_2(\mathbb{C})$ and the Lorentz group 32

Chapter 3. Quaternions, Clifford algebras and some associated groups 37
1. Algebras 37
2. Linear algebra over a division algebra 39
3. Quaternions 41
4. Quaternionic matrix groups 44
5. The real Clifford algebras 45
6. The spinor groups 49
7. The centres of spinor groups 52
8. Finite subgroups of spinor groups 53

Chapter 4. Matrix groups as Lie groups 55
1. Smooth manifolds 55
2. Tangent spaces and derivatives 55
3. Lie groups 58
4. Some examples of Lie groups 59
5. Some useful formula in matrix groups 62
6. Matrix groups are Lie groups 66
7. Not all Lie groups are matrix groups 69

Chapter 5. Homogeneous spaces 73

1. Homogeneous spaces as manifolds 73
2. Homogeneous spaces as orbits 74
3. Projective spaces 76
4. Grassmannians 79

Chapter 6. Connectivity of matrix groups 81
1. Connectivity of manifolds 81
2. Examples of path connected matrix groups 82
3. The path components of a Lie group 84
4. Another connectivity result 86

Chapter 7. Compact connected Lie groups and their maximal tori 89
1. Tori 89
2. Maximal tori in compact Lie groups 91
3. The normalizer and Weyl group of a maximal torus 93

Bibliography 97

Index 99

Problem sets 1
Problems on Chapter 1 1
Problems on Chapter 2 4
Problems on Chapter 3 6
Problems on Chapter 4 8
Problems on Chapters 5 and 6 9
Problems on Chapter 7 10

CHAPTER 1

Real and complex matrix groups

1. Groups of matrices

In these notes we will usually consider the cases of the fields $\Bbbk = \mathbb{R}$, the real numbers, and $\Bbbk = \mathbb{C}$, the complex numbers. However, the general framework of this section is applicable for any (commutative) field $\Bbbk$. Actually, much of it applies to the case of a general division algebra, with the example of the *quaternions* discussed in Chapter 3 being of most interest to us.

Let $\mathrm{M}_{m,n}(\Bbbk)$ be the set of $m \times n$ matrices with entries in $\Bbbk$. We will denote (i, j) entry of an $m \times n$ matrix A by A_{ij} or a_{ij},

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

We set $\mathrm{M}_n(\Bbbk) = \mathrm{M}_{n,n}(\Bbbk)$. Then $\mathrm{M}_n(\Bbbk)$ is a (not usually commutative) ring under the usual addition and multiplication of matrices, with identity I_n. Recall the *determinant* function $\det\colon \mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk$.

Proposition 1.1. $\det\colon \mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk$ *has the following properties.*

a) *For* $A, B \in \mathrm{M}_n(\Bbbk)$, $\det(AB) = \det A \det B$.
b) $\det I_n = 1$.
c) $A \in \mathrm{M}_n(\Bbbk)$ *is invertible if and only if* $\det A \neq 0$.

We use the notation

$$\mathrm{GL}_n(\Bbbk) = \{A \in \mathrm{M}_n(\Bbbk) : \det A \neq 0\}$$

for the set of invertible $n \times n$ matrices, and

$$\mathrm{SL}_n(\Bbbk) = \{A \in \mathrm{M}_n(\Bbbk) : \det A = 1\} \subseteq \mathrm{GL}_n(\Bbbk)$$

for the set of $n \times n$ *unimodular* matrices.

Theorem 1.2. *The sets* $\mathrm{GL}_n(\Bbbk)$, $\mathrm{SL}_n(\Bbbk)$ *are groups under matrix multiplication. Furthermore,* $\mathrm{SL}_n(\Bbbk)$ *is a subgroup of* $\mathrm{GL}_n(\Bbbk)$, *i.e.,* $\mathrm{SL}_n(\Bbbk) \leqslant \mathrm{GL}_n(\Bbbk)$.

$\mathrm{GL}_n(\Bbbk)$ is called the $n \times n$ *general linear group* , while $\mathrm{SL}_n(\Bbbk)$ is called the $n \times n$ *special linear* or *unimodular group.* When $\Bbbk = \mathbb{R}$ or $\Bbbk = \mathbb{C}$ we will refer to $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{C})$ as the real and complex general linear groups. Of course, we can also consider subgroups of these groups, but before doing so we consider the topology of $\mathrm{M}_n(\mathbb{R})$ and $\mathrm{M}_n(\mathbb{C})$.

2. Groups of matrices as metric spaces

In this section we assume that $\Bbbk = \mathbb{R}, \mathbb{C}$. We may view $\mathrm{M}_n(\Bbbk)$ as a vector space over $\Bbbk$ of dimension n^2. We will define a *norm* on $\mathrm{M}_n(\Bbbk)$ as follows. Let $\Bbbk^n$ be the set of $n \times 1$ matrices over $\Bbbk$, and for $\mathbf{x} \in \Bbbk^n$ let

$$|\mathbf{x}| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}, \quad \text{where} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

For $A \in \mathrm{M}_n(\Bbbk)$ consider the set

$$\mathcal{S}_A = \left\{ \frac{|A\mathbf{x}|}{|\mathbf{x}|} : \mathbf{0} \neq \mathbf{x} \in \Bbbk^n \right\}.$$

It turns out that $\mathcal{S}_A$ is bounded and so we can define the real number

$$\|A\| = \sup \mathcal{S}_A.$$

Putting

$$\mathcal{S}_A^1 = \left\{ \frac{|A\mathbf{x}|}{|\mathbf{x}|} : \mathbf{x} \in \Bbbk^n,\ |\mathbf{x}| = 1 \right\},$$

we have

$$\|A\| = \sup \mathcal{S}_A^1 = \max \mathcal{S}_A^1,$$

since $\{\mathbf{x} \in \Bbbk^n : |\mathbf{x}| = 1\}$ is compact.

Remark 1.3. The following gives a procedure for calculating $\|A\|$; it may be familiar from numerical linear algebra where it is also used.

All the eigenvalues of the positive hermitian matrix A^*A are non-negative real numbers, hence it has a largest non-negative real eigenvalue λ. Then

$$\|A\| = \sqrt{\lambda}.$$

In fact, for any unit eigenvector $\mathbf{v}$ of A^*A for the eigenvalue λ, $\|A\| = |A\mathbf{v}|$.

When A is real, $A^*A = A^TA$ is real positive symmetric and there are unit eigenvectors $\mathbf{w} \in \mathbb{R}^n \subseteq \mathbb{C}^n$ of A^*A for the eigenvalue λ for which $\|A\| = |A\mathbf{w}|$. In particular, this shows that $\|A\|$ is independent of whether A is viewed as a real or complex matrix.

Proposition 1.4. *$\|\ \|$ is a $\Bbbk$-norm on $\mathrm{M}_n(\Bbbk)$, i.e.,*

a) $\|tA\| = |t|\,\|A\|$ *for* $t \in \Bbbk$, $A \in \mathrm{M}_n(\Bbbk)$;
b) $\|AB\| \leqslant \|A\|\,\|B\|$ *for* $A, B \in \mathrm{M}_n(\Bbbk)$;
c) $\|A + B\| \leqslant \|A\| + \|B\|$ *for* $A, B \in \mathrm{M}_n(\Bbbk)$;
d) $\|A\| = 0$ *if and only if* $A = 0$.

This norm $\|\ \|$ is called the *operator* or *sup* (= *supremum*) *norm*. We define a metric ρ on $\mathrm{M}_n(\Bbbk)$ by

$$\rho(A, B) = \|A - B\|.$$

Associated to this metric is a natural topology on $\mathrm{M}_n(\Bbbk)$, which allows us to define continuous functions $\mathrm{M}_n(\Bbbk) \longrightarrow X$ into a topological space X.

For $A \in \mathrm{M}_n(\Bbbk)$ and $r > 0$, let

$$\mathrm{N}_{\mathrm{M}_n(\Bbbk)}(A; r) = \{B \in \mathrm{M}_n(\Bbbk) : \|B - A\| < r\},$$

which is the *open disc of radius* r in $\mathrm{M}_n(\Bbbk)$. Similarly if $Y \subseteq \mathrm{M}_n(\Bbbk)$ and $A \in Y$, set

$$\mathrm{N}_Y(A; r) = \{B \in Y : \|B - A\| < r\} = \mathrm{N}_{\mathrm{M}_n(\Bbbk)}(A; r) \cap Y.$$

Then a subset $V \subseteq Y$ is open in Y if and only if for every $A \in V$, there is a $\delta > 0$ such that $\mathrm{N}_Y(A; \delta) \subseteq V$.

Definition 1.5. Let $Y \subseteq \mathrm{M}_n(\Bbbk)$ and $(X, \mathcal{T})$ be a topological space. Then a function $f \colon Y \longrightarrow X$ is *continuous* or a *continuous map* if for every $A \in Y$ and $U \in \mathcal{T}$ such that $f(A) \in U$, there is a $\delta > 0$ for which

$$B \in \mathrm{N}_Y(A; \delta) \Longrightarrow f(B) \in U.$$

Equivalently, f is continuous if and only if for $U \in \mathcal{T}$, $f^{-1}U \subseteq Y$ is open in Y.

Recall that for a topological space $(X, \mathcal{T})$, a subset $W \subseteq X$ is closed if $X - W \subseteq X$ is open. Yet another alternative formulation of the definition of continuity is that f is continuous if and only if for every closed subset $W \subseteq X$, $f^{-1}W \subseteq Y$ is closed in Y.

In particular we may take $X = \Bbbk$ and $\mathcal{T}$ to be the natural metric space topology associated to the standard norm on $\Bbbk$ and consider continuous functions $Y \longrightarrow \Bbbk$.

PROPOSITION 1.6. *For $1 \leqslant r, s \leqslant n$, the coordinate function*

$$\mathrm{coord}_{rs}\colon \mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk; \quad \mathrm{coord}_{rs}(A) = A_{rs}$$

is continuous.

PROOF. For the standard unit basis vectors $\mathbf{e}_i$ $(1 \leqslant i \leqslant n)$ of $\Bbbk^n$, we have

$$\begin{aligned} |A_{rs}| &\leqslant \sqrt{\sum_{i=1}^{n} |A_{is}|^2} \\ &= \left| \sum_{i=1}^{n} A_{is}\mathbf{e}_i \right| \\ &= |A\mathbf{e}_s| \\ &\leqslant \|A\|. \end{aligned}$$

So for $A, A' \in \mathrm{M}_n(\Bbbk)$,

$$|A'_{rs} - A_{rs}| \leqslant \|A' - A\|.$$

Now given $A \in \mathrm{M}_n(\Bbbk)$ and $\varepsilon > 0$, $\|A' - A\| < \varepsilon$ implies $|A'_{rs} - A_{rs}| < \varepsilon$. This shows that the function coord_{rs} is continuous at every $A \in \mathrm{M}_n(\Bbbk)$. □

COROLLARY 1.7. *If $f\colon \Bbbk^{n^2} \longrightarrow \Bbbk$ is continuous, then the associated function*

$$F\colon \mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk; \quad F(A) = f((A_{ij})_{1\leqslant i,j\leqslant n}),$$

is continuous.

COROLLARY 1.8. *The determinant* $\det\colon \mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk$ *and trace* $\mathrm{tr}\colon \mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk$ *are continuous functions.*

PROOF. The determinant is the composite of the continuous function $\mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk^{n^2}$ (which identifies $\mathrm{M}_n(\Bbbk)$ with $\Bbbk^{n^2}$) and a polynomial function $\Bbbk^{n^2} \longrightarrow \Bbbk$ (which is also continuous). Similarly for the trace,

$$\mathrm{tr}\, A = \sum_{i=1}^{n} A_{ii}.$$

□

There is a sort of converse of these results.

PROPOSITION 1.9. *For $A \in \mathrm{M}_n(\Bbbk)$,*

$$\|A\| \leqslant \sum_{i,j=1}^{n} |A_{ij}|.$$

Proof. Let $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ with $|\mathbf{x}| = 1$. Then since each $|x_k| \leqslant 1$,

$$\begin{aligned}
|A\mathbf{x}| &= |x_1 A\mathbf{e}_1 + \cdots + x_n A\mathbf{e}_n| \\
&\leqslant |x_1 A\mathbf{e}_1| + \cdots + |x_n A\mathbf{e}_n| \\
&\leqslant |A\mathbf{e}_1| + \cdots + |A\mathbf{e}_n| \\
&\leqslant \sqrt{\sum_{i=1}^{n} A_{i1}^2} + \cdots + \sqrt{\sum_{i=1}^{n} A_{in}^2} \\
&\leqslant \sum_{i,j=1}^{n} |A_{ij}|.
\end{aligned}$$

Since this is true for all vectors $\mathbf{x}$ with $|\mathbf{x}| = 1$, by definition of $\|A\|$,

$$\|A\| \leqslant \sum_{i,j=1}^{n} |A_{ij}|.$$

□

In fact, $\mathrm{M}_n(\Bbbk)$ is *complete* with respect to the norm $\|\ \|$.

Definition 1.10. A sequence $\{A_r\}_{r\geqslant 0}$ for which the following holds is a *Cauchy sequence*.

- For every $\varepsilon > 0$, there is an N such that $r, s > N$ implies $\|A_r - A_s\| < \varepsilon$.

Theorem 1.11. *For $\Bbbk = \mathbb{R}, \mathbb{C}$, every Cauchy sequence $\{A_r\}_{r\geqslant 0}$ in $\mathrm{M}_n(\Bbbk)$ has a limit $\lim_{r\to\infty} A_r$. Furthermore,*

$$(\lim_{r\to\infty} A_r)_{ij} = \lim_{r\to\infty} (A_r)_{ij}.$$

Proof. By Proposition 1.6, the limit on the right hand side exists, so it is sufficient to show that the required matrix limit is the matrix A with

$$A_{ij} = \lim_{r\to\infty} (A_r)_{ij}.$$

The sequence $\{A_r - A\}_{r\geqslant 0}$ satisfies

$$\|A_r - A\| \leqslant \sum_{i,j=1}^{n} |(A_r)_{ij} - A_{ij}| \to 0$$

as $r \to \infty$, so by Proposition 1.9, $A_r \to A$. □

It can be shown that the metric topologies induced by $\|\ \|$ and the usual norm on $\Bbbk^{n^2}$ agree in the sense that they have the same open sets (actually this is true for any two norms on $\Bbbk^{n^2}$). We summarise this in a useful criterion whose proof is left as an exercise.

Proposition 1.12. *A function $F\colon \mathrm{M}_m(\Bbbk) \longrightarrow \mathrm{M}_n(\Bbbk)$ is continuous with respect to the norms $\|\ \|$ if and only if each of the component functions $F_{rs}\colon \mathrm{M}_m(\Bbbk) \longrightarrow \Bbbk$ is continuous.*

A function $f\colon \mathrm{M}_m(\Bbbk) \longrightarrow \Bbbk$ is continuous with respect to the norm $\|\ \|$ and the usual metric on $\Bbbk$ if and only if it is continuous when viewed as a function $\Bbbk^{m^2} \longrightarrow \Bbbk$.

We now consider the topology of some subsets of $\mathrm{M}_n(\Bbbk)$, in particular some groups of matrices.

Proposition 1.13. *If $\Bbbk = \mathbb{R}, \mathbb{C}$,*

a) $\mathrm{GL}_n(\Bbbk) \subseteq \mathrm{M}_n(\Bbbk)$ *is an open subset;*
b) $\mathrm{SL}_n(\Bbbk) \subseteq \mathrm{M}_n(\Bbbk)$ *is a closed subset.*

Proof. We have seen that the function $\det\colon \mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk$ is continuous. Then

$$\mathrm{GL}_n(\Bbbk) = \mathrm{M}_n(\Bbbk) - \det^{-1}\{0\},$$

which is open since $\{0\}$ is closed, hence (a) holds. Similarly,

$$\mathrm{SL}_n(\Bbbk) = \det^{-1}\{1\} \subseteq \mathrm{GL}_n(\Bbbk),$$

which is closed in $\mathrm{M}_n(\Bbbk)$ and $\mathrm{GL}_n(\Bbbk)$ since $\{1\}$ is closed in $\Bbbk$, so (b) is true. □

The addition and multiplication maps $\mathrm{add}, \mathrm{mult}\colon \mathrm{M}_n(\Bbbk) \times \mathrm{M}_n(\Bbbk) \longrightarrow \mathrm{M}_n(\Bbbk)$ are also continuous where we take the product metric space topology on the domain. Finally, the inverse map

$$\mathrm{inv}\colon \mathrm{GL}_n(\Bbbk) \longrightarrow \mathrm{GL}_n(\Bbbk); \quad \mathrm{inv}(A) = A^{-1},$$

is also continuous since each entry of A^{-1} has the form

$$\frac{\text{polynomial in } A_{ij}\text{'s}}{\det A}$$

which is a continuous function of the entries of A and so is a continuous function of A itself.

Definition 1.14. Let G be a topological space and view $G \times G$ as the product space (i.e., give it the product topology). Suppose that G is also a group with multiplication map $\mathrm{mult}\colon G \times G \longrightarrow G$ and inverse map $\mathrm{inv}\colon G \longrightarrow G$. Then G is a *topological group* if $\mathrm{mult}, \mathrm{inv}$ are continuous.

The most familiar examples are obtained from arbitrary groups G given discrete topologies. In particular all finite groups can be viewed this way.

Theorem 1.15. *For $\Bbbk = \mathbb{R}, \mathbb{C}$, each of the groups $\mathrm{GL}_n(\Bbbk), \mathrm{SL}_n(\Bbbk)$ is a topological group with the evident multiplication and inverse maps and the subspace topologies inherited from $\mathrm{M}_n(\Bbbk)$.*

3. Matrix groups

Definition 1.16. A subgroup $G \leqslant \mathrm{GL}_n(\Bbbk)$ which is also a closed subspace is called a *matrix group over* $\Bbbk$ or a $\Bbbk$-*matrix group*. If we wish to make the value of n explicit, we say that G is a *matrix subgroup of* $\mathrm{GL}_n(\Bbbk)$.

Before considering some examples and properties, we record the following useful fact.

Proposition 1.17. *Let $G \leqslant \mathrm{GL}_n(\Bbbk)$ be a matrix subgroup and $H \leqslant G$ a closed subgroup of G. Then $H \leqslant \mathrm{GL}_n(\Bbbk)$ is a matrix subgroup.*

Proof. Every sequence $\{A_n\}_{n\geqslant 0}$ in H with a limit in $\mathrm{GL}_n(\Bbbk)$ actually has its limit in G since each $A_n \in H \subseteq G$ and G is closed in $\mathrm{GL}_n(\Bbbk)$. Since H is closed in G, this means that $\{A_n\}_{n\geqslant 0}$ has a limit in H. So H is closed in $\mathrm{GL}_n(\Bbbk)$, showing it is a matrix subgroup. □

Example 1.18. $\mathrm{SL}_n(\Bbbk) \leqslant \mathrm{GL}_n(\Bbbk)$ is a matrix group over $\Bbbk$.

Proof. By Proposition 1.13, $\mathrm{SL}_n(\Bbbk)$ is closed in $\mathrm{M}_n(\Bbbk)$ and $\mathrm{SL}_n(\Bbbk) \subseteq \mathrm{GL}_n(\Bbbk)$. □

Definition 1.19. A closed subgroup $H \leqslant G$ of a matrix group G is called a *matrix subgroup* of G.

Proposition 1.20. *A matrix subgroup $H \leqslant G$ of a matrix group G is a matrix group.*

Proof. This is a direct consequence of Proposition 1.17. □

EXAMPLE 1.21. We can consider $\mathrm{GL}_n(\Bbbk)$ as a subgroup of $\mathrm{GL}_{n+1}(\Bbbk)$ by identifying the $n \times n$ matrix $A = [a_{ij}]$ with

$$\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

and it is easily verified that $\mathrm{GL}_n(\Bbbk)$ is closed in $\mathrm{GL}_{n+1}(\Bbbk)$, hence $\mathrm{GL}_n(\Bbbk)$ is a matrix subgroup of $\mathrm{GL}_{n+1}(\Bbbk)$.

Restricting this embedding to $\mathrm{SL}_n(\Bbbk)$ we find that it embeds as a closed subgroup of $\mathrm{SL}_{n+1}(\Bbbk) \leqslant \mathrm{GL}_{n+1}(\Bbbk)$. Hence $\mathrm{SL}_n(\Bbbk)$ is a matrix subgroup of $\mathrm{SL}_{n+1}(\Bbbk)$.

More generally, any matrix subgroup of $\mathrm{GL}_n(\Bbbk)$ can also be viewed as a matrix subgroup of $\mathrm{GL}_{n+1}(\Bbbk)$ with the aid of this embedding.

Given a matrix subgroup $G \leqslant \mathrm{GL}_n(\Bbbk)$, it will often be useful to restrict the determinant to a function $\det_G\colon G \longrightarrow \Bbbk^\times$, where $\det_G A = \det A$; we usually write this as det when no ambiguity can arise. This is a continuous group homomorphism.

When $\Bbbk = \mathbb{R}$, we set

$$\mathbb{R}^+ = \{t \in \mathbb{R} : t > 0\}, \quad \mathbb{R}^- = \{t \in \mathbb{R} : t < 0\}, \quad \mathbb{R}^\times = \mathbb{R}^+ \cup \mathbb{R}^-.$$

Notice that $\mathbb{R}^+$ is a subgroup of $\mathrm{GL}_1(\mathbb{R}) = \mathbb{R}^\times$ which is both closed and open as a subset, while $\mathbb{R}^-$ is an open subset; hence $\mathbb{R}^+$ and $\mathbb{R}^-$ are *clopen* subsets, i.e., both closed and open. For $G \leqslant \mathrm{GL}_n(\mathbb{R})$,

$$\det\nolimits_G^{-1} \mathbb{R}^+ = G \cap \det^{-1} \mathrm{GL}_n(\mathbb{R}),$$

and also

$$G = \det\nolimits_G^{-1} \mathbb{R}^+ \cup \det\nolimits_G^{-1} \mathbb{R}^-.$$

Hence G is a disjoint union of the clopen subsets

$$G^+ = \det\nolimits_G^{-1} \mathbb{R}^+, \quad G^- = \det\nolimits_G^{-1} \mathbb{R}^-.$$

Since $I_n \in G^+ = \det_G^{-1} \mathbb{R}^+$, the component G^+ is never empty. Indeed, G^+ is a closed subgroup of G, hence it is a matrix subgroup of $\mathrm{GL}_n(\mathbb{R})$. When $G^- \neq \emptyset$, the space G is not connected since it is the union of two disjoint open subsets. When $G^- = \emptyset$, $G = G^+$ may or may not be connected.

If $\Bbbk = \mathbb{R}, \mathbb{C}$, recall that a subset $X \subseteq \Bbbk^m$ is *compact* if and only of it is closed and bounded. Identifying subsets of $\mathrm{M}_n(\Bbbk)$ with subsets of $\Bbbk^{n^2}$, we can specify compact subsets of $\mathrm{M}_n(\Bbbk)$. A matrix group $G \leqslant \mathrm{GL}_n(\Bbbk)$ is *compact* if it is compact as a subset of $\mathrm{M}_n(\Bbbk) \supseteq \mathrm{GL}_n(\Bbbk)$. The following result is standard for metric spaces.

PROPOSITION 1.22. *$X \subseteq \mathrm{M}_n(\Bbbk)$ is compact if and only if the following two conditions are satisfied:*

- *there is a $b \in \mathbb{R}^+$ such that for all $A \in X$, $\|A\| \leqslant b$;*
- *every Cauchy sequence $\{C_n\}_{n \geqslant 0}$ in X has a limit in X.*

Finally, we have the following characterisation of compact sets which is usually taken as the definition of a *compact topological space*.

THEOREM 1.23 (Heine-Borel Theorem). *$X \subseteq \mathrm{M}_n(\Bbbk)$ is compact if and only if every open cover $\{U_\alpha\}_{\alpha \in \Lambda}$ of X contains a finite subcover $\{U_{\alpha_1}, \ldots, U_{\alpha_k}\}$.*

4. Some examples of matrix groups

In this section we discuss some important examples of real and complex matrix groups.

For $n \geqslant 1$, an $n \times n$ matrix $A = [a_{ij}]$ is *upper triangular* if it has the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & \cdots & a_{1n} \\ 0 & a_{21} & \ddots & \ddots & \ddots & a_{2n} \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_{n-2\,n-2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & a_{n-1\,n-1} & \vdots \\ 0 & 0 & \cdots & 0 & 0 & a_{nn} \end{bmatrix},$$

i.e., $a_{ij} = 0$ if $i < j$. A matrix is *unipotent* if it is upper triangular and also has all diagonal entries equal to 1, i.e., $a_{ij} = 0$ if $i < j$ and $a_{ii} = 1$.

The *upper triangular* or *Borel subgroup*) of $\mathrm{GL}_n(\Bbbk)$ is

$$\mathrm{UT}_n(\Bbbk) = \{A \in \mathrm{GL}_n(\Bbbk) : A \text{ is upper triangular}\},$$

while the *unipotent subgroup* of $\mathrm{GL}_n(\Bbbk)$ is

$$\mathrm{SUT}_n(\Bbbk) = \{A \in \mathrm{GL}_n(\Bbbk) : A \text{ is unipotent}\}.$$

It is easy to see that $\mathrm{UT}_n(\Bbbk)$ and $\mathrm{SUT}_n(\Bbbk)$ are closed subgroups of $\mathrm{GL}_n(\Bbbk)$. Notice also that $\mathrm{SUT}_n(\Bbbk) \leqslant \mathrm{UT}_n(\Bbbk)$ and is a closed subgroup.

For the case

$$\mathrm{SUT}_2(\Bbbk) = \left\{ \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \in \mathrm{GL}_2(\Bbbk) : t \in \Bbbk \right\} \leqslant \mathrm{GL}_2(\Bbbk),$$

the function

$$\theta \colon \Bbbk \longrightarrow \mathrm{SUT}_2(\Bbbk); \quad \theta(t) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix},$$

is a continuous group homomorphism which is an isomorphism with continuous inverse. This allows us to view $\Bbbk$ as a matrix group.

The *n-dimensional affine group over* $\Bbbk$ is

$$\mathrm{Aff}_n(\Bbbk) = \left\{ \begin{bmatrix} A & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} : A \in \mathrm{GL}_n(\Bbbk),\ \mathbf{t} \in \Bbbk^n \right\} \leqslant \mathrm{GL}_{n+1}(\Bbbk).$$

This is clearly a closed subgroup of $\mathrm{GL}_{n+1}(\Bbbk)$. If we identify the element $\mathbf{x} \in \Bbbk^n$ with $\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \in \Bbbk^{n+1}$, then since

$$\begin{bmatrix} A & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} = \begin{bmatrix} A\mathbf{x} + \mathbf{t} \\ 1 \end{bmatrix},$$

we obtain an action of $\mathrm{Aff}_n(\Bbbk)$ on $\Bbbk^n$. Transformations of $\Bbbk^n$ having the form $\mathbf{x} \mapsto A\mathbf{x}+\mathbf{t}$ with A invertible are called *affine transformations* and they preserve lines (i.e., translates of 1-dimensional subspaces of the $\Bbbk$-vector space $\Bbbk^n$). The associated geometry is *affine geometry* has $\mathrm{Aff}_n(\Bbbk)$ as its symmetry group. Notice that we can view the vector space $\Bbbk^n$ itself as the *translation subgroup* of $\mathrm{Aff}_n(\Bbbk)$,

$$\mathrm{Trans}_n(\Bbbk) = \left\{ \begin{bmatrix} I_n & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} : \mathbf{t} \in \Bbbk^n \right\} \leqslant \mathrm{Aff}_n(\Bbbk),$$

and this is a closed subgroup.

For $n \geqslant 1$,

$$\mathrm{O}(n) = \{A \in \mathrm{GL}_n(\mathbb{R}) : A^T A = I_n\},$$

is the $n \times n$ *real orthogonal group* where A^T is the *transpose* of $A = [a_{ij}]$,

$$(A^T)_{ij} = a_{ji}.$$

It is easy to see that every *orthogonal matrix* $A \in \mathrm{O}(n)$ has an inverse, namely A^T. Moreover, the product of two orthogonal matrices is orthogonal since $(AB)^T = B^T A^T$. Hence $\mathrm{O}(n) \subseteq \mathrm{GL}_n(\mathbb{R})$. If $A, B \in \mathrm{O}(n)$ then

$$(AB)^T(AB) = B^T A^T AB = BI_n B^T = BB^T = I_n,$$

hence $\mathrm{O}(n)$ is closed under multiplication. Notice also that $I_n \in \mathrm{O}(n)$. Together these facts imply that $\mathrm{O}(n) \leqslant \mathrm{GL}_n(\mathbb{R})$, i.e., $\mathrm{O}(n)$ is a subgroup of $\mathrm{GL}_n(\mathbb{R})$.

The single matrix equation $A^T A = I_n$ is equivalent to n^2 equations for the n^2 real numbers a_{ij},

$$\sum_{k=1}^{n} a_{ki} a_{kj} = \delta_{ij} \tag{1.1}$$

where the *Kronecker symbol* δ_{ij} is defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

This means that $\mathrm{O}(n)$ is a closed subset of $\mathrm{M}_n(\mathbb{R})$ and hence of $\mathrm{GL}_n(\mathbb{R})$.

Let us consider the determinant function restricted to $\mathrm{O}(n)$, $\det\colon \mathrm{O}(n) \longrightarrow \mathbb{R}^\times$. Then for $A \in \mathrm{O}(n)$,

$$\det I_n = \det(A^T A) = \det A^T \det A = (\det A)^2,$$

hence $\det A = \pm 1$. So we have

$$\mathrm{O}(n) = \mathrm{O}(n)^+ \cup \mathrm{O}(n)^-,$$

where

$$\mathrm{O}(n)^+ = \{A \in \mathrm{O}(n) : \det A = 1\}, \quad \mathrm{O}(n)^- = \{A \in \mathrm{O}(n) : \det A = -1\}.$$

The subgroup $\mathrm{SO}(n) = \mathrm{O}(n)^+$ is called the $n \times n$ *special orthogonal group.*

One of the main reasons for the study of these groups $\mathrm{SO}(n), \mathrm{O}(n)$ is their relationship with *isometries* where an *isometry* of $\mathbb{R}^n$ is a distance preserving function $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$. If such an isometry fixes the origin $\mathbf{0}$ then it is actually a linear transformation and so with respect to say the standard basis corresponds to a matrix A. The isometry condition is equivalent to the fact that

$$A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y} \quad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n),$$

which is in turn equivalent to the condition that $A^T A = I_n$, i.e., A is orthogonal. Elements of $\mathrm{SO}(n)$ are called *direct isometries* or *rotations;* elements of $\mathrm{O}(n)^-$ are sometimes called *indirect isometries.*

A more general situation is associated with an $n \times n$ real symmetric matrix Q. Then there is an analogue of the orthogonal group,

$$\mathrm{O}_Q = \{A \in \mathrm{GL}_n(\mathbb{R}) : A^T Q A = Q\}.$$

It is easy to see that this is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ and so is a matrix group. Moreover, if $\det Q \neq 0$, for $A \in \mathrm{O}_Q$ we have $\det A = \pm 1$. We can also define

$$\mathrm{O}_Q^+ = \det^{-1} \mathbb{R}^+, \quad \mathrm{O}_Q^- = \det^{-1} \mathbb{R}^-$$

and can write O_Q as a disjoint union of clopen subsets $\mathrm{O}_Q = \mathrm{O}_Q^+ \cup \mathrm{O}_Q^-$ where O_Q^+ is a subgroup.

An important example of this occurs in relativity where $n = 4$ and

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

The *Lorentz group* Lor is the closed subgroup of $\mathrm{O}_Q^+ \cap \mathrm{SL}_2(\mathbb{R})$ which preserves each of the two connected components of the hyperboloid

$$x_1^2 + x_2^2 + x_3^2 - x_4^2 = -1.$$

A similar construction can be carried out starting with an $n \times n$ real *skew symmetric* matrix S, i.e., $S^T = -S$. If $\det S \neq 0$ then it turns out that n has to be even, so $n = 2m$. The standard example is built up from 2×2 blocks

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

and we get

$$J_{2m} = \begin{bmatrix} J & O_2 & \cdots & O_2 \\ O_2 & J & \cdots & O_2 \\ \vdots & \vdots & \ddots & \vdots \\ O_2 & O_2 & \cdots & J \end{bmatrix}.$$

The matrix group

$$\mathrm{Symp}_{2m}(\mathbb{R}) = \{A \in \mathrm{GL}_{2m}(\mathbb{R}) : A^T J_{2m} A = J_{2m}\} \leqslant \mathrm{GL}_{2m}(\mathbb{R}),$$

is called the $2m \times 2m$ real *symplectic group*. It is easily checked that $\mathrm{Symp}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$, but in general $\mathrm{Symp}_{2m}(\mathbb{R}) \neq \mathrm{SL}_{2m}(\mathbb{R})$.

Symplectic geometry has become extremely important and is the natural geometry associated to Hamiltonian mechanics and therefore to quantum mechanics; it is also important as an area of differential geometry and in the study of 4-dimensional manifolds. The symplectic groups are the natural symmetry groups of such geometries.

For $A = [a_{ij}] \in \mathrm{M}_n(\mathbb{C})$,

$$A^* = (\overline{A})^T = \overline{(A^T)},$$

is the *hermitian conjugate* of A, i.e., $(A^*)_{ij} = \overline{a_{ji}}$. The $n \times n$ *unitary group* is the subgroup

$$\mathrm{U}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) : A^*A = I\} \leqslant \mathrm{GL}_n(\mathbb{C}).$$

Again the unitary condition amounts to n^2 equations for the n^2 complex numbers a_{ij} (compare Equation (1.1)),

$$\sum_{k=1}^{n} \overline{a}_{ki} a_{kj} = \delta_{ij}. \tag{1.2}$$

By taking real and imaginary parts, these equations actually give $2n^2$ bilinear equations in the $2n^2$ real and imaginary parts of the a_{ij}, although there is some redundancy.

The $n \times n$ *special unitary group* is

$$\mathrm{SU}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) : A^*A = I \text{ and } \det A = 1\} \leqslant \mathrm{U}(n).$$

Again we can specify that a matrix is special unitary by requiring that its entries satisfy the (n^2+1) equations

$$\begin{cases} \sum_{k=1}^{n} \overline{a}_{ki} a_{kj} = \delta_{ij} \quad (1 \leqslant i, j \leqslant n), \\ \det A = 1. \end{cases} \tag{1.3}$$

Of course, $\det A$ is a polynomial in the a_{ij}. Notice that $\mathrm{SU}(n)$ is a normal subgroup of $\mathrm{U}(n)$, $\mathrm{SU}(n) \triangleleft \mathrm{U}(n)$.

The dot product on $\mathbb{R}^n$ can be extended to $\mathbb{C}^n$ by setting

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^* \mathbf{y} = \sum_{k=1}^{n} \overline{x}_k y_k,$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Note that this is not $\mathbb{C}$-linear but satisfies

$$(u\mathbf{x}) \cdot (v\mathbf{y}) = \overline{u}v(\mathbf{x} \cdot \mathbf{y}).$$

This dot product allows us to define the length of a complex vector by

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

since $\mathbf{x} \cdot \mathbf{x}$ is a non-negative real number which is zero only when $\mathbf{x} = \mathbf{0}$. Then a matrix $A \in \mathrm{M}_n(\mathbb{C})$ is unitary if and only if

$$A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y} \quad (\mathbf{x}, \mathbf{y} \in \mathbb{C}^n).$$

5. Complex matrix groups as real matrix groups

Recall that the complex numbers can be viewed as a 2-dimensional real vector space, with basis $1, i$ for example. Similarly, every $n \times n$ complex matrix $Z = [z_{ij}]$ can also be viewed as a $2n \times 2n$ real matrix as follows.

We identify each complex number $z = x + yi$ with a 2×2 real matrix by defining a function

$$\rho \colon \mathbb{C} \longrightarrow \mathrm{M}_2(\mathbb{R}); \quad \rho(x + yi) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}.$$

This turns out to be an injective ring homomorphism, so we can view $\mathbb{C}$ as a subring of $\mathrm{M}_2(\mathbb{R})$, i.e.,

$$\operatorname{im} \rho = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{M}_2(\mathbb{R}) : d = a,\ c = -b \right\}.$$

Notice that complex conjugation corresponds to transposition, i.e.,

$$\rho(\overline{z}) = \rho(z)^T. \tag{1.4}$$

More generally, given $Z = [z_{ij}] \in \mathrm{M}_n(\mathbb{C})$ with $z_{rs} = x_{rs} + y_{rs}i$, we can write

$$Z = [x_{ij}] + i[y_{ij}]$$

where the two $n \times n$ matrices $X = [x_{ij}]$, $Y = [y_{ij}]$ are real symmetric.

Define a function

$$\rho_n \colon \mathrm{M}_n(\mathbb{C}) \longrightarrow \mathrm{M}_{2n}(\mathbb{R}); \quad \rho_n(Z) = \begin{bmatrix} X & -Y \\ Y & X \end{bmatrix},$$

which is an injective ring homomorphism.

Let J_{2n} denote the $2n \times 2n$ real matrix with block form

$$J_{2n} = \begin{bmatrix} O_n & -I_n \\ I_n & O_n \end{bmatrix}.$$

Notice that $J_{2n}^2 = -I_{2n}$ and $J_{2n}^T = -J_{2n}$. We have

$$\rho_n(Z) = \begin{bmatrix} X & O_n \\ O_n & X \end{bmatrix} + \begin{bmatrix} Y & O_n \\ O_n & Y \end{bmatrix} J_{2n};$$
$$\rho_n(\overline{Z}) = \rho_n(Z)^T.$$

Notice that $\rho_n(\mathrm{GL}_n(\mathbb{C})) \leqslant \mathrm{GL}_{2n}(\mathbb{R})$, so any matrix subgroup $G \leqslant \mathrm{GL}_n(\mathbb{C})$ can be viewed as a matrix subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$ by identifying it with its image $\rho_n G$ under ρ_n (this uses the fact that ρ_n is continuous).

6. Continuous homomorphisms of matrix groups

In group theory the notion of a homomorphism of groups is central. For matrix groups we need to be careful about topological properties as well as the algebraic ones.

Definition 1.24. Let G, H be two matrix groups. A group homomorphism $\varphi\colon G \longrightarrow H$ is a *continuous homomorphism of matrix groups* if it is continuous and its image $\operatorname{im}\varphi = \varphi G \leqslant H$ is a closed subspace of H.

Example 1.25. The function

$$\varphi\colon \mathrm{SUT}_2(\mathbb{R}) \longrightarrow \mathrm{U}(1); \quad \varphi\left(\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\right) = [e^{2\pi i t}]$$

is a continuous surjective group homomorphism, so it is a continuous homomorphism of matrix groups.

To see why this definition is necessary, consider the following example.

Example 1.26. Let

$$G = \left\{ \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} \in \mathrm{SUT}_1(\mathbb{R}) : n \in \mathbb{Z} \right\}.$$

Then G is a closed subgroup of $\mathrm{SUT}_1(\mathbb{R})$, so it is a matrix group.

For any irrational number $r \in \mathbb{R} - \mathbb{Q}$, the function

$$\varphi\colon G \longrightarrow \mathrm{U}(1); \quad \varphi\left(\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}\right) = \left[e^{2\pi i r n}\right]$$

is a continuous group homomorphism. But its image is a dense proper subset of $\mathrm{U}(1)$. So φ is not a continuous homomorphism of matrix groups.

The point of this example is that φG has limit points in $\mathrm{U}(1)$ which are not in φG, whereas G is discrete as a subspace of $\mathrm{SUT}_2(\mathbb{R})$.

Whenever we have a homomorphism of matrix groups $\varphi\colon G \longrightarrow H$ which is a homeomorphism (i.e., a bijection with continuous inverse) we say that φ is a *continuous isomorphism of matrix groups* and regard G and H as essentially identical as matrix groups.

Proposition 1.27. *Let $\varphi\colon G \longrightarrow H$ be a continuous homomorphism of matrix groups. Then $\ker\varphi \leqslant G$ is a closed subgroup, hence $\ker\varphi$ is a matrix group.*

The quotient group $G/\ker\varphi$ can be identified with the matrix group φG by the usual quotient isomorphism $\overline{\varphi}\colon G/\ker\varphi \longrightarrow \varphi G$.

PROOF. Since φ is continuous, whenever it makes sense in G,

$$\lim_{n\to\infty} \varphi(A_n) = \varphi(\lim_{n\to\infty} A_n),$$

which implies that a limit of elements of $\ker\varphi$ in G is also in $\ker\varphi$. So $\ker\varphi$ is a closed subset of G.

The fact that $\ker\varphi \leqslant G$ is a matrix group follows from Proposition 1.17. $\square$

REMARK 1.28. $G/\ker\varphi$ has a natural *quotient topology* which is not obviously a metric topology. Then $\overline{\varphi}$ is always a homoeomorphism.

REMARK 1.29. Not every closed normal matrix subgroup $N \triangleleft G$ of a matrix group G gives rise to a matrix group G/N; there are examples for which G/N is a *Lie group* but not a matrix group. This is one of the most important differences between matrix groups and Lie groups (we will see later that every matrix group is a Lie group). One consequence is that certain important matrix groups have quotients which are not matrix groups and therefore have no faithful finite dimensional representations; such groups occur readily in Quantum Physics, where their infinite dimensional representations play an important rôle.

7. Continuous group actions

In ordinary group theory, the notion of a *group action* is fundamental. Suitably formulated, it amounts to the following. An *action* μ of a group G on a set X is a function

$$\mu\colon G\times X \longrightarrow X$$

for which we usually write $\mu(g,x) = gx$ if there is no danger of ambiguity, satisfying the following conditions for all $g,h\in G$ and $x \in X$ and with ι being the identity element of G:

- $(gh)x = g(hx)$, i.e., $\mu(gh,x) = \mu(g,\mu(h,x))$;
- $\iota x = x$.

There are two important notions associated to such an action.

For $x\in X$, the *stabilizer of* x is

$$\mathrm{Stab}_G(x) = \{g\in G : gx = x\} \subseteq G,$$

while the *orbit of* x is

$$\mathrm{Orb}_G(x) = \{gx\in X : g\in G\} \subseteq X.$$

THEOREM 1.30. *Let G act on X.*
a) *For $x\in X$, $\mathrm{Stab}_G(x)\leqslant G$, i.e., $\mathrm{Stab}_G(x)$ is a subgroup of G.*
b) *For $x,y\in X$, $y\in\mathrm{Orb}_G(x)$ if and only if $\mathrm{Orb}_G(y)=\mathrm{Orb}_G(x)$.*
For $x\in X$, there is a bijection

$$\varphi\colon G/\mathrm{Stab}_G(x)\longrightarrow \mathrm{Orb}_G(x);\quad \varphi(g) = gx.$$

Furthermore, this is G-equivariant in the sense that for all $g,h\in G$,

$$\varphi((hg)\,\mathrm{Stab}_G(x)) = h\varphi(g\,\mathrm{Stab}_G(x)).$$

c) *If $y\in\mathrm{Orb}_G(x)$, then for any $t\in G$ with $y=tx$,*

$$\mathrm{Stab}_G(y) = t\,\mathrm{Stab}_G(x)t^{-1}.$$

For a topological group there is a notion of *continuous group action* on a topological space.

DEFINITION 1.31. Let G be a topological group and X a topological space. Then a group action $\mu\colon G\times X\longrightarrow X$ is a *continuous group action* if the function μ is continuous.

In this definition $G \times X$ has the *product topology.* When G and X are metric spaces this can be obtained from a suitable metric. Details of this can be found in the the first Problem Set.

If X is Hausdorff then any one-element subset $\{x\}$ is closed and $\mathrm{Stab}_G(x) \leqslant G$ is a closed subgroup. This provides a useful way of producing closed subgroups.

8. The matrix exponential and logarithm functions

Let $\Bbbk = \mathbb{R}$ or $\mathbb{C}$. The power series

$$\mathrm{Exp}(X) = \sum_{n\geqslant 0} \frac{1}{n!} X^n, \quad \mathrm{Log}(X) = \sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n} X^n,$$

have radii of convergence (r. o. c) ∞ and 1 respectively. If $z \in \mathbb{C}$, the series $\mathrm{Exp}(z)$, $\mathrm{Log}(z)$ converge absolutely whenever $|z| <$ r. o. c.

Let $A \in \mathrm{M}_n(\Bbbk)$. The matrix valued series

$$\mathrm{Exp}(A) = \sum_{n\geqslant 0} \frac{1}{n!} A^n = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots ,$$

$$\mathrm{Log}(A) = \sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n} A^n = A - \frac{1}{2}A^2 + \frac{1}{3}A^3 - \frac{1}{4}A^4 + \cdots ,$$

will converge provided $\|A\| <$ r. o. c. So $\mathrm{Exp}(A)$ makes sense for every $A \in \mathrm{M}_n(\Bbbk)$ while $\mathrm{Log}(A)$ only exists if $\|A\| < 1$.

PROPOSITION 1.32. *Let $A \in \mathrm{M}_n(\Bbbk)$.*

a) *For $u, v \in \mathbb{C}$, $\mathrm{Exp}((u+v)A) = \mathrm{Exp}(uA)\,\mathrm{Exp}(vA)$.*
b) *$\mathrm{Exp}(A) \in \mathrm{GL}_n(\Bbbk)$ and $\mathrm{Exp}(A)^{-1} = \mathrm{Exp}(-A)$.*

PROOF.
a) Expanding the series gives

$$\begin{aligned}\mathrm{Exp}((u+v)A) &= \sum_{n\geqslant 0} \frac{1}{n!}(u+v)^n A^n \\ &= \sum_{n\geqslant 0} \frac{(u+v)^n}{n!} A^n.\end{aligned}$$

By a series of manipulations that can be justified since these series are all absolutely convergent,

$$\begin{aligned}\mathrm{Exp}(uA)\,\mathrm{Exp}(vA) &= \left(\sum_{r\geqslant 0} \frac{u^r}{r!} A^r\right)\left(\sum_{s\geqslant 0} \frac{v^s}{s!} A^s\right) \\ &= \sum_{\substack{r\geqslant 0\\ s\geqslant 0}} \frac{u^r v^s}{r!s!} A^{r+s} \\ &= \sum_{n\geqslant 0} \left(\sum_{r=0}^{n} \frac{u^r v^{n-r}}{r!(n-r)!}\right) A^n \\ &= \sum_{n\geqslant 0} \frac{1}{n!} \left(\sum_{r=0}^{n} \binom{n}{r} u^r v^{n-r}\right) A^n \\ &= \sum_{n\geqslant 0} \frac{(u+v)^n}{n!} A^n \\ &= \mathrm{Exp}((u+v)A).\end{aligned}$$

b) From part (a),

$$I = \mathrm{Exp}(O) = \mathrm{Exp}((1+(-1))A) = \mathrm{Exp}(A)\,\mathrm{Exp}(-A),$$

so $\mathrm{Exp}(A)$ is invertible with inverse $\mathrm{Exp}(-A)$. □

Using these series we define the *exponential function*

$$\exp\colon \mathrm{M}_n(\Bbbk) \longrightarrow \mathrm{GL}_n(\Bbbk); \quad \exp(A) = \mathrm{Exp}(A).$$

PROPOSITION 1.33. *If $A, B \in \mathrm{M}_n(\Bbbk)$ commute then*

$$\exp(A+B) = \exp(A)\exp(B).$$

PROOF. Again we expand the series and perform a sequence of manipulations all of which can be justified.

$$\begin{aligned}
\exp(A)\exp(B) &= \left(\sum_{r\geqslant 0} \frac{1}{r!}A^r\right)\left(\sum_{s\geqslant 0} \frac{1}{s!}B^s\right) \\
&= \sum_{\substack{r\geqslant 0 \\ s\geqslant 0}} \frac{1}{r!s!}A^r B^s \\
&= \sum_{n\geqslant 0}\left(\sum_{r=0}^{n} \frac{1}{r!(n-r)!}A^r B^{n-r}\right) \\
&= \sum_{n\geqslant 0}\frac{1}{n!}\left(\sum_{r=0}^{n} \binom{n}{r}A^r B^{n-r}\right) \\
&= \sum_{n\geqslant 0}\frac{1}{n!}(A+B)^n \\
&= \mathrm{Exp}(A+B).
\end{aligned}$$

Notice that we make crucial use of the commutativity of A and B in the identity

$$\sum_{r=0}^{n}\binom{n}{r}A^r B^{n-r} = (A+B)^n.$$

□

Define the *logarithmic function*

$$\log\colon \mathrm{N}_{\mathrm{M}_n(\Bbbk)}(I;1) \longrightarrow \mathrm{M}_n(\Bbbk); \quad \log(A) = \mathrm{Log}(A-I).$$

Then for $\|A-I\|<1$,

$$\log(A) = \sum_{n\geqslant 1}\frac{(-1)^{n-1}}{n}(A-I)^n.$$

PROPOSITION 1.34. *The functions* $\exp$ *and* $\log$ *satisfy*

a) *if $\|A-I\|<1$, then $\exp(\log(A)) = A$;*
b) *if $\|\exp(B)-I\|<1$, then $\log(\exp(B)) = B$.*

PROOF. These results follow from the formal identities between power series

$$\sum_{m\geqslant 0}\frac{1}{m!}\left(\sum_{n\geqslant 1}\frac{(-1)^{n-1}}{n}(X-1)^n\right)^m = X,$$

$$\sum_{n\geqslant 1}\frac{(-1)^{n-1}}{n}\left(\sum_{m\geqslant 1}\frac{1}{m!}X^m\right)^n = X,$$

proved by comparing coefficients. □

The functions $\exp, \log$ are continuous and in fact infinitely differentiable on their domains. By continuity of exp at O, there is a $\delta_1 > 0$ such that

$$\mathrm{N}_{\mathrm{M}_n(\Bbbk)}(O;\delta_1) \subseteq \exp^{-1} \mathrm{N}_{\mathrm{GL}_n(\Bbbk)}(I;1).$$

In fact we can actually take $\delta_1 = \log 2$ since

$$\exp \mathrm{N}_{\mathrm{M}_n(\Bbbk)}(O;r) \subseteq \mathrm{N}_{\mathrm{M}_n(\Bbbk)}(I;e^r - 1).$$

Hence we have

PROPOSITION 1.35. *The exponential function* exp *is injective when restricted to the open subset* $\mathrm{N}_{\mathrm{M}_n(\Bbbk)}(O;\ln 2) \subseteq \mathrm{M}_n(\Bbbk)$, *hence it is locally a diffeomorphism at O with local inverse* log.

It will sometimes be useful to have a formula for the derivative of exp at an arbitrary $A \in \mathrm{M}_n(\Bbbk)$. When $B \in \mathrm{M}_n(\Bbbk)$ commutes with A,

$$\frac{\mathrm{d}}{\mathrm{d}\,t}_{|t=0} \exp(A + tB) = \lim_{h\to 0} \frac{1}{h}\left(\exp(A + hB) - \exp(A)\right) = \exp(A)B = B\exp(A). \tag{1.5}$$

However, the general situation is more complicated.

For a variable X consider the series

$$F(X) = \sum_{r\geqslant 0} \frac{1}{(k+1)!} X^k = \frac{\exp(X) - 1}{X}$$

which has infinite radius of convergence. If we have a linear operator Φ on $\mathrm{M}_n(\mathbb{C})$ we can apply the convergent series of operators

$$F(\Phi) = \sum_{r\geqslant 0} \frac{1}{(k+1)!} \Phi^k$$

to elements of $\mathrm{M}_n(\mathbb{C})$. In particular we can consider

$$\Phi(C) = AC - CA = \operatorname{ad} A(C),$$

where

$$\operatorname{ad} A \colon \mathrm{M}_n(\mathbb{C}) \longrightarrow \mathrm{M}_n(\mathbb{C}); \quad \operatorname{ad} A(C) = AC - CA,$$

is viewed as a $\mathbb{C}$-linear operator. Then

$$F(\operatorname{ad} A)(C) = \sum_{r\geqslant 0} \frac{1}{(k+1)!} (\operatorname{ad} A)^k(C).$$

PROPOSITION 1.36. *For $A, B \in \mathrm{M}_n(\mathbb{C})$ we have*

$$\frac{\mathrm{d}}{\mathrm{d}\,t}_{|t=0} \exp(A + tB) = F(\operatorname{ad} A)(B)\, \exp(A).$$

In particular, if $A = O$ or more generally if $AB = BA$,

$$\frac{\mathrm{d}}{\mathrm{d}\,t}_{|t=0} \exp(A + tB) = B\exp(A).$$

PROOF. We begin by observing that if $D = \dfrac{\mathrm{d}}{\mathrm{d}\,s}$ and $f(s)$ is a smooth function of the real variable s, then

$$F(D)_{|s=0} f(s) = \int_0^1 f(s)\, \mathrm{d}\,s. \tag{1.6}$$

This holds since the Taylor expansion of a smooth function g satisfies

$$\sum_{r\geqslant 1} \frac{1}{k!} D^k g(s) = g(s+1) - g(s),$$

hence taking $g(s) = \int f(s)\,\mathrm{d}\,s$ to be an indefinite integral of f we obtain

$$\sum_{r\geqslant 0} \frac{1}{(k+1)!} D^k f(s) = g(s+1) - g(s).$$

Evaluating at $s = 0$ gives the Equation (1.6).

Now note that the matrix valued function

$$\varphi(s) = \exp(sA) B \exp((1-s)A)$$

satisfies

$$\begin{aligned}
\varphi(s) &= \exp(sA) B \exp(A) \exp(-sA) \\
&= \exp(s \operatorname{ad} A)(B \exp(A)) \\
&= \exp(s \operatorname{ad} A)(B) \exp(A),
\end{aligned}$$

since for $m, n \geqslant 1$

$$, (\operatorname{ad} A)^m (BA^n) = (\operatorname{ad} A)^m (B)\, A^n.$$

So

$$F(D)(\varphi(s)) = \left(\sum_{k\geqslant 0} \frac{((s+1)^{k+1} - s^{k+1})}{(k+1)!} (\operatorname{ad} A)^k \right) (B) \exp(A)$$

giving

$$\begin{aligned}
F(D)(\varphi(s))_{|s=0} &= \left(\sum_{k\geqslant 0} \frac{1}{(k+1)!} (\operatorname{ad} A)^k \right) (B) \exp(A) \\
&= F(\operatorname{ad} A)(B) \exp(A).
\end{aligned}$$

□

CHAPTER 2

Lie algebras for matrix groups

1. Differential equations in matrices

Let $A \in \mathrm{M}_n(\mathbb{R})$. Let $(a,b) \subseteq \mathbb{R}$ be the open interval with endpoints a, b and $a < b$; we will usually assume that $a < 0 < b$. We will use the standard notation

$$\alpha'(t) = \frac{\mathrm{d}}{\mathrm{d}\,t}\,\alpha(t).$$

Consider the first order differential equation

$$\alpha'(t) = \alpha(t)A, \tag{2.1}$$

where $\alpha\colon (a,b) \longrightarrow \mathrm{M}_n(\mathbb{R})$ is assumed to be a differentiable function.

If $n = 1$ then taking A to be a non-zero real number we know that the general solution is $\alpha(t) = ce^{At}$ where $\alpha(0) = c$. Hence there is a unique solution subject to this boundary condition. In fact this solution is given by a power series

$$\alpha(t) = \sum_{k\geqslant 0} \frac{t^k}{k!}\alpha(0).$$

This is indicative of the general situation.

Theorem 2.1. *For $A, C \in \mathrm{M}_n(\mathbb{R})$ with A non-zero, and $a < 0 < b$, the differential equation of* (2.1) *has a unique solution $\alpha\colon (a,b) \longrightarrow \mathrm{M}_n(\mathbb{R})$ for which $\alpha(0) = C$. Furthermore, if C is invertible then so is $\alpha(t)$ for $t \in (a,b)$, hence $\alpha\colon (a,b) \longrightarrow \mathrm{GL}_n(\mathbb{R})$.*

Proof. First we will solve the equation subject to the boundary condition $\alpha(0) = I$. For $t \in (a,b)$, by Chapter 1 Section 8 the series

$$\sum_{k\geqslant 0} \frac{t^k}{k!}A^k = \sum_{k\geqslant 0} \frac{1}{k!}(tA)^k = \exp(tA)$$

converges, so the function

$$\alpha\colon (a,b) \longrightarrow \mathrm{M}_n(\mathbb{R}); \quad \alpha(t) = \exp(tA),$$

is defined and differentiable with

$$\alpha'(t) = \sum_{k\geqslant 1} \frac{t^{k-1}}{(k-1)!}A^k = \exp(tA)A = A\exp(tA).$$

Hence α satisfies the above differential equation with boundary condition $\alpha(0) = I$. Notice also that whenever $s, t, (s+t) \in (a,b)$,

$$\alpha(s+t) = \alpha(s)\alpha(t).$$

In particular, this shows that $\alpha(t)$ is always invertible with $\alpha(t)^{-1} = \alpha(-t)$.

One solution subject to $\alpha(0) = C$ is easily seen to be $\alpha(t) = C\exp(tA)$. If β is a second such solution then $\gamma(t) = \beta(t)\exp(-tA)$ satisfies

$$\begin{aligned}\gamma'(t) &= \beta'(t)\exp(-tA) + \beta(t)\frac{\mathrm{d}}{\mathrm{d}\,t}\exp(-tA)\\ &= \beta'(t)\exp(-tA) - \beta(t)\exp(-tA)A\\ &= \beta(t)A\exp(-tA) - \beta(t)\exp(-tA)A\\ &= O.\end{aligned}$$

Hence $\gamma(t)$ is a constant function with $\gamma(t) = \gamma(0) = C$. Thus $\beta(t) = C\exp(tA)$, and this is the unique solution subject to $\beta(0) = C$. If C is invertible so is $C\exp(tA)$ for all t. □

2. One parameter subgroups

Let $G \leqslant \mathrm{GL}_n(\Bbbk)$ be a matrix group and let $\varepsilon > 0$ or $\varepsilon = \infty$.

DEFINITION 2.2. A *one parameter semigroup in G* is a continuous function $\gamma\colon (-\varepsilon, \varepsilon) \longrightarrow G$ which is differentiable at 0 and satisfies

$$\gamma(s + t) = \gamma(s)\gamma(t)$$

whenever $s, t, (s + t) \in (-\varepsilon, \varepsilon)$. We will refer to the last condition as the *homomorphism property*.

If $\varepsilon = \infty$ then $\gamma\colon \mathbb{R} \longrightarrow G$ is called a *one parameter group in G* or *one parameter subgroup of G*.

Notice that for a one parameter semigroup in G, $\gamma(0) = I$.

PROPOSITION 2.3. *Let $\gamma\colon (-\varepsilon, \varepsilon) \longrightarrow G$ be a one parameter semigroup in G. Then γ is differentiable at every $t \in (-\varepsilon, \varepsilon)$ and*

$$\gamma'(t) = \gamma'(0)\gamma(t) = \gamma(t)\gamma'(0).$$

PROOF. For small $h \in \mathbb{R}$ we have

$$\gamma(h)\gamma(t) = \gamma(h + t) = \gamma(t + h) = \gamma(t)\gamma(h).$$

Hence

$$\begin{aligned}\gamma'(t) &= \lim_{h\to 0}\frac{1}{h}(\gamma(t + h) - \gamma(t))\\ &= \lim_{h\to 0}\frac{1}{h}(\gamma(h) - I)\gamma(t)\\ &= \gamma'(0)\gamma(t),\end{aligned}$$

and similarly

$$\gamma'(t) = \gamma(t)\gamma'(0).$$

□

PROPOSITION 2.4. *Let $\gamma\colon (-\varepsilon, \varepsilon) \longrightarrow G$ be a one parameter semigroup in G. Then there is a unique extension to a one parameter group $\widetilde{\gamma}\colon \mathbb{R} \longrightarrow G$ in G, i.e., such that for all $t \in (-\varepsilon, \varepsilon)$, $\widetilde{\gamma}(t) = \gamma(t)$.*

PROOF. Let $t \in \mathbb{R}$. Then for a large enough natural number m, $t/m \in (-\varepsilon, \varepsilon)$. Hence

$$\gamma(t/m), \gamma(t/m)^m \in G.$$

Similarly, for a second such natural number n,

$$\gamma(t/n), \gamma(t/n)^n \in G.$$

Then since $mn \geqslant m, n$ we have $t/mn \in (-\varepsilon, \varepsilon)$ and

$$\begin{aligned}\gamma(t/n)^n &= \gamma(mt/mn)^n \\ &= \gamma(t/mn)^{mn} \\ &= \gamma(nt/mn)^m \\ &= \gamma(t/m)^m.\end{aligned}$$

So $\gamma(t/n)^n = \gamma(t/m)^m$ showing that we get a well defined element of G for every real number t. This defines a function

$$\widetilde{\gamma}\colon \mathbb{R} \longrightarrow G; \quad \widetilde{\gamma}(t) = \gamma(t/n)^n \quad \text{for large } n.$$

It is easy to see that $\widetilde{\gamma}$ is a one parameter group in G. □

We can now determine the form of all one parameter groups in G.

THEOREM 2.5. *Let $\gamma\colon \mathbb{R} \longrightarrow G$ be a one parameter group in G. Then it has the form*

$$\gamma(t) = \exp(tA)$$

for some $A \in \mathrm{M}_n(\Bbbk)$.

PROOF. Let $A = \gamma'(0)$. By Proposition 2.3 this means that γ satisfies the differential equation

$$\gamma'(t) = A, \quad \gamma(0) = I.$$

By Theorem 2.1, this has the unique solution $\gamma(t) = \exp(tA)$. □

REMARK 2.6. We cannot yet reverse this process and decide for which $A \in \mathrm{M}_n(\Bbbk)$ the one parameter group

$$\gamma\colon \mathbb{R} \longrightarrow \mathrm{GL}_n(\Bbbk); \quad \gamma(t) = \exp(tA)$$

actually takes values in G. The answer involves the *Lie algebra* of G. Notice that we also have a curious phenomenon in the fact that although the definition of a one parameter group only involves first order differentiability, the general form $\exp(tA)$ is always infinitely differentiable and indeed analytic as a function of t. This is an important characteristic of much of Lie theory, namely that conditions of first order differentiability and even continuity often lead to much stronger conclusions.

3. Curves, tangent spaces and Lie algebras

Throughout this section, let $G \leqslant \mathrm{GL}_n(\Bbbk)$ be a matrix group.

DEFINITION 2.7. A *differentiable curve in G* is a function

$$\gamma\colon (a, b) \longrightarrow G \subseteq \mathrm{M}_n(\Bbbk)$$

for which the derivative $\gamma'(t)$ exists at each $t \in (a, b)$.

Here we define the derivative as an element of $\mathrm{M}_n(\Bbbk)$ by

$$\gamma'(t) = \lim_{s \to t} \frac{1}{(s-t)} \left(\gamma(s) - \gamma(t)\right),$$

provided this limit exists. We will usually assume that $a < 0 < b$.

DEFINITION 2.8. The *tangent space to G at $U \in G$* is

$$\mathrm{T}_U\, G = \{\gamma'(0) \in \mathrm{M}_n(\Bbbk) : \gamma \text{ a differentiable curve in } G \text{ with } \gamma(0) = U\}.$$

PROPOSITION 2.9. *$\mathrm{T}_U\, G$ is a real vector subspace of $\mathrm{M}_n(\Bbbk)$.*

PROOF. Suppose that α, β are differentiable curves in G for which $\alpha(0) = \beta(0) = U$. Then

$$\gamma\colon \operatorname{dom}\alpha \cap \operatorname{dom}\beta \longrightarrow G; \quad \gamma(t) = \alpha(t)U^{-1}\beta(t),$$

is also a differentiable curve in G with $\gamma(0) = U$. The Product Rule now gives

$$\gamma'(t) = \alpha'(t)U^{-1}\beta(t) + \alpha(t)U^{-1}\beta'(t),$$

hence

$$\gamma'(0) = \alpha'(0)U^{-1}\beta(0) + \alpha(0)U^{-1}\beta'(0) = \alpha'(0) + \beta'(0),$$

which shows that T_U is closed under addition.

Similarly, if $r \in \mathbb{R}$ and α is a differentiable curve in G with $\alpha(0) = U$, then $\eta(t) = \alpha(rt)$ defines another such curve. Since

$$\eta'(0) = r\alpha'(0),$$

we see that $\mathrm{T}_U\, G$ is closed under real scalar multiplication. □

DEFINITION 2.10. The *dimension* of the real matrix group G is

$$\dim G = \dim_{\mathbb{R}} \mathrm{T}_I\, G.$$

If G is complex then its *complex dimension*

$$\dim_{\mathbb{C}} G = \dim_{\mathbb{C}} \mathrm{T}_I\, G.$$

We will adopt the notation $\mathfrak{g} = \mathrm{T}_I\, G$ for this real vector subspace of $\mathrm{M}_n(\Bbbk)$. In fact, $\mathfrak{g}$ has a more interesting algebraic structure, namely that of a real *Lie algebra*.

DEFINITION 2.11. A $\Bbbk$-*Lie algebra* consists of a vector space $\mathfrak{a}$ over a field $\Bbbk$, equipped with a $\Bbbk$-bilinear map $[\ ,\ ]\colon \mathfrak{a} \times \mathfrak{a} \longrightarrow \mathfrak{a}$ such that for $x, y, z \in \mathfrak{a}$,

(Skew symmetry) $\qquad [x, y] = -[y, x],$

(Jacobi identity) $\qquad [x, [y, z]] + [y, [x, z]] + [z, [x, y]] = 0.$

Here $\Bbbk$-bilinear means that for $x_1, x_2, x, y_1, y_2, y \in \mathfrak{a}$ and $r_1, r_2, r, s_1, s_2, s \in \Bbbk$,

$$[r_1x_1 + r_2x_2, y] = r_1[x_1, y] + r_2[x_2, y],$$
$$[x, s_1y_1 + s_2y_2] = s_1[x, y_1] + s_2[x, y_2].$$

$[\ ,\ ]$ is called the *Lie bracket* of the Lie algebra $\mathfrak{a}$.

EXAMPLE 2.12. Let $\Bbbk = \mathbb{R}$ and $\mathfrak{a} = \mathbb{R}^3$ and set

$$[\mathbf{x}, \mathbf{y}] = \mathbf{x} \times \mathbf{y},$$

the *vector* or *cross product*. For the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$,

$$[\mathbf{e}_1, \mathbf{e}_2] = -[\mathbf{e}_2, \mathbf{e}_1] = \mathbf{e}_3, \quad [\mathbf{e}_2, \mathbf{e}_3] = -[\mathbf{e}_3, \mathbf{e}_2] = \mathbf{e}_1, \quad [\mathbf{e}_3, \mathbf{e}_1] = -[\mathbf{e}_1, \mathbf{e}_3] = \mathbf{e}_2. \tag{2.2}$$

Then $\mathbb{R}^3$ equipped with this bracket operation is an $\mathbb{R}$-Lie algebra. In fact, as we will see later, this is the Lie algebra of $\mathrm{SO}(3)$ and also of $\mathrm{SU}(2)$ in disguise.

Given two matrices $A, B \in \mathrm{M}_n(\Bbbk)$, their *commutator* is

$$[A, B] = AB - BA.$$

This is a $\Bbbk$-bilinear function $\mathrm{M}_n(\Bbbk) \times \mathrm{M}_n(\Bbbk) \longrightarrow \mathrm{M}_n(\Bbbk)$ satisfying the conditions of Definition 2.11. Recall that A, B *commute* if $AB = BA$.

PROPOSITION 2.13. $[A, B] = O_n$ *if and only if* A, B *commute.*

Suppose that $\mathfrak{a}$ is a $\Bbbk$-vector subspace of $\mathrm{M}_n(\Bbbk)$. Then $\mathfrak{a}$ is a $\Bbbk$-*Lie subalgebra* of $\mathrm{M}_n(\Bbbk)$ if it is closed under taking commutators of pairs of elements in $\mathfrak{a}$, i.e., if $A, B \in \mathfrak{a}$ then $[A, B] \in \mathfrak{a}$. Of course $\mathrm{M}_n(\Bbbk)$ is a $\Bbbk$-Lie subalgebra of itself.

THEOREM 2.14. *For $\Bbbk = \mathbb{R}$ or $\Bbbk = \mathbb{C}$, if $G \leqslant \mathrm{GL}_n(\Bbbk)$ is a matrix subgroup, then $\mathfrak{g}$ is an $\mathbb{R}$-Lie subalgebra of $\mathrm{M}_n(\Bbbk)$.*

If $G \leqslant \mathrm{GL}_m(\mathbb{C})$ is a matrix subgroup and $\mathfrak{g}$ is a $\mathbb{C}$-subspace of $\mathrm{M}_m(\mathbb{C})$, then $\mathfrak{g}$ is a $\mathbb{C}$-Lie subalgebra.

PROOF. We will show that for two differentiable curves α, β in G with $\alpha(0) = \beta(0) = I_n$, there is such a curve γ with $\gamma'(0) = [\alpha'(0), \beta'(0)]$.

Consider the function

$$F\colon \operatorname{dom}\alpha \times \operatorname{dom}\beta \longrightarrow G; \quad F(s,t) = \alpha(s)\beta(t)\alpha(s)^{-1}.$$

This is clearly continuous and differentiable with respect to each of the variables s, t. For each $s \in \operatorname{dom}\alpha$, $F(s, \ )$ is a differentiable curve in G with $F(s, 0) = I_n$. Differentiating gives

$$\frac{\mathrm{d}\, F(s,t)}{\mathrm{d}\, t}_{|t=0} = \alpha(s)\beta'(0)\alpha(s)^{-1},$$

and so

$$\alpha(s)\beta'(0)\alpha(s)^{-1} \in \mathfrak{g}.$$

Since $\mathfrak{g}$ is a closed subspace of $\mathrm{M}_n(\Bbbk)$, whenever this limit exists we also have

$$\lim_{s\to 0} \frac{1}{s}\left(\alpha(s)\beta'(0)\alpha(s)^{-1} - \beta'(0)\right) \in \mathfrak{g}.$$

We will use the following easily verified matrix version of the usual rule for differentiating an inverse:

$$\frac{\mathrm{d}}{\mathrm{d}\, t}\left(\alpha(t)^{-1}\right) = -\alpha(t)^{-1}\alpha'(t)\alpha(t)^{-1}. \tag{2.3}$$

We have

$$\begin{aligned}
\lim_{s\to 0} \frac{1}{s}\left(\alpha(s)\beta'(0)\alpha(s)^{-1} - \beta'(0)\right) =& \frac{\mathrm{d}}{\mathrm{d}\, s}_{|s=0} \alpha(s)\beta'(0)\alpha(s)^{-1} \\
=& \alpha'(0)\beta'(0)\alpha(0) - \alpha(0)\beta'(0)\alpha(0)^{-1}\alpha'(0)\alpha(0)^{-1} \\
& \qquad\qquad\qquad\qquad [\text{by Equation (2.3)}] \\
=& \alpha'(0)\beta'(0)\alpha(0) - \alpha(0)\beta'(0)\alpha'(0) \\
=& \alpha'(0)\beta'(0) - \beta'(0)\alpha'(0) \\
=& [\alpha'(0), \beta'(0)].
\end{aligned}$$

This shows that $[\alpha'(0), \beta'(0)] \in \mathfrak{g}$, hence it must be of the form $\gamma'(0)$ for some differentiable curve. The second part follows easily. $\square$

So for each matrix group G there is a Lie algebra $\mathfrak{g} = \mathrm{T}_I\, G$. A suitable type of homomorphism $G \longrightarrow H$ between matrix groups gives rise to a linear transformation $\mathfrak{g} \longrightarrow \mathfrak{h}$ respecting the Lie algebra structures.

DEFINITION 2.15. Let $G \leqslant \mathrm{GL}_n(\Bbbk)$, $H \leqslant \mathrm{GL}_m(\Bbbk)$ be matrix groups and $\varphi\colon G \longrightarrow H$ a continuous map. Then φ is said to be a *differentiable map* if for every differentiable curve $\gamma\colon (a, b) \longrightarrow G$, the composite curve $\varphi\circ\gamma\colon (a, b) \longrightarrow H$ is differentiable, with derivative

$$(\varphi\circ\gamma)'(t) = \frac{\mathrm{d}}{\mathrm{d}\, t}\varphi(\gamma(t)),$$

and if whenever two differentiable curves $\alpha, \beta\colon (a,b) \longrightarrow G$ both satisfy the conditions

$$\alpha(0) = \beta(0), \quad \alpha'(0) = \beta'(0),$$

then

$$(\varphi \circ \alpha)'(0) = (\varphi \circ \beta)'(0).$$

Such a φ is a *differentiable homomorphism* if it is also a group homomorphism. A continuous homomorphism of matrix groups that is also a differentiable map is called a *Lie homomorphism.*

We will see later that the technical restriction in this definition is unnecessary. For now we note that if $\varphi\colon G \longrightarrow H$ is the restriction of a differentiable map $\Phi\colon \mathrm{GL}_n(\Bbbk) \longrightarrow \mathrm{GL}_m(\Bbbk)$ then φ is also a differentiable map.

PROPOSITION 2.16. *Let G, H, K be matrix groups and $\varphi\colon G \longrightarrow H$, $\theta\colon H \longrightarrow K$ be differentiable homomorphisms.*

a) *For each $A \in G$ there is an $\mathbb{R}$-linear transformation $\mathrm{d}\,\varphi\colon \mathrm{T}_A\, G \longrightarrow \mathrm{T}_{\varphi(A)}\, H$ given by*

$$\mathrm{d}\,\varphi_A(\gamma'(0)) = (\varphi \circ \gamma)'(0),$$

for every differentiable curve $\gamma\colon (a,b) \longrightarrow G$ with $\gamma(0) = A$.

b) *We have*

$$\mathrm{d}\,\theta_{\varphi(A)} \circ \mathrm{d}\,\varphi_A = \mathrm{d}(\theta \circ \varphi)_A.$$

c) *For the identity map $\mathrm{Id}_G\colon G \longrightarrow G$ and $A \in G$,*

$$\mathrm{d}\,\mathrm{Id}_G = \mathrm{Id}_{\mathrm{T}_A\, G}\,.$$

PROOF. a) The definition of $\mathrm{d}\,\varphi_A$ makes sense since by the definition of differentiability, given $X \in \mathrm{T}_A\, G$, for any curve γ with

$$\gamma(0) = A, \quad \gamma'(0) = X,$$

$(\varphi \circ \gamma)'(0)$ depends only on X and not on γ. Linearity is established using similar ideas to the proof of Proposition 2.9.

The identities of (b) and (c) are straightforward to verify. □

If $\varphi\colon G \longrightarrow H$ is a differentiable homomorphism then since $\varphi(I) = I$, $\mathrm{d}\,\varphi_I\colon \mathrm{T}_I\, G \longrightarrow \mathrm{T}_I\, H$ is a linear transformation called the *derivative* of φ which will usually be denoted

$$\mathrm{d}\,\varphi\colon \mathfrak{g} \longrightarrow \mathfrak{h}.$$

DEFINITION 2.17. Let $\mathfrak{g}, \mathfrak{h}$ be Lie algebras over a field $\Bbbk$. A $\Bbbk$-linear transformation $\Phi\colon \mathfrak{g} \longrightarrow \mathfrak{h}$ is a *homomorphism of Lie algebras* if

$$\Phi([x,y]) = [\Phi(x), \Phi(y)] \quad (x, y \in \mathfrak{g}).$$

THEOREM 2.18. *Let G, H be matrix groups and $\varphi\colon G \longrightarrow H$ a differentiable homomorphism. Then the derivative $\mathrm{d}\,\varphi\colon \mathfrak{g} \longrightarrow \mathfrak{h}$ is a homomorphism of Lie algebras.*

PROOF. Following ideas and notation in the proof of Theorem 2.14, for differentiable curves α, β in G with $\alpha(0) = \beta(0) = I$, we can use the composite function $\varphi \circ F$ given by

$$\varphi \circ F(s,t) = \varphi(F(s,t)) = \varphi(\alpha(s))\varphi(\beta(t))\varphi(\alpha(s))^{-1},$$

to deduce

$$\mathrm{d}\,\varphi([\alpha'(0), \beta'(0)]) = [\mathrm{d}\,\varphi(\alpha'(0)), \mathrm{d}\,\varphi(\beta'(0))]).$$

□

4. Some Lie algebras of matrix groups

The Lie algebras of $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{C})$. Let us start with the matrix group $\mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{M}_n(\mathbb{R})$. For $A \in \mathrm{M}_n(\mathbb{R})$ and $\varepsilon > 0$ there is a differentiable curve

$$\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{M}_n(\mathbb{R}); \quad \alpha(t) = I + tA.$$

For $t \neq 0$, the roots of the equation $\det(t^{-1}I + A) = 0$ are of the form $t = -1/\lambda$ where λ is a non-zero eigenvalue of A. Hence if

$$\varepsilon < \min \left\{ \frac{1}{|\lambda|} : \lambda \text{ a non-zero eigenvalue of } A \right\},$$

then $\operatorname{im} \alpha \subseteq \mathrm{GL}_n(\mathbb{R})$, so we might as well view α a function $\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{GL}_n(\mathbb{R})$. Calculating the derivative we find that $\alpha'(t) = A$, hence $\alpha'(0) = A$. This shows that $A \in \mathrm{T}_I \mathrm{GL}_n(\mathbb{R})$. Since $A \in \mathrm{M}_n(\mathbb{R})$ was arbitrary, we have

$$\begin{cases} \mathfrak{gl}_n(\mathbb{R}) = \mathrm{T}_I \mathrm{GL}_n(\mathbb{R}) = \mathrm{M}_n(\mathbb{R}), \\ \dim \mathrm{GL}_n(\mathbb{R}) = n^2. \end{cases} \tag{2.4}$$

Similarly,

$$\begin{cases} \mathfrak{gl}_n(\mathbb{C}) = \mathrm{T}_I \mathrm{GL}_n(\mathbb{C}) = \mathrm{M}_n(\mathbb{C}), \\ \dim_{\mathbb{C}} \mathrm{GL}_n(\mathbb{C}) = n^2, \\ \dim \mathrm{GL}_n(\mathbb{C}) = 2n^2. \end{cases} \tag{2.5}$$

For $\mathrm{SL}_n(\mathbb{R}) \leqslant \mathrm{GL}_n(\mathbb{R})$, suppose that $\alpha\colon (a, b) \longrightarrow \mathrm{SL}_n(\mathbb{R})$ is a curve lying in $\mathrm{SL}_n(\mathbb{R})$ and satisfying $\alpha(0) = I$. For $t \in (a, b)$ we have $\det \alpha(t) = 1$, so

$$\frac{\mathrm{d}(\det \alpha(t))}{\mathrm{d}\, t} = 0.$$

Lemma 2.19. *We have*

$$\frac{\mathrm{d}(\det \alpha(t))}{\mathrm{d}\, t}_{|t=0} = \operatorname{tr} \alpha'(0).$$

Proof. Recall that for $A \in \mathrm{M}_n(\Bbbk)$,

$$\operatorname{tr} A = \sum_{i=1}^{n} A_{ii}.$$

It is easy to verify that the operation $\partial = \dfrac{\mathrm{d}}{\mathrm{d}\, t}_{|t=0}$ on functions has the *derivation property*

$$\partial(\gamma_1 \gamma_2) = (\partial \gamma_1)\gamma_2(0) + \gamma_1(0) \partial \gamma_2. \tag{2.6}$$

Put $a_{ij} = \alpha(t)_{ij}$ and notice that when $t = 0$,

$$a_{ij} = \delta_{ij}.$$

Write C_{ij} for the cofactor matrix obtained from $\alpha(t)$ by deleting the i th row and j th column. By expanding along the n th row we obtain

$$\det \alpha(t) = \sum_{j=1}^{n} (-1)^{n+j} a_{nj} \det C_{nj}$$

Then

$$\begin{aligned} \partial \det \alpha(t) &= \sum_{j=1}^{n} (-1)^{n+j} \left((\partial a_{nj}) \det C_{nj} + a_{nj} (\partial \det C_{nj}) \right) \\ &= \sum_{j=1}^{n} (-1)^{n+j} \left((\partial a_{nj}) \det C_{nj} \right) + (\partial \det C_{nn}). \end{aligned}$$

For $t = 0$, $\det C_{nj} = \delta_{jn}$ since $\alpha(0) = I$, hence

$$\partial \det(\alpha(t)) = \partial a_{nn} + \partial \det C_{nn}.$$

We can repeat this calculation with the $(n-1) \times (n-1)$ matrix C_{nn} and so on. This gives

$$\begin{aligned}\partial \det(\alpha(t)) &= \partial a_{nn} + \partial a_{(n-1)(n-1)} + \partial \det C_{(n-1)(n-1)} \\ &\vdots \\ &= \partial a_{nn} + \partial a_{(n-1)(n-1)} + \cdots + \partial a_{11)} \\ &= \operatorname{tr} \alpha'(0).\end{aligned}$$

□

So we have $\operatorname{tr} \alpha'(0) = 0$ and hence

$$\mathfrak{sl}_n(\mathbb{R}) = \mathrm{T}_I \operatorname{SL}_n(\mathbb{R}) \subseteq \ker \operatorname{tr} \subseteq \mathrm{M}_n(\mathbb{R}).$$

If $A \in \ker \operatorname{tr} \subseteq \mathrm{M}_n(\mathbb{R})$, the function

$$\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{M}_n(\mathbb{R}); \quad \alpha(t) = \exp(tA) = \sum_{k \geqslant 0} \frac{t^k}{k!} A^k,$$

is defined for every $\varepsilon > 0$ and satisfies the boundary conditions

$$\alpha(0) = I, \quad \alpha'(0) = A.$$

We will use the following result for which another proof appears in Chapter 4, Section 5.

Lemma 2.20. *For $A \in \mathrm{M}_n(\mathbb{C})$ we have*

$$\det \exp(A) = e^{\operatorname{tr} A}.$$

Proof using differential equations. Consider the curve

$$\gamma\colon \mathbb{R} \longrightarrow \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^\times; \quad \gamma(t) = \det \exp(tA).$$

Then

$$\begin{aligned}\gamma'(t) &= \lim_{h \to 0} \frac{1}{h} (\det \exp((t+h)A) - \det \exp(tA)) \\ &= \det \exp(tA) \lim_{h \to 0} \frac{1}{h} (\det \exp(hA) - 1) \\ &= \det \exp(tA) \operatorname{tr} A \\ &= \gamma \operatorname{tr} A\end{aligned}$$

by Lemma 2.19 applied to the curve $t \mapsto \det \exp(tA)$. So α satisfies the same differential equation and initial condition as the curve $t \mapsto e^{t \operatorname{tr} A}$. By the uniqueness part of Theorem 2.1,

$$\alpha(t) = \det \exp(tA) = e^{t \operatorname{tr} A}.$$

□

Proof using Jordan Canonical Form. If $S \in \mathrm{GL}_n(\mathbb{C})$,

$$\begin{aligned}\det \exp(SAS^{-1}) &= \det \left(S \exp(A) S^{-1}\right) \\ &= \det S \det \exp(A) \det S^{-1} \\ &= \det \exp A,\end{aligned}$$

and

$$e^{\operatorname{tr} SAS^{-1}} = e^{\operatorname{tr} A}.$$

So it suffices to prove the identity for SAS^{-1} for a suitably chosen invertible matrix S. Using for example the theory of Jordan Canonical Forms, there is a suitable choice of such an S for which

$$B = SAS^{-1} = D + N,$$

with D diagonal, N strictly upper triangular and $N_{ij} = 0$ whenever $i \geqslant j$. Then N is nilpotent, i.e., $N^k = O_n$ for large k.

We have

$$\begin{aligned}
\exp(B) &= \sum_{k\geqslant 0} \frac{1}{k!}(D+N)^k \\
&= \left(\sum_{k\geqslant 0} \frac{1}{k!} D^k\right) + \sum_{k\geqslant 0} \frac{1}{(k+1)!}\left((D+N)^{k+1} - D^{k+1}\right) \\
&= \exp(D) + \sum_{k\geqslant 0} \frac{1}{(k+1)!} N(D^k + D^{k-1}N + \cdots + N^k).
\end{aligned}$$

Now for $k \geqslant 0$, the matrix

$$N(D^k + D^{k-1}N + \cdots + N^k)$$

is strictly upper triangular, hence

$$\exp(B) = \exp(D) + N',$$

where N' is strictly upper triangular. If $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, on calculating the determinant we find that

$$\begin{aligned}
\det \exp(A) &= \det \exp(B) \\
&= \det \exp(D) \\
&= \det \operatorname{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}) \\
&= e^{\lambda_1} \cdots e^{\lambda_n} \\
&= e^{\lambda_1 + \cdots + \lambda_n}.
\end{aligned}$$

Since $\operatorname{tr} D = \lambda_1 + \cdots + \lambda_n$, this implies

$$\det \exp(A) = e^{\operatorname{tr} D}.$$

□

Using this Lemma and the function α, we obtain

$$\begin{cases} \mathfrak{sl}_n(\mathbb{R}) = \mathrm{T}_I \mathrm{SL}_n(\mathbb{R}) = \ker \operatorname{tr} \subseteq \mathrm{M}_n(\mathbb{R}), \\ \dim \mathrm{SL}_n(\mathbb{R}) = n^2 - 1. \end{cases} \tag{2.7}$$

Working over $\mathbb{C}$ we also have

$$\begin{cases} \mathfrak{sl}_n(\mathbb{C}) = \mathrm{T}_I \mathrm{SL}_n(\mathbb{C}) = \ker \operatorname{tr} \subseteq \mathrm{M}_n(\mathbb{C}), \\ \dim_{\mathbb{C}} \mathrm{SL}_n(\mathbb{C}) = n^2 - 1, \\ \dim \mathrm{SL}_n(\mathbb{C}) = 2n^2 - 2. \end{cases} \tag{2.8}$$

The Lie algebras of $\mathrm{UT}_n(\Bbbk)$ **and** $\mathrm{SUT}_n(\Bbbk)$**.** For $n \geqslant 1$ and $\Bbbk = \mathbb{R}, \mathbb{C}$, recall the upper triangular and unipotent subgroups of $\mathrm{GL}_n(\Bbbk)$. Let

$$\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{UT}_n(\mathbb{R})$$

be a differentiable curve with $\alpha(0) = I$. Then $\alpha'(t)$ is upper triangular. Moreover, using the argument for $\mathrm{GL}_n(\Bbbk)$ we see that given any upper triangular matrix $A \in \mathrm{M}_n(\Bbbk)$, there is a curve

$$\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{UT}_n(\Bbbk); \quad \alpha(t) = I + tA,$$

where $\varepsilon > 0$ has to be chosen small and $\alpha'(0) = A$. We then have

$$\begin{cases} \mathfrak{ut}_n(\Bbbk) = \mathrm{T}_I \, \mathrm{UT}_n(\Bbbk) = \text{set of all upper triangular matrices in } \mathrm{M}_n(\Bbbk), \\ \dim \mathfrak{ut}_n(\Bbbk) = \dbinom{n+1}{2} \dim_{\mathbb{R}} \Bbbk. \end{cases} \tag{2.9}$$

An upper triangular matrix $A \in \mathrm{M}_n(\Bbbk)$ is *strictly upper triangular* if all its diagonal entries are 0, i.e., $a_{ii} = 0$. Then

$$\begin{cases} \mathfrak{sut}_n(\Bbbk) = \mathrm{T}_I \, \mathrm{SUT}_n(\Bbbk) = \text{set of all strictly upper triangular matrices in } \mathrm{M}_n(\Bbbk), \\ \dim \mathfrak{sut}_n(\Bbbk) = \dbinom{n}{2} \dim_{\mathbb{R}} \Bbbk. \end{cases} \tag{2.10}$$

The Lie algebras of $\mathrm{O}(n)$ **and** $\mathrm{SO}(n)$**.** Let $\mathrm{O}(n)$ be the $n \times n$ orthogonal group, i.e.,

$$\mathrm{O}(n) = \{A \in \mathrm{GL}_n(\mathbb{R}) : A^T A = I\} \leqslant \mathrm{GL}_n(\mathbb{R}).$$

Given a curve $\alpha\colon (a, b) \longrightarrow \mathrm{O}(n)$ satisfying $\alpha(0) = I$ we have

$$\frac{\mathrm{d}}{\mathrm{d}\, t} \alpha(t)^T \alpha(t) = O,$$

and so

$$\alpha'(t)^T \alpha(t) + \alpha(t)^T \alpha'(t) = O,$$

implying

$$\alpha'(0)^T + \alpha'(0) = O.$$

Thus we must have $\alpha'(0)^T = -\alpha'(0)$, i.e., $\alpha'(0)$ is skew symmetric. Thus

$$\mathfrak{o}(n) = \mathrm{T}_I \, \mathrm{O}(n) \subseteq \text{Sk-Sym}_n(\mathbb{R}),$$

the set of $n \times n$ real skew symmetric matrices.

On the other hand, if $A \in \text{Sk-Sym}_n(\mathbb{R})$, for $\varepsilon > 0$ we can consider the curve

$$\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{GL}_n(\mathbb{R}); \quad \alpha(t) = \exp(tA).$$

Then

$$\begin{aligned} \alpha(t)^T \alpha(t) &= \exp(tA)^T \exp(tA) \\ &= \exp(tA^T) \exp(tA) \\ &= \exp(-tA) \exp(tA) \\ &= I. \end{aligned}$$

Hence we can view α as a curve $\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{O}(n)$. Since $\alpha'(0) = A$, this shows that

$$\text{Sk-Sym}_n(\mathbb{R}) \subseteq \mathfrak{o}(n) = \mathrm{T}_I \, \mathrm{O}(n)$$

and so

$$\mathfrak{o}(n) = \mathrm{T}_I \, \mathrm{O}(n) = \text{Sk-Sym}_n(\mathbb{R}).$$

Notice that if $A \in \text{Sk-Sym}_n(\mathbb{R})$ then

$$\operatorname{tr} A = \operatorname{tr} A^T = \operatorname{tr}(-A) = -\operatorname{tr} A,$$

hence $\operatorname{tr} A = 0$. By Lemma 2.20 we have

$$\det \exp(tA) = 1,$$

hence $\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{SO}(n)$ where $\mathrm{SO}(n)$ is the $n \times n$ special orthogonal group. So we have actually shown that

$$\mathfrak{so}(n) = \mathrm{T}_I \, \mathrm{SO}(n) = \mathfrak{o}(n) = \mathrm{T}_I \, \mathrm{O}(n) = \text{Sk-Sym}_n(\mathbb{R}).$$

The Lie algebras of $\mathrm{U}(n)$ **and** $\mathrm{SU}(n)$**.** Now consider the $n \times n$ unitary group

$$\mathrm{U}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) : A^*A = I\}.$$

For a curve α in $\mathrm{U}(n)$ satisfying $\alpha(0) = I$, we obtain

$$\alpha'(0)^* + \alpha'(0) = 0$$

and so $\alpha'(0)^* = -\alpha'(0)$, i.e., $\alpha(0)$ is skew hermitian. So

$$\mathfrak{u}(n) = \mathrm{T}_I \, \mathrm{U}(n) \subseteq \text{Sk-Herm}_n(\mathbb{C}),$$

the set of all $n \times n$ skew hermitian matrices.

If $H \in \text{Sk-Herm}_n(\mathbb{C})$ then the curve

$$\eta\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{GL}_n(\mathbb{C}); \quad \eta(t) = \exp(tH)$$

satisfies

$$\begin{aligned}
\eta(t)^*\eta(t) &= \exp(tH)^* \exp(tH) \\
&= \exp(tH^*) \exp(tH) \\
&= \exp(-tH) \exp(tH) \\
&= I.
\end{aligned}$$

Hence we can view η as a curve $\eta\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{U}(n)$. Since $\eta'(0) = H$, this shows that

$$\text{Sk-Herm}_n(\mathbb{C}) \subseteq \mathfrak{u}(n) = \mathrm{T}_I \, \mathrm{U}(n).$$

Hence

$$\mathfrak{u}(n) = \mathrm{T}_I \, \mathrm{U}(n) \subseteq \text{Sk-Herm}_n(\mathbb{C}).$$

The special unitary group $\mathrm{SU}(n)$ can be handled in a similar way. Again we have

$$\mathfrak{su}(n) = \mathrm{T}_I \, \mathrm{SU}(n) \subseteq \text{Sk-Herm}_n(\mathbb{C}).$$

But also if $\eta\colon (a, b) \longrightarrow \mathrm{SU}(n)$ is a curve with $\eta(0) = I$ then as in the analysis for $\mathrm{SL}_n(\mathbb{R})$,

$$\operatorname{tr} \eta'(0) = 0.$$

Writing

$$\text{Sk-Herm}^0_n(\mathbb{C}) = \{H \in \text{Sk-Herm}_n(\mathbb{C}) : \operatorname{tr} H = 0\},$$

this gives $\mathfrak{su}(n) \subseteq \text{Sk-Herm}^0_n(\mathbb{C})$. On the other hand, if $H \in \text{Sk-Herm}^0_n(\mathbb{C})$ then the curve

$$\eta\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{U}(n); \quad \eta(t) = \exp(tH),$$

takes values in $\mathrm{SU}(n)$ by Lemma 2.20 and has $\eta'(0) = H$. Hence

$$\mathfrak{su}(n) = \mathrm{T}_I \, \mathrm{SU}(n) \subseteq \text{Sk-Herm}^0_n(\mathbb{C}).$$

REMARK 2.21. Later, we will see that for a matrix group $G \leqslant \mathrm{GL}_n(\mathbb{R})$, the following are true and can be used in determining Lie algebras of matrix groups as above.

- The function
$$\exp_G\colon \mathfrak{g} \longrightarrow \mathrm{GL}_n(\mathbb{R}); \quad \exp_G(X) = \exp(X),$$
has image contained in G, $\exp_G \mathfrak{g} \subseteq G$; so we will normally write $\exp_G\colon \mathfrak{g} \longrightarrow G$ for the exponential on G and sometimes even just exp.
- If G is compact and connected then $\exp_G \mathfrak{g} = G$.
- There is an open disc $\mathrm{N}_{\mathfrak{g}}(O; r) \subseteq \mathfrak{g}$ on which exp is injective and gives a homeomorphism $\exp\colon \mathrm{N}_{\mathfrak{g}}(O; r) \longrightarrow \exp \mathrm{N}_{\mathfrak{g}}(O; r)$ where $\exp \mathrm{N}_{\mathfrak{g}}(O; r) \subseteq G$ is in fact an open subset.

5. SO(3) and SU(2)

In this section we will discuss the groups $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$ and their Lie algebras in detail. The Lie algebras are both 3-dimensional real vector spaces, having for example the following bases:

$$\mathfrak{so}(3)\colon \quad P = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$$
$$\mathfrak{su}(2)\colon \quad H = \frac{1}{2}\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad E = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad F = \frac{1}{2}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

The non-trivial Lie brackets are then

$$[P, Q] = R, \qquad [Q, R] = P, \qquad [R, P] = Q, \tag{2.11a}$$
$$[H, E] = F, \qquad [E, F] = H, \qquad [F, H] = E. \tag{2.11b}$$

This means that the $\mathbb{R}$-linear isomorphism

$$\varphi\colon \mathfrak{su}(2) \longrightarrow \mathfrak{so}(3); \quad \varphi(xH + yE + zF) = xP + yQ + zR \quad (x, y, z \in \mathbb{R}), \tag{2.12}$$

satisfies

$$\varphi([U, V]) = [\varphi(U), \varphi(V)],$$

hence is an *isomorphism of* $\mathbb{R}$*-Lie algebras*. Thus these Lie algebras look the same algebraically. This suggests that there might be a close relationship between the groups themselves. Before considering this, notice also that for the Lie algebra of Example 2.12, the $\mathbb{R}$-linear transformation

$$\mathbb{R}^3 \longrightarrow \mathfrak{so}(3); \quad x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 \longmapsto xP + yQ + zR,$$

is an isomorphism of $\mathbb{R}$-Lie algebras by Formulæ (2.2).

Now we will construct a Lie homomorphism $\mathrm{SU}(2) \longrightarrow \mathrm{SO}(3)$ whose derivative at I is φ. Recall the *adjoint action* of Ad of $\mathrm{SU}(2)$ on $\mathfrak{su}(2)$ by

$$\mathrm{Ad}_A(U) = AUA^{-1} = AUA^* \quad (A \in \mathrm{SU}(2),\ U \in \mathfrak{su}(2)).$$

Then each Ad_A is an $\mathbb{R}$-linear isomorphism $\mathfrak{su}(2) \longrightarrow \mathfrak{su}(2)$.

We can define a *real inner product* $(\ \mid\ )$ on $\mathfrak{su}(2)$ by

$$(X \mid Y) = -\operatorname{tr}(XY) \quad (X, Y \in \mathfrak{su}(2)).$$

Introducing the elements

$$\hat{H} = \sqrt{2}H = \frac{1}{\sqrt{2}}\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad \hat{E} = \sqrt{2}E = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \hat{F} = \sqrt{2}F = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix},$$

we obtain an $\mathbb{R}$-linear isomorphism

$$\theta\colon \mathbb{R}^3 \longrightarrow \mathfrak{su}(2); \quad \theta(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3) = x\hat{H} + y\hat{E} + z\hat{F}, \tag{2.13}$$

which is an isometry since $\hat{H}, \hat{E}, \hat{F}$ form an orthonormal basis of $\mathfrak{su}(2)$ with respect to $(\quad|\quad)$, i.e.,

$$(\hat{H} \mid \hat{H}) = (\hat{E} \mid \hat{E}) = (\hat{F} \mid \hat{F}) = 1, \tag{2.14a}$$

$$(\hat{H} \mid \hat{E}) = (\hat{H} \mid \hat{F}) = (\hat{E} \mid \hat{F}) = 0. \tag{2.14b}$$

Remark 2.22. It would perhaps be more natural to rescale the inner product $(\ |\ )$ so that H, E, F were all unit vectors. This would certainly make many of the following formulæ neater as well making the Lie bracket in SU(2) correspond exactly with the vector product in $\mathbb{R}^3$. However, our choice of $(\ |\ )$ agrees with the conventional one for SU(n).

Proposition 2.23. *$(\ |\ )$ is a real symmetric bilinear form on $\mathfrak{su}(2)$ which is positive definite. It is* invariant *in the sense that*

$$([Z, X] \mid Y) + (X \mid [Z, Y]) = 0 \quad (X, Y, Z \in \mathfrak{su}(2)).$$

Proof. The $\mathbb{R}$-bilinearity is clear, as is the symmetry. For positive definiteness, notice that for $x, x', y, y', z, z' \in \mathbb{R}$,

$$(x\hat{H} + y\hat{E} + z\hat{F} \mid x'\hat{H} + y'\hat{E} + z'\hat{F}) = xx' + yy' + zz'$$

and in particular,

$$(x\hat{H} + y\hat{E} + z\hat{F} \mid x\hat{H} + y\hat{E} + z\hat{F}) = x^2 + y^2 + z^2 \geqslant 0,$$

with equality precisely when $x = y = z = 0$.

The invariance is checked by a calculation. □

Also, for $A \in \mathrm{SU}(2)$ and $X, Y \in \mathfrak{su}(2)$,

$$\begin{aligned}
(AXA^* \mid AYA^*) &= -\operatorname{tr}(AXA^*AYA^*) \\
&= -\operatorname{tr}(AXYA^*) \\
&= -\operatorname{tr}(AXYA^{-1}) \\
&= -\operatorname{tr}(XY) \\
&= (X \mid Y),
\end{aligned}$$

hence Ad_A is actually an orthogonal linear transformation with respect to this inner product. Using the orthonormal basis $\hat{H}, E, \hat{F}$, we can identify $\mathfrak{su}(2)$ with $\mathbb{R}^3$ and $(\ |\ )$ with the usual inner product $\cdot$, then each Ad_A corresponds to an element of O(3) which we will still write as Ad_A. It is then easy to see that the function

$$\overline{\mathrm{Ad}}\colon \mathrm{SU}(2) \longrightarrow \mathrm{O}(3); \quad \overline{\mathrm{Ad}}(A) = \mathrm{Ad}_A \in \mathrm{O}(3),$$

is a continuous homomorphism of groups. In fact, SU(2) is path connected, as is SO(3); so since $\overline{\mathrm{Ad}}(I) = I$,

$$\overline{\mathrm{Ad}}\,\mathrm{SU}(2) \subseteq \mathrm{SO}(3),$$

hence we will redefine

$$\overline{\mathrm{Ad}}\colon \mathrm{SU}(2) \longrightarrow \mathrm{SO}(3); \quad \overline{\mathrm{Ad}}(A) = \mathrm{Ad}_A .$$

PROPOSITION 2.24. *The continuous homomorphism of matrix groups*

$$\overline{\mathrm{Ad}}\colon \mathrm{SU}(2) \longrightarrow \mathrm{SO}(3);\quad \overline{\mathrm{Ad}}(A) = \mathrm{Ad}_A,$$

is smooth, has $\ker \overline{\mathrm{Ad}} = \{\pm I\}$ *and is surjective.*

PROOF. The identification of the kernel is an easy exercise. The remaining statements can be proved using ideas from Chapter 4, especially Section 1. We will give a direct proof that $\ker \overline{\mathrm{Ad}}$ is surjective to illustrate some important special geometric aspects of this example.

We can view an element of $\mathfrak{su}(2)$ as a vector in $\mathbb{R}^3$ by identifying the orthonormal basis vectors $\hat{H}$, $\hat{E}$, $\hat{F}$ with $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. From Equations (2.11), the non-trivial brackets of these basis elements are as follows:

$$[\hat{H},\hat{E}] = \sqrt{2}\hat{F},\quad [\hat{E},\hat{F}] = \sqrt{2}\hat{H},\quad [\hat{F},\hat{H}] = \sqrt{2}\hat{E}. \tag{2.15}$$

So apart from the factors of $\sqrt{2}$, this behaves exactly like the vector product on $\mathbb{R}^3$.

PROPOSITION 2.25. *For* $U_1 = x_1\hat{H} + y_1\hat{E} + z_1\hat{F}$, $U_2 = x_2\hat{H} + y_2\hat{E} + z_2\hat{F} \in \mathfrak{su}(2)$,

$$[U_1,U_2] = \sqrt{2}\left(\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}\hat{H} - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}\hat{E} + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\hat{F}\right).$$

PROOF. This follows from the formula

$$\begin{aligned} x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 &= (x_1\mathbf{e}_1 + y_1\mathbf{e}_2 + z_1\mathbf{e}_3) \times (x_2\mathbf{e}_1 + y_2\mathbf{e}_2 + z_2\mathbf{e}_3) \\ &= \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\mathbf{e}_3. \end{aligned}$$

$\square$

We can similarly calculate a product of elements of $\mathfrak{su}(2)$ in terms of the dot and cross products. However, note that in general if $U_1, U_2 \in \mathfrak{su}(2)$ then $U_1U_2 \notin \mathfrak{su}(2)$.

PROPOSITION 2.26. $U_1 = x_1\hat{H} + y_1\hat{E} + z_1\hat{F}$, $U_2 = x_2\hat{H} + y_2\hat{E} + \hat{F} \in \mathfrak{su}(2)$,

$$\begin{aligned} U_1U_2 &= -\frac{(x_1x_2 + y_1y_2 + z_1z_2)}{2}I + \frac{1}{\sqrt{2}}\left(\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}\hat{H} - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}\hat{E} + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\hat{F}\right) \\ &= -\frac{(U_1 \mid U_2)}{2}I + \frac{1}{2}[U_1,U_2]. \end{aligned}$$

PROOF. Calculation! $\square$

COROLLARY 2.27. *If* $U_1, U_2 \in \mathfrak{su}(2)$ *are orthogonal, i.e.,* $(U_1 \mid U_2) = 0$, *then*

$$U_1U_2 = \frac{1}{2}[U_1,U_2] \in \mathfrak{su}(2).$$

Next we will examine the effect of $A \in \mathrm{SU}(2)$ acting as an $\mathbb{R}$-linear transformation on $\mathfrak{su}(2)$ which we will identify with $\mathbb{R}^3$. Note that A can be uniquely written as

$$A = \begin{bmatrix} u & v \\ -\overline{v} & \overline{u} \end{bmatrix} \tag{2.16}$$

for $u, v \in \mathbb{C}$ and $|u|^2 + |v|^2 = 1$. This allows us to express A in the form

$$A = \cos\theta I + S,$$

where S is skew hermitian and $\operatorname{Re} u = \cos\theta$ for $\theta \in [0,\pi]$, so $\sin\theta \geqslant 0$. A calculation gives

$$S^2 = -((\operatorname{Im} u)^2 + |v|^2)I = -\sin^2\theta I, \tag{2.17a}$$

$$(S \mid S) = 2\sin^2\theta. \tag{2.17b}$$

Since $A \in \mathrm{SU}(2)$, we have

$$A^{-1} = A^* = \cos\theta I - S.$$

Notice that for any $t \in \mathbb{R}$,

$$\mathrm{Ad}_A(tS) = A(tS)A^{-1} = tS.$$

On the other hand, if $U \in \mathfrak{su}(2)$ with $(S \mid U) = 0$, then by the above results,

$$\begin{aligned}
\mathrm{Ad}_A(U) &= (\cos\theta I + S)U(\cos\theta I - S) \\
&= (\cos\theta U + SU)(\cos\theta I - S) \\
&= \cos^2\theta U + \cos\theta SU - \cos\theta US - SUS \\
&= \cos^2\theta U + \cos\theta[S, U] - SUS.
\end{aligned}$$

A further calculation using properties of the vector product shows that

$$SUS = \frac{(S \mid S)}{2}U.$$

By Equation (2.17b), whenever $(S \mid U) = 0$ we have

$$\begin{aligned}
\mathrm{Ad}_A(U) &= (\cos^2\theta - \sin^2\theta)U + \cos\theta[S, U] \\
&= (\cos 2\theta)U + \cos\theta[S, U] \\
&= (\cos 2\theta)U + \sqrt{2}\cos\theta\sin\theta[\hat{S}, U] \\
&= \cos 2\theta\, U + \sin 2\theta \hat{S} \times U,
\end{aligned}$$

where $\hat{S} = \dfrac{1}{\sqrt{2}\sin\theta}S$ is of unit length. Noting that U and $\hat{S} \times U$ are orthogonal to S, we see that the effect of Ad_A on U is to rotate it in the plane orthogonal to S (and spanned by U and $\hat{S} \times U$) through the angle θ.

We can now see that every element $R \in \mathrm{SO}(3)$ has the form Ad_A for some $A \in \mathrm{SU}(2)$. This follows from the facts that the eigenvalues of R have modulus 1 and $\det R = 1$. Together these show that at least one of the eigenvalues of R must be 1 with corresponding eigenvector $\mathbf{v}$ say, while the other two have the form $e^{\pm\varphi i} = \cos \pm i \sin\varphi$ for some φ. Now we can take $A = \cos(\varphi/2)I + S$ where $S \in \mathfrak{su}(2)$ is chosen to correspond to a multiple of $\mathbf{v}$ and $(S \mid S) = 2\sin^2(\varphi/2)$. If we choose $-\varphi$ in place of φ we obtain $-A$ in place of A. $\square$

Let $B \in \mathfrak{su}(2)$. Then the curve

$$\beta \colon \mathbb{R} \longrightarrow \mathrm{SU}(2); \quad \beta(t) = \exp(tB),$$

gives rise to the curve

$$\overline{\beta} \colon \mathbb{R} \longrightarrow \mathrm{SO}(3); \quad \overline{\beta}(t) = \overline{\mathrm{Ad}}_{\beta(t)}.$$

We can differentiate $\overline{\beta}$ at $t = 0$ to obtain and element of $\mathfrak{so}(3)$ which R^3 identified with $\mathfrak{su}(2)$ by the formula:

$$\begin{aligned}
\overline{\beta}'(0)(X) &= \frac{\mathrm{d}}{\mathrm{d}\,t}\exp(tB)X\exp(-tB)_{|t=0} \\
&= BX - XB = [B, X].
\end{aligned}$$

For example when $B = H$,

$$[H, H] = 0, \quad [H, E] = F, \quad [H, F] = -E,$$

hence the matrix of H acting on $\mathfrak{su}(2)$ relative to the basis H, E, F is

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = R.$$

Similarly,

$$[E, H] = -F, \quad [E, E] = 0, \quad [E, F] = \hat{H},$$

giving the matrix

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} = Q,$$

and

$$[F, H] = E, \quad [F, E] = -H, \quad [F, F] = 0,$$

giving the matrix

$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = P.$$

So the corresponding derivative map is

$$\mathrm{d}\,\overline{\mathrm{Ad}} \colon \mathfrak{su}(2) \longrightarrow \mathfrak{so}(3); \quad \mathrm{d}\,\overline{\mathrm{Ad}}(xH + yE + zF) = xR + yQ + zP.$$

Apart from the change in order, this is the obvious isomorphism between these two Lie algebras.

To summarize, we have proved the following.

THEOREM 2.28. $\overline{\mathrm{Ad}} \colon SU(2) \longrightarrow \mathrm{SO}(3)$ *is a surjective Lie homomorphism with* $\ker \overline{\mathrm{Ad}} = \{\pm I\}$. *Furthermore, the derivative* $\mathrm{d}\,\overline{\mathrm{Ad}} \colon \mathfrak{su}(2) \longrightarrow \mathfrak{so}(3)$ *is an isomorphism of* $\mathbb{R}$*-Lie algebras.*

6. $\mathrm{SL}_2(\mathbb{C})$ and the Lorentz group

Let us now consider the Lie algebra $\mathrm{SL}_2(\mathbb{C})$, $\mathfrak{sl}_2(\mathbb{C})$. By Equation (2.8),

$$\mathfrak{sl}_2(\mathbb{C}) = \ker \mathrm{tr} \subseteq \mathrm{M}_2(\mathbb{C})$$

and $\dim_{\mathbb{C}} \mathfrak{sl}_2(\mathbb{C}) = 3$. The following matrices form a $\mathbb{C}$-basis for $\mathfrak{sl}_2(\mathbb{C})$:

$$H' = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad E' = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad F' = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

The elements $H', iH', E', iE', F', iF'$ form an $\mathbb{R}$-basis and $\dim \mathfrak{sl}_2(\mathbb{C}) = 6$. Notice also that $\mathfrak{su}(2) \subseteq \mathfrak{sl}_2(\mathbb{C})$ and the elements $H, E, F \in \mathfrak{su}(2)$ form a $\mathbb{C}$-basis of $\mathfrak{sl}_2(\mathbb{C})$, so H, iH, E, iE, F, iF form an $\mathbb{R}$-basis. The Lie brackets of H', E', F' are determined by

$$[H', E'] = 2E', \qquad [H', F'] = -2F', \qquad [E', F'] = H'.$$

Notice that the subspaces spanned by each of the pairs H', E' and H', F' are $\mathbb{C}$-Lie subalgebras. In fact, H', E' span the Lie algebra $\mathfrak{ut}_2(\mathbb{C})$ of the group of upper triangular complex matrices, while H', F' spans the Lie algebra of the group of *lower triangular* complex matrices.

Given the existence of the double covering homomorphism $\overline{\mathrm{Ad}} \colon \mathrm{SU}(2) \longrightarrow \mathrm{SO}(3)$ of Section 5, it seems reasonable to ask if a similar homomorphism exists for $\mathrm{SL}_2(\mathbb{C})$. It does, but we need to use the special Lorentz group Lor and then obtain a double covering homomorphism $\mathrm{SL}_2(\mathbb{C}) \longrightarrow \mathrm{Lor}$ which appears in Physics in connection with spinors and twistors.

Next we will determine the $\mathbb{R}$-Lie algebra of $\mathrm{Lor} \leqslant \mathrm{SL}_4(\mathbb{R})$, $\mathfrak{lor}$. Let $\alpha\colon (-\varepsilon,\varepsilon) \longrightarrow \mathrm{Lor}$ be a differentiable curve with $\alpha(0) = I$. By definition, for $t \in (-\varepsilon,\varepsilon)$ we have

$$\alpha(t)Q\alpha(t)^T = Q,$$

where

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Differentiating and setting $t = 0$ we obtain

$$\alpha'(0)Q + Q\alpha'(0)^T = O,$$

giving

$$\begin{bmatrix} \alpha'(0)_{11} & \alpha'(0)_{12} & \alpha'(0)_{13} & -\alpha'(0)_{14} \\ \alpha'(0)_{21} & \alpha'(0)_{22} & \alpha'(0)_{23} & -\alpha'(0)_{24} \\ \alpha'(0)_{31} & \alpha'(0)_{32} & \alpha'(0)_{33} & -\alpha'(0)_{34} \\ \alpha'(0)_{41} & \alpha'(0)_{42} & \alpha'(0)_{43} & -\alpha'(0)_{44} \end{bmatrix} + \begin{bmatrix} \alpha'(0)_{11} & \alpha'(0)_{21} & \alpha'(0)_{31} & \alpha'(0)_{41} \\ \alpha'(0)_{12} & \alpha'(0)_{22} & \alpha'(0)_{32} & \alpha'(0)_{42} \\ \alpha'(0)_{13} & \alpha'(0)_{23} & \alpha'(0)_{33} & \alpha'(0)_{43} \\ -\alpha'(0)_{14} & -\alpha'(0)_{42} & -\alpha'(0)_{34} & -\alpha'(0)_{44} \end{bmatrix} = O.$$

So we have

$$\alpha'(0) = \begin{bmatrix} 0 & \alpha'(0)_{12} & \alpha'(0)_{13} & \alpha'(0)_{14} \\ -\alpha'(0)_{12} & 0 & \alpha'(0)_{23} & \alpha'(0)_{24} \\ -\alpha'(0)_{13} & -\alpha'(0)_{23} & 0 & \alpha'(0)_{34} \\ \alpha'(0)_{14} & \alpha'(0)_{24} & \alpha'(0)_{34} & 0 \end{bmatrix}.$$

Notice that the trace of such a matrix is zero.

In fact, every matrix of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ a_{14} & a_{24} & a_{34} & 0 \end{bmatrix}$$

or equivalently satisfying $AQ + QA^T = O$ (and hence $\operatorname{tr} A = 0$), is in $\mathfrak{lor}$. This holds since there is a curve

$$\alpha\colon \mathbb{R} \longrightarrow \mathrm{GL}_4(\mathbb{R}); \quad \alpha(t) = \exp(tA),$$

with $\alpha'(0) = A$ which satisfies

$$Q\exp(tA)^T = \exp(tA)Q\exp(tA^T) = \exp(tA)\exp(-tA)Q = Q,$$

since $QA^T = -AQ$, and by Lemma 2.20,

$$\det\exp(tA) = e^{\operatorname{tr}(tA)} = 1,$$

and moreover it preserves the components of the hyperboloid $x_1^2 + x_2^2 + x_3^2 - x_4^2 = -1$; all of this shows that $\exp(tA) \in \mathrm{Lor}$. Therefore we might as well redefine

$$\alpha\colon \mathbb{R} \longrightarrow \mathrm{Lor}; \quad \alpha(t) = \exp(tA).$$

We have shown that

$$\mathfrak{lor} = \{A \in \mathrm{M}_4(\mathbb{R}) : AQ + QA^T = O\} = \left\{ A \in \mathrm{M}_4(\mathbb{R}) : A = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ a_{14} & a_{24} & a_{34} & 0 \end{bmatrix} \right\}. \tag{2.18}$$

We also have

$$\dim \mathrm{Lor} = \dim \mathfrak{lor} = 6. \tag{2.19}$$

An $\mathbb{R}$-basis for $\mathfrak{lor}$ consists of the elements

$$P_{12} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad P_{13} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad P_{14} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

$$P_{23} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad P_{24} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad P_{34} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The non-trivial brackets for these are

$$\begin{array}{lllll} [P_{12}, P_{13}] = P_{23}, & [P_{12}, P_{14}] = P_{24}, & [P_{12}, P_{23}] = -P_{13}, & [P_{12}, P_{24}] = -P_{14}, & [P_{12}, P_{34}] = 0, \\ [P_{13}, P_{14}] = P_{34}, & [P_{13}, P_{23}] = P_{12}, & [P_{13}, P_{24}] = 0, & [P_{13}, P_{34}] = 0, & \\ [P_{14}, P_{23}] = 0, & [P_{14}, P_{24}] = -P_{12}, & [P_{14}, P_{34}] = -P_{13}, & & \\ [P_{23}, P_{24}] = P_{34}, & [P_{23}, P_{34}] = -P_{24}, & [P_{24}, P_{34}] = -P_{23}. & & \end{array}$$

We will now define the homomorphism $\mathrm{SL}_2(\mathbb{C}) \longrightarrow \mathrm{Lor}$. To do this we will identify the 2×2 skew hermitian matrices $\text{Sk-Herm}_2(\mathbb{C})$ with $\mathbb{R}^4$ by

$$\begin{bmatrix} (t+x)i & y+zi \\ -y+zi & (t-x)i \end{bmatrix} \longleftrightarrow x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 + t\mathbf{e}_4.$$

Define an $\mathbb{R}$-bilinear inner product on $\text{Sk-Herm}_2(\mathbb{C})$ by the formula

$$\langle S_1 | S_2 \rangle = \frac{1}{4}(\det(S_1 + S_2) - \det(S_1 - S_2)). \tag{2.20}$$

When $S_1 = S_2 = S$ we obtain

$$\langle S | S \rangle = \frac{1}{4}(\det 2S - \det O) = \det S,$$

It is easy to check that

$$\left\langle \begin{bmatrix} (t_1+x_1)i & y_1+z_1i \\ -y_1+z_1i & (t_1-x_1)i \end{bmatrix} \middle| \begin{bmatrix} (t_2+x_2)i & y_2+z_2i \\ -y_2+z_2i & (t_2-x_2)i \end{bmatrix} \right\rangle = x_1x_2 + y_1y_2 + z_1z_2 - t_1t_2, \tag{2.21}$$

which is the Lorentzian inner product on $\mathbb{R}^4$, which is also given by

$$x_1x_2 + y_1y_2 + z_1z_2 - t_1t_2 = \begin{bmatrix} x_1 & y_1 & z_1 & t_1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \\ t_2 \end{bmatrix}$$

The *polarization identity* allows us to write

$$\langle S_1 | S_2 \rangle = \frac{1}{4}(\det(S_1 + S_2) - \det(S_1 - S_2)). \tag{2.22}$$

Now observe that for $A \in \mathrm{SL}_2(\mathbb{C})$ and $S \in \text{Sk-Herm}_2(\mathbb{C})$,

$$(ASA^*)^* = AS^*A^* = -ASA^*,$$

so $ASA^* \in \text{Sk-Herm}_2(\mathbb{C})$. By Equation (2.22), for $S_1, S_2 \in \text{Sk-Herm}_2(\mathbb{C})$, and the fact that $\det A = 1 = \det A^*$,

$$\begin{aligned}\langle AS_1A^*|AS_2A^*\rangle &= \frac{1}{4}(\det A(S_1+S_2)A^* - \det A(S_1-S_2)A^*)\\ &= \frac{1}{4}(\det A\det(S_1+S_2)\det A^* - \det A\det(S_1-S_2)\det A^*)\\ &= \frac{1}{4}(\det(S_1+S_2) - \det(S_1-S_2))\\ &= \langle S_1|S_2\rangle\,.\end{aligned}$$

Hence the function

$$\text{Sk-Herm}_2(\mathbb{C}) \longrightarrow \text{Sk-Herm}_2(\mathbb{C}); \quad S \mapsto ASA^*,$$

is an $\mathbb{R}$-linear transformation preserving the inner product $\langle\ |\ \rangle$. We can identify this with an $\mathbb{R}$-linear transformation $\widetilde{\mathrm{Ad}}_A\colon \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ which preserves the Lorentzian inner product. In fact, $\det\widetilde{\mathrm{Ad}}_A = 1$ and Ad_A preserves the components of the hyperboloid $x^2+y^2+z^2-t^2 = -1$. Let

$$\widetilde{\mathrm{Ad}}\colon \mathrm{SL}_2(\mathbb{C}) \longrightarrow \mathrm{Lor}; \quad \widetilde{\mathrm{Ad}}(A) = \widetilde{\mathrm{Ad}}_A.$$

$\widetilde{\mathrm{Ad}}$ is homomorphism since

$$\widetilde{\mathrm{Ad}}_{AB}(S) = AB(S)(AB)^* = AB(S)B^*A^* = \widetilde{\mathrm{Ad}}_A(\widetilde{\mathrm{Ad}}_B(S)) = \widetilde{\mathrm{Ad}}_A\widetilde{\mathrm{Ad}}_B(S).$$

It is also continuous. Also, $A \in \ker\widetilde{\mathrm{Ad}}$ if and only if $ASA^* = S$ for all $S \in \text{Sk-Herm}_2(\mathbb{C})$, and it is easy to see that this occurs exactly when $A = \pm I$. This shows that $\ker\widetilde{\mathrm{Ad}} = \{\pm I\}$.

THEOREM 2.29. $\widetilde{\mathrm{Ad}}\colon \mathrm{SL}_2(\mathbb{C}) \longrightarrow \mathrm{Lor}$ *is a continuous surjective Lie homomorphism with* $\ker\widetilde{\mathrm{Ad}} = \{\pm I\}$*, hence* $\mathrm{SL}_2(\mathbb{C})\{\pm I\} \cong \mathrm{Lor}$*. Furthermore the derivative* $\mathfrak{sl}_2(\mathbb{C}) \longrightarrow \mathfrak{lor}$ *is an isomorphism of* $\mathbb{R}$*-Lie algebras.*

We will not prove that $\widetilde{\mathrm{Ad}}$ is surjective but merely consider what happens at the Lie algebra level. As in the case of $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$, we can determine the derivative $\mathrm{d}\,\widetilde{\mathrm{Ad}}$ by considering for each $C \in \mathfrak{sl}_2(\mathbb{C})$, the curve

$$\gamma\colon \mathbb{R} \longrightarrow \mathrm{Lor}; \quad \gamma(t) = \exp(tC),$$

which gives rise to the curve

$$\overline{\gamma}\colon \mathbb{R} \longrightarrow \mathrm{Lor}; \quad \overline{\gamma}(t) = \widetilde{\mathrm{Ad}}_{\gamma(t)}.$$

Using as an $\mathbb{R}$-basis for $\text{Sk-Herm}_2(\mathbb{C})$ the vectors

$$V_1 = \begin{bmatrix} i & 0\\ 0 & -i\end{bmatrix}, \quad V_2 = \begin{bmatrix} 0 & 1\\ -1 & 0\end{bmatrix}, \quad V_3 = \begin{bmatrix} 0 & i\\ i & 0\end{bmatrix}, \quad V_4 = \begin{bmatrix} i & 0\\ 0 & i\end{bmatrix},$$

we can determine the action of $\widetilde{\mathrm{Ad}}_{\gamma(t)}$ on $\text{Sk-Herm}_2(\mathbb{C})$ and interpret it as an element of Lor. Differentiating we obtain the action of C as an element of $\mathfrak{lor}$ and so $\mathrm{d}\,\widetilde{\mathrm{Ad}}(C)$. For $X \in \text{Sk-Herm}_2(\mathbb{C})$ we have

$$\widetilde{\mathrm{Ad}}_{\gamma(t)}(X) = \exp(tC)X\exp(tC)^* = \exp(tC)X\exp(tC^*),$$

hence

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\widetilde{\mathrm{Ad}}_{\gamma(t)}(X)|t = 0 = CX + XC^*.$$

So for the $\mathbb{R}$-basis H, iH, E, iE, F, iF of $\mathfrak{sl}_2(\mathbb{C})$, we have

$$H(x_1V_1 + x_2V_2 + x_3V_3 + x_4V_4) + (x_1V_1 + x_2V_2 + x_3V_3 + x_4V_4)H^* = x_2V_3 - x_3V_2,$$

so

$$\mathrm{d}\,\widetilde{\mathrm{Ad}}(H) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Here is the complete list written in terms of the matrices P_{rs} which we know form an $\mathbb{R}$-basis of Lor:

$$\mathrm{d}\,\widetilde{\mathrm{Ad}}(H) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = P_{23}, \qquad \mathrm{d}\,\widetilde{\mathrm{Ad}}(iH) = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} = -P_{14},$$

$$\mathrm{d}\,\widetilde{\mathrm{Ad}}(E) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = -P_{13}, \qquad \mathrm{d}\,\widetilde{\mathrm{Ad}}(iE) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} = -P_{24},$$

$$\mathrm{d}\,\widetilde{\mathrm{Ad}}(F) = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = P_{12}, \qquad \mathrm{d}\,\widetilde{\mathrm{Ad}}(iF) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix} = -P_{34}.$$

This shows that $\mathrm{d}\,\widetilde{\mathrm{Ad}}(C)$ maps a basis for $\text{Sk-Herm}_2(\mathbb{C})$ to one for $\mathfrak{lor}$ and hence it is an isomorphism of Lie algebras.

CHAPTER 3

Quaternions, Clifford algebras and some associated groups

1. Algebras

In this section $\Bbbk$ will denote any field, although our main interest will be in the cases $\mathbb{R}, \mathbb{C}$.

Definition 3.1. A *finite dimensional* (*associative* and *unital*) *algebra* A is a finite dimensional $\Bbbk$-vector space which is an associative and unital ring such that for all $r, s \in \Bbbk$ and $a, b \in A$,

$$(ra)(sb) = (rs)(ab).$$

If A is a ring then A is a *commutative* $\Bbbk$*-algebra.*

If every non-zero element $u \in A$ is a unit, i.e., is invertible, then A is a *division algebra.*

In this last equation, ra and sb are scalar products in the vector space structure, while $(rs)(ab)$ is the scalar product of rs with the ring product ab. Furthermore, if $1 \in \Bbbk$ is the unit of A, for $t \in \Bbbk$, the element $t1 \in A$ satisfies

$$(t1)a = ta = t(a1) = a(t1).$$

If $\dim \Bbbk A > 0$, then $1 \neq 0$, and the function

$$\eta \colon \Bbbk \longrightarrow A; \quad \eta(t) = t1$$

is an injective ring homomorphism; we usually just write t for $\eta(t) = t1$.

Example 3.2. For $n \geqslant 1$, $\mathrm{M}_n(\Bbbk)$ is a $\Bbbk$-algebra. Here we have $\eta(t) = tI_n$. For $n > 1$, $\mathrm{M}_n(\Bbbk)$ is non-commutative.

Example 3.3. The ring of complex numbers $\mathbb{C}$ is an $\mathbb{R}$-algebra. Here we have $\eta(t) = t$. $\mathbb{C}$ is commutative. Notice that $\mathbb{C}$ is a commutative division algebra.

A commutative division algebra is usually called a *field* while a non-commutative division algebra is called a *skew field.* In French *corps* ($\sim$ field) is often used in sense of possibly non-commutative division algebra.

In any algebra, the set of units of A forms a group $A^\times$ under multiplication, and this contains $\Bbbk^\times$. For $A = \mathrm{M}_n(\Bbbk)$, $\mathrm{M}_n(\Bbbk)^\times = \mathrm{GL}_n(\Bbbk)$.

Definition 3.4. Let A, B be two $\Bbbk$-algebras. A $\Bbbk$-linear transformation that is also a ring homomorphism is called a $\Bbbk$*-algebra homomorphism* or *homomorphism of* $\Bbbk$*-algebras.*

A homomorphism of $\Bbbk$-algebras $\varphi \colon A \longrightarrow B$ which is also an isomorphism of rings or equivalently of $\Bbbk$-vector spaces is called *isomorphism of* $\Bbbk$*-algebras.*

Notice that the unit $\eta \colon \Bbbk \longrightarrow A$ is always a homomorphism of $\Bbbk$-algebras. There are obvious notions of kernel and image for such homomorphisms, and of subalgebra.

Definition 3.5. Given two $\Bbbk$-algebras A, B, their *direct product* has underlying set $A \times B$ with sum and product

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2), \quad (a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2).$$

The zero is $(0, 0)$ while the unit is $(1, 1)$.

It is easy to see that there is an isomorphism of $\Bbbk$-algebras $A \times B \cong B \times A$.

Given a $\Bbbk$-algebra A, it is also possible to consider the ring $\mathrm{M}_n(A)$ consisting of $m \times m$ matrices with entries in A; this is also a $\Bbbk$-algebra of dimension

$$\dim_{\Bbbk} \mathrm{M}_m(A) = m^2 \dim_{\Bbbk} A.$$

It is often the case that a $\Bbbk$-algebra A contains a subalgebra $\Bbbk_1 \subseteq A$ which is also a field. In that case A can be viewed as a over $\Bbbk_1$ in two different ways, corresponding to left and right multiplication by elements of $\Bbbk_1$. Then for $t \in \Bbbk_1$, $a \in A$,

$$\text{(Left scalar multiplication)} \qquad t \cdot a = ta;$$

$$\text{(Right scalar multiplication)} \qquad a \cdot t = at.$$

These give different $\Bbbk_1$-vector space structures unless all elements of $\Bbbk_1$ commute with all elements of A, in which case $\Bbbk_1$ is said to be a *central subfield* of A. We sometimes write ${}_{\Bbbk_1}A$ and $A_{\Bbbk_1}$ to indicate which structure is being considered. $\Bbbk_1$ is itself a finite dimensional commutative $\Bbbk$-algebra of some dimension $\dim_{\Bbbk} \Bbbk_1$.

PROPOSITION 3.6. *Each of the $\Bbbk_1$-vector spaces ${}_{\Bbbk_1}A$ and $A_{\Bbbk_1}$ is finite dimensional and in fact*

$$\dim_{\Bbbk} A = \dim_{\Bbbk_1}({}_{\Bbbk_1}A) \dim_{\Bbbk} \Bbbk_1 = \dim_{\Bbbk_1} A_{\Bbbk_1} \dim_{\Bbbk} \Bbbk_1.$$

EXAMPLE 3.7. Let $\Bbbk = \mathbb{R}$ and $A = \mathrm{M}_2(\mathbb{R})$, so $\dim_{\mathbb{R}} A = 4$. Let

$$\Bbbk_1 = \left\{ \begin{bmatrix} x & y \\ -y & x \end{bmatrix} : x, y \in \mathbb{R} \right\} \subseteq \mathrm{M}_2(\mathbb{R}).$$

Then $\Bbbk_1 \cong \mathbb{C}$ so is a subfield of $\mathrm{M}_2(\mathbb{R})$, but it is not a central subfield. Also $\dim_{\Bbbk_1} A = 2$.

EXAMPLE 3.8. Let $\Bbbk = \mathbb{R}$ and $A = \mathrm{M}_2(\mathbb{C})$, so $\dim_{\mathbb{R}} A = 8$. Let

$$\Bbbk_1 = \left\{ \begin{bmatrix} x & y \\ -y & x \end{bmatrix} : x, y \in \mathbb{R} \right\} \subseteq \mathrm{M}_2(\mathbb{C}).$$

Then $\Bbbk_1 \cong \mathbb{C}$ so is subfield of $\mathrm{M}_2(\mathbb{C})$, but it is not a central subfield. Here $\dim_{\Bbbk_1} A = 4$.

Given a $\Bbbk$-algebra A and a subfield $\Bbbk_1 \subseteq A$ containing $\Bbbk$ (possibly equal to $\Bbbk$), an element $a \in A$ acts on A by left multiplication:

$$a \cdot u = au \quad (u \in A).$$

This is always a $\Bbbk$-linear transformation of A, and if we view A as the $\Bbbk_1$-vector space $A_{\Bbbk_1}$, it is always a $\Bbbk_1$-linear transformation. Given a $\Bbbk_1$-basis $\{v_1, \ldots, v_m\}$ for $A_{\Bbbk_1}$, there is an $m \times m$ matrix $\rho(a)$ with entries in $\Bbbk_1$ defined by

$$\lambda(a)v_j = \sum_{r=1}^{m} \lambda(a)_{rj} v_r.$$

It is easy to check that

$$\lambda \colon A \longrightarrow \mathrm{M}_m(\Bbbk_1); \quad a \longmapsto \lambda(a)$$

is a homomorphism of $\Bbbk$-algebras, called the *left regular representation* of A over $\Bbbk_1$ with respect to the basis $\{v_1, \ldots, v_m\}$.

LEMMA 3.9. $\lambda \colon A \longrightarrow \mathrm{M}_m(\Bbbk_1)$ *has trivial kernel* $\ker \lambda = 0$, *hence it is an injection.*

PROOF. If $a \in \ker \lambda$ then $\lambda(a)(1) = 0$, giving $a1 = 0$, so $a = 0$. □

DEFINITION 3.10. The $\Bbbk$-algebra A is *simple* if it has only one proper two sided ideal, namely (0), hence every non-trivial $\Bbbk$-algebra homomorphism $\theta \colon A \longrightarrow B$ is an injection.

PROPOSITION 3.11. *Let* $\Bbbk$ *be a field.*

i) *For a division algebra* $\mathbb{D}$ *over* $\Bbbk$, $\mathbb{D}$ *is simple.*

ii) *For a simple* $\Bbbk$*-algebra* A, $\mathrm{M}_n(A)$ *is simple. In particular,* $\mathrm{M}_n(\Bbbk)$ *is a simple* $\Bbbk$*-algebra.*

On restricting the left regular representation to the group of units of $A^\times$, we obtain an injective group homomorphism

$$\lambda^\times \colon A^\times \longrightarrow \mathrm{GL}_m(\Bbbk_1); \quad \lambda^\times(a)(u) = au,$$

where $\Bbbk_1 \subseteq A$ is a subfield containing $\Bbbk$ and we have chosen a $\Bbbk_1$-basis of $A_{\Bbbk_1}$. Because

$$A^\times \cong \operatorname{im} \lambda^\times \leqslant \mathrm{GL}_m(\Bbbk_1),$$

$A^\times$ and its subgroups give groups of matrices.

Given a $\Bbbk$-basis of A, we obtain a group homomorphism

$$\rho^\times \colon A^\times \longrightarrow \mathrm{GL}_n(\Bbbk); \quad \rho^\times(a)(u) = ua^{-1}.$$

We can combine $\lambda^\times$ and $\rho^\times$ to obtain two further group homomorphisms

$$\lambda^\times \times \rho^\times \colon A^\times \times A^\times \longrightarrow \mathrm{GL}_n(\Bbbk); \quad \lambda^\times \times \rho^\times(a,b)(u) = aub^{-1},$$
$$\Delta \colon A^\times \longrightarrow \mathrm{GL}_n(\Bbbk); \quad \Delta(a)(u) = aua^{-1}.$$

Notice that these have non-trivial kernels,

$$\ker \lambda^\times \times \rho^\times = \{(1,1), (-1,-1)\}, \quad \ker \Delta = \{1, -1\}.$$

2. Linear algebra over a division algebra

Throughout this section, let $\mathbb{D}$ be a finite dimensional division algebra over a field $\Bbbk$.

DEFINITION 3.12. A (*right*) $\mathbb{D}$*-vector space* V is a right $\mathbb{D}$-module, i.e., an abelian group with a right scalar multiplication by elements of $\mathbb{D}$ so that for $u, v \in V$, $x, y \in \mathbb{D}$,

$$v(xy) = (vx)y,$$
$$v(x+y) = vx + vy,$$
$$(u+v)x = ux + vx,$$
$$v1 = v.$$

All the obvious notions of $\mathbb{D}$-linear transformations, subspaces, kernels and images make sense as do notions of spanning set and linear independence over $\mathbb{D}$.

THEOREM 3.13. *Let* V *be a* $\mathbb{D}$*-vector space. Then* V *has a* $\mathbb{D}$*-basis.*

If V *has a finite spanning set over* $\mathbb{D}$ *then it has a finite* $\mathbb{D}$*-basis; furthermore any two such finite bases have the same number of elements.*

DEFINITION 3.14. A $\mathbb{D}$-vector space V with a finite basis is called *finite dimensional* and the number of elements in a basis is called the *dimension of* V *over* $\mathbb{D}$, denoted $\dim_{\mathbb{D}} V$.

For $n \geqslant 1$, we can view $\mathbb{D}^n$ as the set of $n \times 1$ column vectors with entries in $\mathbb{D}$ and this becomes a $\mathbb{D}$-vector space with the obvious scalar multiplication

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} x = \begin{bmatrix} z_1 x \\ z_2 x \\ \vdots \\ z_n x \end{bmatrix}$$

PROPOSITION 3.15. *Let V, W be two finite dimensional vector spaces over $\mathbb{D}$, of dimensions $\dim_{\mathbb{D}} V = m$, $\dim_{\mathbb{D}} W = n$ and with bases $\{v_1, \ldots, v_m\}$, $\{w_1, \ldots, w_n\}$. Then a $\mathbb{D}$-linear transformation $\varphi\colon V \longrightarrow W$ is given by*

$$\varphi(v_j) = \sum_{r=1}^{n} w_r a_{rj}$$

for unique elements $a_{ij} \in \mathbb{D}$. Hence if

$$\varphi\left(\sum_{s=1}^{n} v_s x_s\right) = \sum_{s=1}^{n} w_r y_r,$$

then

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

In particular, for $V = \mathbb{D}^m$ and $W = \mathbb{D}^n$, every $\mathbb{D}$-linear transformation is obtained in this way from left multiplication by a fixed matrix.

This is of course analogous to what happens over a field except that we are careful to keep the scalar action on the right and the matrix action on the left.

We will be mainly interested in linear transformations which we will identify with the corresponding matrices. If $\theta\colon \mathbb{D}^k \longrightarrow \mathbb{D}^m$ and $\varphi\colon \mathbb{D}^m \longrightarrow \mathbb{D}^n$ are $\mathbb{D}$-linear transformations with corresponding matrices $[\theta], [\varphi]$, then

$$[\theta][\varphi] = [\theta \circ \varphi]. \tag{3.1}$$

Also, the identity and zero functions $\mathrm{Id}, 0\colon \mathbb{D}^m \longrightarrow \mathbb{D}^m$ have $[\mathrm{Id}] = I_m$ and $[0] = O_m$.

Notice that given a $\mathbb{D}$-linear transformation $\varphi\colon V \longrightarrow W$, we can 'forget' the $\mathbb{D}$-structure and just view it as a $\Bbbk$-linear transformation. Given $\mathbb{D}$-bases $\{v_1, \ldots, v_m\}$, $\{w_1, \ldots, w_n\}$ and a basis $\{b_1, \ldots, b_d\}$ say for $\mathbb{D}$, the elements

$$v_r b_t \quad (r = 1, \ldots, m,\ t = 1, \ldots, d),$$
$$w_s b_t \quad (s = 1, \ldots, n,\ t = 1, \ldots, d)$$

form $\Bbbk$-bases for V, W as $\Bbbk$-vector spaces.

We denote the set of al $m \times n$ matrices with entries in $\mathbb{D}$ by $\mathrm{M}_{m,n}(\mathbb{D})$ and $\mathrm{M}_n(\mathbb{D}) = \mathrm{M}_{n,n}(\mathbb{D})$. Then $\mathrm{M}_n(\mathbb{D})$ is a $\Bbbk$-algebra of dimension $\dim \mathrm{M}_n(\mathbb{D}) = n^2 \dim_{\Bbbk} \mathbb{D}$. The group of units of $\mathrm{M}_n(\mathbb{D})$ is denoted $\mathrm{GL}_n(\mathbb{D})$. However, for non-commutative $\mathbb{D}$ there is no determinant function so we cannot define an analogue of the special linear group. We can however use the left regular representation to overcome this problem with the aid of some algebra.

PROPOSITION 3.16. *Let A be algebra over a field $\Bbbk$ and $B \subseteq A$ a finite dimensional subalgebra. If $u \in B$ is a unit in A then $u^{-1} \in B$, hence u is a unit in B.*

PROOF. Since B is finite dimensional, the powers u^k $(k \geqslant 0)$ are linearly dependent over $\Bbbk$, so for some $t_r \in \Bbbk$ $(r = 0, \ldots, \ell)$ with $t_\ell \neq 0$ and $\ell \geqslant 1$, there is a relation

$$\sum_{r=0}^{\ell} t_r u^r = 0.$$

If we choose k suitably and multiply by a non-zero scalar, then we can assume that

$$u^k - \sum_{r=k+1}^{\ell} t_r u^r = 0.$$

If v is the inverse of u in A, then multiplication by v^{k+1} gives

$$v - \sum_{r=k+1}^{\ell} t_r u^{r-k-1} = 0,$$

from which we obtain

$$v = \sum_{r=k+1}^{\ell} t_r u^{r-k-1} \in B.$$

$\square$

For a division algebra $\mathbb{D}$, each matrix $A \in \mathrm{M}_n(\mathbb{D})$ acts by multiplication on the left of $\mathbb{D}^n$. For any subfield $\Bbbk_1 \subseteq \mathbb{D}$ containing $\Bbbk$, A induces a (right) $\Bbbk_1$-linear transformation,

$$\mathbb{D}^n \longrightarrow \mathbb{D}^n; \quad \mathbf{x} \longmapsto A\mathbf{x}.$$

If we choose a $\Bbbk_1$-basis for $\mathbb{D}$, A gives rise to a matrix $\Lambda_A \in \mathrm{M}_{nd}(\Bbbk_1)$ where $d = \dim_{\Bbbk_1} \mathbb{D}_{\Bbbk_1}$. It is easy to see that the function

$$\Lambda\colon\ \mathrm{M}_n(\mathbb{D}) \longrightarrow \mathrm{M}_{nd}(\Bbbk_1); \quad \Lambda(A) = \Lambda_A,$$

is a ring homomorphism with $\ker \Lambda = 0$. This allows us to identify $\mathrm{M}_n(\mathbb{D})$ with the subring $\operatorname{im} \Lambda \subseteq \mathrm{M}_{nd}(\Bbbk_1)$.

Applying Proposition 3.16 we see that A is invertible in $\mathrm{M}_n(\mathbb{D})$ if and only if Λ_A is invertible in $\mathrm{M}_{nd}(\Bbbk_1)$. But the latter is true if and only if $\det \Lambda_A \neq 0$.

Hence to determine invertibility of $A \in \mathrm{M}_n(\mathbb{D})$, it suffices to consider $\det \Lambda_A$ using a subfield $\Bbbk_1$. The resulting function

$$\mathrm{Rdet}_{\Bbbk_1}\colon\ \mathrm{M}_n(\mathbb{D}) \longrightarrow \Bbbk_1; \quad \mathrm{Rdet}_{\Bbbk_1}(A) = \det \Lambda_A,$$

is called the $\Bbbk_1$-*reduced determinant* of $\mathrm{M}_n(\mathbb{D})$ and is a group homomorphism. It is actually true that $\det \Lambda_A \in \Bbbk$, not just in $\Bbbk_1$, although we will not prove this here.

PROPOSITION 3.17. *$A \in \mathrm{M}_n(\mathbb{D})$ is invertible if and only if $\mathrm{Rdet}_{\Bbbk_1}(A) \neq 0$ for some subfield $\Bbbk_1 \subseteq \mathbb{D}$ containing $\Bbbk$.*

3. Quaternions

PROPOSITION 3.18. *If A is a finite dimensional commutative $\mathbb{R}$-division algebra then either $A = \mathbb{R}$ or there is an isomorphism of $\mathbb{R}$-algebras $A \cong \mathbb{C}$.*

PROOF. Let α. Since A is a finite dimensional $\mathbb{R}$-vector space, the powers $1, \alpha, \alpha^2, \ldots, \alpha^k, \ldots$ must be linearly dependent, say

$$t_0 + t_1\alpha + \cdots + t_m\alpha^m = 0 \tag{3.2}$$

for some $t_j \in \mathbb{R}$ with $m \geqslant 1$ and $t_m \neq 0$. We can choose m to be minimal with these properties. If $t_0 = 0$, then

$$t_1 + t_2\alpha + t_3\alpha^2 + \cdots + t_m\alpha^{m-1} = 0,$$

contradicting minimality; so $t_0 \neq 0$. In fact, the polynomial $p(X) = t_0 + t_1X + \cdots + t_mX^m \in \mathbb{R}[X]$ is irreducible since if $p(X) = p_1(X)p_2(X)$ then since A is a division algebra, either $p_1(\alpha) = 0$ or $p_2(\alpha) = 0$, which would contradict minimality if both $\deg p_1(X) > 0$ and $\deg p_2(X) > 0$.

Consider the $\mathbb{R}$-subspace

$$\mathbb{R}(\alpha) = \{\sum_{j=0}^{k} s_j\alpha^j : s_j \in \mathbb{R}\} \subseteq A.$$

Then $\mathbb{R}(\alpha)$ is easily seen to be a $\mathbb{R}$-subalgebra of A. The elements $1, \alpha, \alpha^2, \ldots, \alpha^{m-1}$ form a basis by Equation (3.2), hence $\dim_{\mathbb{R}} \mathbb{R}(\alpha) = m$.

Let $\gamma \in \mathbb{C}$ be any complex root of the irreducible polynomial $t_0 + t_1 X + \cdots + t_m X^m \in \mathbb{R}[X]$ which certainly exists by the Fundamental Theorem of Algebra). There is an $\mathbb{R}$-linear transformation which is actually an injection,

$$\varphi\colon \mathbb{R}(\alpha) \longrightarrow \mathbb{C}; \quad \varphi(\sum_{j=0}^{m-1} s_j \alpha^j) = \sum_{j=0}^{m-1} s_j \gamma^j.$$

It is easy to see that this is actually an $\mathbb{R}$-algebra homomorphism. Hence $\varphi\mathbb{R}(\alpha) \subseteq \mathbb{C}$ is a subalgebra. But as $\dim_{\mathbb{R}} \mathbb{C} = 2$, this implies that $m = \dim_{\mathbb{R}} \mathbb{R}(\alpha) \leqslant 2$. If $m = 1$, then by Equation (3.2), $\alpha \in \mathbb{R}$. If $m = 2$, then $\varphi\mathbb{R}(\alpha) = \mathbb{C}$.

So either $\dim_{\mathbb{R}} A = 1$ and $A = \mathbb{R}$, or $\dim_{\mathbb{R}} A > 1$ and we can choose an $\alpha \in A$ with $\mathbb{C} \cong \mathbb{R}(\alpha)$. This means that we can view A as a finite dimensional $\mathbb{C}$-algebra. Now for any $\beta \in A$ there is polynomial

$$q(X) = u_0 + u_1 X + \cdots + u_\ell X^\ell \in \mathbb{C}[X]$$

with $\ell \geqslant 1$ and $u_\ell \neq 0$. Again choosing ℓ to be minimal with this property, $q(X)$ is irreducible. But then since $q(X)$ has a root in $\mathbb{C}$, $\ell = 1$ and $\beta \in \mathbb{C}$. This shows that $A = \mathbb{C}$ whenever $\dim_{\mathbb{R}} A > 1$. □

The above proof actually shows that if A is a finite dimensional $\mathbb{R}$-division algebra, then either $A = \mathbb{R}$ or there is a subalgebra isomorphic to $\mathbb{C}$. However, the question of what finite dimensional $\mathbb{R}$-division algebras exist is less easy to decide. In fact there is only one other up to isomorphism, the skew field of *quaternions* $\mathbb{H}$. We will now show how to construct this skew field.

Let

$$\mathbb{H} = \left\{ \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} : z, w \in \mathbb{C} \right\} \subseteq \mathrm{M}_2(\mathbb{C}).$$

It is easy to see that $\mathbb{H}$ is a subring of $\mathrm{M}_2(\mathbb{C})$ and is in fact an $\mathbb{R}$-subalgebra where we view $\mathrm{M}_2(\mathbb{C})$ as an $\mathbb{R}$-algebra of dimension 8. It also contains a copy of $\mathbb{C}$, namely the $\mathbb{R}$-subalgebra

$$\left\{ \begin{bmatrix} z & 0 \\ 0 & \overline{z} \end{bmatrix} : z \in \mathbb{C} \right\} \subseteq \mathbb{H}.$$

However, $\mathbb{H}$ is not a $\mathbb{C}$-algebra since for example

$$\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = - \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \neq \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Notice that if $z, w \in \mathbb{C}$, then $z = 0 = w$ if and only if $|z|^2 + |w|^2 = 0$. We have

$$\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix} = \begin{bmatrix} |z|^2 + |w|^2 & 0 \\ 0 & |z|^2 + |w|^2 \end{bmatrix},$$

hence $\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}$ is invertible if and only if $\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \neq O$; furthermore in that case,

$$\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}^{-1} = \begin{bmatrix} \dfrac{\overline{z}}{|z|^2 + |w|^2} & \dfrac{-w}{|z|^2 + |w|^2} \\ \dfrac{\overline{w}}{|z|^2 + |w|^2} & \dfrac{z}{|z|^2 + |w|^2} \end{bmatrix}$$

which is in $\mathbb{H}$. So an element of $\mathbb{H}$ is invertible in $\mathbb{H}$ if and only if it is invertible as a matrix. Notice that

$$\mathrm{SU}(2) = \{A \in \mathbb{H} : \det A = 1\} \leqslant \mathbb{H}^\times.$$

It is useful to define on $\mathbb{H}$ a norm in the sense of Proposition 1.4:

$$\left| \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \right| = \det \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} = |z|^2 + |w|^2.$$

Then

$$\mathrm{SU}(2) = \{A \in \mathbb{H} : |A| = 1\} \leqslant \mathbb{H}^{\times}.$$

As an $\mathbb{R}$-basis of $\mathbb{H}$ we have the matrices

$$\mathbf{1} = I, \quad \mathbf{i} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

These satisfy the equations

$$\mathbf{i}^2 = \mathbf{j}^2 = k^2 = -\mathbf{1}, \quad \mathbf{ij} = \mathbf{k} = -\mathbf{ji}, \quad \mathbf{k} = \mathbf{i} = -\mathbf{kj}, \quad \mathbf{ki} = \mathbf{j} = -\mathbf{ik}.$$

This should be compared with the vector product on $\mathbb{R}^3$ as discussed in Example 2.12. From now on we will write quaternions in the form

$$q = x\mathbf{i} + \mathbf{j} + z\mathbf{k} + t\mathbf{1} \quad (x, y, z, t \in \mathbb{R}).$$

q is a *pure quaternion* if and only if $t = 0$; q is a *real quaternion* if and only if $x = y = z = 0$. We can identify the pure quaternion $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ with the element $x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 \in \mathbb{R}^3$. Using this identification we see that the scalar and vector products on $\mathbb{R}^3$ are related to quaternion multiplication by the following.

PROPOSITION 3.19. *For two pure quaternions $q_1 = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}$, $q_2 = x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}$,*

$$q_1 q_2 = -(x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}) \cdot (x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}) + (x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}) \times (x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}).$$

In particular, $q_1 q_2$ is a pure quaternion if and only if q_1 and q_2 are orthogonal, in which case $q_1 q_2$ is orthogonal to each of them.

The following result summarises the general situation about solutions of $X^2 + 1 = 0$.

PROPOSITION 3.20. *The quaternion $q = x\mathbf{i} + \mathbf{j} + z\mathbf{k} + t\mathbf{1}$ satisfies $q^2 + \mathbf{1} = \mathbf{0}$ if and only if $t = 0$ and $x^2 + y^2 + z^2 = 1$.*

PROOF. This easily follows from Proposition 3.19. □

There is a quaternionic analogue of complex conjugation, namely

$$q = x\mathbf{i} + \mathbf{j} + z\mathbf{k} + t\mathbf{1} \longmapsto \overline{q} = q^* = -x\mathbf{i} - \mathbf{j} - z\mathbf{k} + t\mathbf{1}.$$

This is 'almost' a ring homomorphism $\mathbb{H} \longrightarrow \mathbb{H}$, in fact it satisfies

(3.3a) $$\overline{(q_1 + q_2)} = \overline{q}_1 + \overline{q}_2;$$

(3.3b) $$\overline{(q_1 q_2)} = \overline{q}_2 \overline{q}_1;$$

(3.3c) $$\overline{q} = q \iff q \text{ is real quaternion};$$

(3.3d) $$\overline{q} = -q \iff q \text{ is a pure quaternion}.$$

Because of Equation (3.3b) this is called a *homomorphism of skew rings* or *anti-homomorphism of rings.* The inverse of a non-zero quaternion q can be written as

(3.4) $$q^{-1} = \frac{1}{(q\overline{q})}\overline{q} = \frac{\overline{q}}{(q\overline{q})}.$$

The real quantity $q\overline{q}$ is the square of the length of the corresponding vector,

$$|q| = \sqrt{q\overline{q}} = \sqrt{x^2 + y^2 + z^2 + t^2}.$$

For $z =$ with $u, v \in \mathbb{R}$, $\overline{z} = u\mathbf{1} - v\mathbf{i}$ is the usual complex conjugation.

In terms of the matrix description of $\mathbb{H}$, quaternionic conjugation is given by hermitian conjugation,

$$\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \longmapsto \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}^* = \begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix}$$

From now on we will write

$$1 = \mathbf{1}, \quad i = \mathbf{i}, \quad j = \mathbf{j}, \quad k = \mathbf{k}.$$

4. Quaternionic matrix groups

The above norm $|\ |$ on $\mathbb{H}$ extends to a norm on $\mathbb{H}^n$, viewed as a right $\mathbb{H}$-vector space. We can define an quaternionic inner product on $\mathbb{H}$ by

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^* \mathbf{y} = \sum_{r=1}^{n} \overline{x}_r y_r,$$

where we define the *quaternionic conjugate* of a vector by

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^* = \begin{bmatrix} \overline{x}_1 & \overline{x}_2 & \cdots & \overline{x}_n \end{bmatrix}.$$

Similarly, for any matrix $[a_{ij}]$ over $\mathbb{H}$ we can define $[a_{ij}]^* = [\overline{a}_{ji}]$.

The length of $\mathbf{x} \in \mathbb{H}^n$ is defined to be

$$|\mathbf{x}| = \sqrt{\mathbf{x}^* \mathbf{x}} = \sqrt{\sum_{r=1}^{n} |x_r|^2}.$$

We can also define a norm on $\mathrm{M}_n(\mathbb{H})$ by the method used in Section 2 of Chapter 1, i.e., for $A \in \mathrm{M}_n(\mathbb{H})$,

$$\|A\| = \sup \left\{ \frac{|A\mathbf{x}|}{|\mathbf{x}|} : \mathbf{0} \neq \mathbf{x} \in \mathbb{H}^n \right\}.$$

Then the analogue of Proposition 1.4 is true for $\|\ \|$ and the norm $|\ |$ on $\mathbb{H}$, although statements involving scalar multiplication need to be formulated with scalars on the right. There is also a resulting metric on $\mathrm{M}_n(\mathbb{H})$,

$$(A, B) \longmapsto \|A - B\|,$$

and we can use this to do analysis on $\mathrm{M}_n(\mathbb{H})$. The multiplication map $\mathrm{M}_n(\mathbb{H}) \times \mathrm{M}_n(\mathbb{H}) \longrightarrow \mathrm{M}_n(\mathbb{H})$ is again continuous, and the group of invertible elements $\mathrm{GL}_n(\mathbb{H}) \subseteq \mathrm{M}_n(\mathbb{H})$ is actually an open subset. This can be proved using either of the reduced determinants

$$\mathrm{Rdet}_{\mathbb{R}} \colon \mathrm{M}_n(\mathbb{H}) \longrightarrow \mathbb{R}, \quad \mathrm{Rdet}_{\mathbb{C}} \colon \mathrm{M}_n(\mathbb{H}) \longrightarrow \mathbb{C},$$

each of which is continuous. By Proposition 3.17,

$$\mathrm{GL}_n(\mathbb{H}) = \mathrm{M}_n(\mathbb{H}) - \mathrm{Rdet}_{\mathbb{R}}^{-1} 0, \tag{3.5a}$$

$$\mathrm{GL}_n(\mathbb{H}) = \mathrm{M}_n(\mathbb{H}) - \mathrm{Rdet}_{\mathbb{C}}^{-1} 0. \tag{3.5b}$$

In either case we see that $\mathrm{GL}_n(\mathbb{H})$ is an open subset of $\mathrm{M}_n(\mathbb{H})$. It is also possible to show that the images of embeddings $\mathrm{GL}_n(\mathbb{H}) \longrightarrow \mathrm{GL}_{4n}(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{H}) \longrightarrow \mathrm{GL}_{2n}(\mathbb{C})$ are closed. So $\mathrm{GL}_n(\mathbb{H})$ and its closed subgroups are real and complex matrix groups.

The $n \times n$ *quaternionic symplectic group* is

$$\mathrm{Sp}(n) = \{A \in \mathrm{GL}_n(\mathbb{H}) : A^* A = I\} \leqslant \mathrm{GL}_n(\mathbb{H}).$$

These are easily seen to satisfy

$$\mathrm{Sp}(n) = \{A \in \mathrm{GL}_n(\mathbb{H}) : \forall \mathbf{x}, \mathbf{y} \in \mathbb{H}^n,\ A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}\}.$$

These groups $\mathrm{Sp}(n)$ form another infinite family of compact connected matrix groups along with familiar examples such as $\mathrm{SO}(n), \mathrm{U}(n), \mathrm{SU}(n)$. There are further examples, the *spinor groups* $\mathrm{Spin}(n)$ whose description involves the *real Clifford algebras* Cl_n.

5. The real Clifford algebras

The sequence of real division algebras $\mathbb{R}, \mathbb{C}, \mathbb{H}$ can be extended by introducing the *real Clifford algebras* Cl_n, where

$$\mathrm{Cl}_0 = \mathbb{R}, \quad \mathrm{Cl}_1 = \mathbb{C}, \quad \mathrm{Cl}_2 = \mathbb{H}, \quad \dim_{\mathbb{R}} = 2^n.$$

There are also complex Clifford algebras, but we will not discuss these. The theory of Clifford algebras and spinor groups is central in modern differential geometry and topology, particularly *Index Theory.* It also appears in Quantum Theory in connection with the Dirac operator. There is also a theory of *Clifford Analysis* in which the complex numbers are replaced by a Clifford algebra and a suitable class of analytic functions are studied; a motivation for this lies in the above applications.

We begin by describing Cl_n as an $\mathbb{R}$-vector space and then explain what the product looks like in terms of a particular basis. There are elements $e_1, e_2, \ldots, e_n \in \mathrm{Cl}_n$ for which

(3.6a)
$$\begin{cases} e_s e_r = -e_s e_r, & \text{if } s \neq r. \\ e_r^2 = -1. \end{cases}$$

Moreover, the elements $e_{i_1} e_{i_2} \cdots e_{i_r}$ for increasing sequences $1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n$ with $0 \leqslant r \leqslant n$, form an $\mathbb{R}$-basis for Cl_n. Thus

(3.6b)
$$\dim_{\mathbb{R}} \mathrm{Cl}_n = 2^n.$$

When $r = 0$, the element $e_{i_1} e_{i_2} \cdots e_{i_r}$ is taken to be 1.

PROPOSITION 3.21. *There are isomorphisms of* $\mathbb{R}$*-algebras*

$$\mathrm{Cl}_1 \cong \mathbb{C}, \quad \mathrm{Cl}_2 \cong \mathbb{H}.$$

PROOF. For Cl_1, the function

$$\mathrm{Cl}_1 \longrightarrow \mathbb{C}; \quad x + y e_1 \longmapsto x + y i \quad (x, y \in \mathbb{R}),$$

is an $\mathbb{R}$-linear ring isomorphism.

Similarly, for Cl_2, the function

$$\mathrm{Cl}_2 \longrightarrow \mathbb{H}; \quad t1 + x e_1 + y e_2 + z e_1 e_2 \longmapsto t1 + x i + y j + z k \quad (t, x, y, z \in \mathbb{R}),$$

is an $\mathbb{R}$-linear ring isomorphism. □

We can order the basis monomials in the e_r by declaring $e_{i_1} e_{i_2} \cdots e_{i_r}$ to be number

$$1 + 2^{i_1 - 1} + 2^{i_2 - 1} + \cdots + 2^{i_r - 1},$$

which should be interpreted as 1 when $r = 0$. Every integer k in the range $1 \leqslant k \leqslant 2^n$ has a unique *binary expansion*

$$k = k_0 + 2k_1 + \cdots + 2^j k_j + \cdots + 2^n k_n,$$

where each $k_j = 0, 1$. This provides a one-one correspondence between such numbers k and the basis monomials of Cl_n. Here are the basis orderings for the first few Clifford algebras.

$$\mathrm{Cl}_1 : \ 1, e_1; \quad \mathrm{Cl}_2 : \ 1, e_1, e_2, e_1 e_2; \quad \mathrm{Cl}_3 : \ 1, e_1, e_2, e_1 e_2, e_3, e_1 e_3, e_2 e_3, e_1 e_2 e_3.$$

Using the left regular representation over $\mathbb{R}$ associated with this basis of Cl_n, we can realise Cl_n as a subalgebra of $\mathrm{M}_{2^n}(\mathbb{R})$.

EXAMPLE 3.22. For Cl_1 we have the basis $\{1, e_1\}$ and we find that

$$\rho(1) = I_2, \quad \rho(e_1) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

So the general formula is

$$\rho(x + ye_1) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \quad (x, y \in \mathbb{R}).$$

For Cl_2 the basis $\{1, e_1, e_2, e_1e_2\}$ leads to a realization in $\mathrm{M}_4(\mathbb{R})$ for which $\rho(1) = I_4$ and

$$\rho(e_1) = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \ \rho(e_2) = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \ \rho(e_1e_2) = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

In all cases the matrices $\rho(e_{i_1}e_{i_2}\cdots e_{i_r})$ are *generalized permutation matrices* all of whose entries are entries $0, \pm 1$ and exactly on non-zero entry in each row and column. These are always orthogonal matrices of determinant 1.

These Clifford algebras have an important universal property which actually characterises them. First notice that there is an $\mathbb{R}$-linear transformation

$$j_n : \mathbb{R}^n \longrightarrow \mathrm{Cl}_n; \quad j_n\left(\sum_{r=1}^{n} x_r \mathbf{e}_r\right) = \sum_{r=1}^{n} x_r e_r.$$

By an easy calculation,

$$j_n\left(\sum_{r=1}^{n} x_r \mathbf{e}_r\right)^2 = -\sum_{r=1}^{n} x_r^2 = -\left|\sum_{r=1}^{n} x_r \mathbf{e}_r\right|^2. \tag{3.7}$$

THEOREM 3.23 (The Universal Property of Clifford algebras). *Let A be a $\mathbb{R}$-algebra and $f : \mathbb{R}^n \longrightarrow A$ an $\mathbb{R}$-linear transformation for which*

$$f(\mathbf{x})^2 = -|\mathbf{x}|^2 1.$$

Then there is a unique homomorphism of $\mathbb{R}$-algebras $F : \mathrm{Cl}_n \longrightarrow A$ for which $F \circ j_n = f$, i.e., for all $\mathbf{x} \in \mathbb{R}^n$,

$$F(j_n(\mathbf{x})) = f(\mathbf{x}).$$

PROOF. The homomorphism F is defined by setting $F(e_r) = f(\mathbf{e}_r)$ and showing that it extends to a ring homomorphism on Cl_n. □

EXAMPLE 3.24. There is an $\mathbb{R}$-linear transformation

$$\alpha_0 : \mathbb{R}^n \longrightarrow \mathrm{Cl}_n; \quad \alpha_0(\mathbf{x}) = -j_n(\mathbf{x}) = j_n(-\mathbf{x}).$$

Then

$$\alpha_0(\mathbf{x})^2 = j_n(-\mathbf{x})^2 = -|\mathbf{x}|^2,$$

so by the Theorem there is a unique homomorphism of $\mathbb{R}$-algebras $\alpha : \mathrm{Cl}_n \longrightarrow \mathrm{Cl}_n$ for which

$$\alpha(j_n(\mathbf{x})) = \alpha_0(\mathbf{x}).$$

Since $j_n(\mathbf{e}_r) = e_r$, this implies

$$\alpha(e_r) = -e_r.$$

Notice that for $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$,

$$\alpha(e_{i_1}e_{i_2}\cdots e_{i_k}) = (-1)^k e_{i_1}e_{i_2}\cdots e_{i_k} \begin{cases} e_{i_1}e_{i_2}\cdots e_{i_k} & \text{if } k \text{ is even,} \\ -e_{i_1}e_{i_2}\cdots e_{i_k} & \text{if } k \text{ is odd.} \end{cases}$$

It is easy to see that α is an isomorphism and hence an automorphism.

This automorphism $\alpha\colon \mathrm{Cl}_n \longrightarrow \mathrm{Cl}_n$ is often called the *canonical automorphism* of Cl_n.

We record explicit form of the next few Clifford algebras. Consider the $\mathbb{R}$-algebra $\mathrm{M}_2(\mathbb{H})$ of dimension 16. Then we can define an $\mathbb{R}$-linear transformation

$$\theta_4\colon \mathbb{R}^4 \longrightarrow \mathrm{M}_2(\mathbb{H}); \quad \theta_4(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + x_4\mathbf{e}_4) = \begin{bmatrix} x_1 i + x_2 j + x_3 k & x_4 k \\ x_4 k & x_1 i + x_2 j - x_3 k \end{bmatrix}.$$

Direct calculation shows that θ_4 satisfies the condition of Theorem 3.23, hence there is a unique $\mathbb{R}$-algebra homomorphism $\Theta_4\colon \mathrm{Cl}_4 \longrightarrow \mathrm{M}_2(\mathbb{H})$ with $\Theta_4 \circ j_4 = \theta_4$. This is in fact an isomorphism of $\mathbb{R}$-algebras, so

$$\mathrm{Cl}_4 \cong \mathrm{M}_2(\mathbb{H}).$$

Since $\mathbb{R} \subseteq \mathbb{R}^2 \subseteq \mathbb{R}^3 \subseteq \mathbb{R}^4$ we obtain compatible homomorphisms

$$\Theta_1\colon \mathrm{Cl}_1 \longrightarrow \mathrm{M}_2(\mathbb{H}), \quad \Theta_2\colon \mathrm{Cl}_2 \longrightarrow \mathrm{M}_2(\mathbb{H}), \quad \Theta_3\colon \mathrm{Cl}_3 \longrightarrow \mathrm{M}_2(\mathbb{H}),$$

which have images

$$\begin{aligned} \operatorname{im}\Theta_1 &= \{zI_2 : z \in \mathbb{C}\}, \\ \operatorname{im}\Theta_2 &= \{qI_2 : q \in \mathbb{H}\}, \\ \operatorname{im}\Theta_3 &= \left\{ \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} : q_1, q_2 \in \mathbb{H} \right\}. \end{aligned}$$

This shows that there is an isomorphism of $\mathbb{R}$-algebras

$$\mathrm{Cl}_3 \cong \mathbb{H} \times \mathbb{H},$$

where the latter is the direct product of Definition 3.5. We also have

$$\mathrm{Cl}_5 \cong \mathrm{M}_4(\mathbb{C}), \quad \mathrm{Cl}_6 \cong \mathrm{M}_8(\mathbb{R}), \quad \mathrm{Cl}_7 \cong \mathrm{M}_8(\mathbb{R}) \times \mathrm{M}_8(\mathbb{R}).$$

After this we have the following periodicity result, where $\mathrm{M}_m(\mathrm{Cl}_n)$ denotes the ring of $m \times m$ matrices with entries in Cl_n.

THEOREM 3.25. *For $n \geqslant 0$,*

$$\mathrm{Cl}_{n+8} \cong \mathrm{M}_{16}(\mathrm{Cl}_n).$$

In the next section we will make use of some more structure in Cl_n. First there is a *conjugation* $\overline{(\ )}\colon \mathrm{Cl}_n \longrightarrow \mathrm{Cl}_n$ defined by

$$\overline{e_{i_1} e_{i_2} \cdots e_{i_k}} = (-1)^k e_{i_k} e_{i_{k-1}} \cdots e_{i_1}$$

whenever $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$, and satisfying

$$\begin{aligned} \overline{x+y} &= \overline{x} + \overline{y}, \\ \overline{tx}&t\overline{x}, \end{aligned}$$

for $x, y \in \mathrm{Cl}_n$ and $t \in \mathbb{R}$. Notice that this is not a ring homomorphism $\mathrm{Cl}_n \longrightarrow \mathrm{Cl}_n$ since for example whenever $r < s$,

$$\overline{e_r e_s} = e_s e_r = -e_r e_s = -\overline{e_r e_s} \neq \overline{e_r e_s}.$$

However, it is a ring *anti-homomorphism* in the sense that for all $x, y \in \mathrm{Cl}_n$,

$$\overline{xy} = \overline{y}\,\overline{x} \quad (x, y \in \mathrm{Cl}_n). \tag{3.8}$$

When $n = 1, 2$ this agrees with the conjugations already defined in $\mathbb{C}$ and $\mathbb{H}$.

Second there is the canonical automorphism $\alpha\colon \mathrm{Cl}_n \longrightarrow \mathrm{Cl}_n$ defined in Example 3.24. We can use α to define a $\pm$-*grading* on Cl_n:

$$\mathrm{Cl}_n^+ = \{u \in \mathrm{Cl}_n : \alpha(u) = u\}, \quad \mathrm{Cl}_n^- = \{u \in \mathrm{Cl}_n : \alpha(u) = -u\}.$$

PROPOSITION 3.26. i) *Every element $v \in \mathrm{Cl}_n$ can be unique expressed in the form $v = v^+ + v^-$ where $v^+ \in \mathrm{Cl}_n^+$ and $v^- \in \mathrm{Cl}_n^-$. Hence as an $\mathbb{R}$-vector space, $\mathrm{Cl}_n = \mathrm{Cl}_n^+ \oplus \mathrm{Cl}_n^-$.*
ii) *This decomposition is multiplicative in the sense that*

$$\begin{aligned} uv \in \mathrm{Cl}_n^+ &\quad \text{if } u, v \in \mathrm{Cl}_n^+ \text{ or } u, v \in \mathrm{Cl}_n^-, \\ uv, vu \in \mathrm{Cl}_n^+ &\quad \text{if } u \in \mathrm{Cl}_n^+ \text{ and } v \in \mathrm{Cl}_n^-. \end{aligned}$$

PROOF. i) The elements

$$v^+ = \frac{1}{2}(v + \alpha(v)), \quad v^- = \frac{1}{2}(v - \alpha(v)),$$

satisfy $\alpha(v^+) = v^+$, $\alpha(v^-) = -v^-$ and $v = v^+ + v^-$. This expression is easily found to be the unique one with these properties and defines the stated vector space direct sum decomposition.
ii) This is easily checked using the fact that α is a ring homomorphism. ☐

Notice that for bases of $\mathrm{Cl}_n^\pm$ we have the monomials

$$\begin{cases} e_{j_1} \cdots e_{j_{2m}} \in \mathrm{Cl}_n^+ & (1 \leqslant j_1 < \cdots < j_{2m} \leqslant n), \\ e_{j_1} \cdots e_{j_{2m+1}} \in \mathrm{Cl}_n^- & (1 \leqslant j_1 < \cdots < j_{2m+1} \leqslant n). \end{cases} \tag{3.9}$$

Finally, we introduce an inner product $\cdot$ and a norm $|\ |$ on Cl_n by defining the distinct monomials $e_{i_1} e_{i_2} \cdots e_{i_k}$ with $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$ to be an orthonormal basis, i.e.,

$$(e_{i_1} e_{i_2} \cdots e_{i_k}) \cdot (e_{j_1} e_{j_2} \cdots e_{j_\ell}) = \begin{cases} 1 & \text{if } \ell = k \text{ and } i_r = j_r \text{ for all } r, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$|x| = \sqrt{x \cdot x}.$$

A more illuminating way to define $\cdot$ is by the formula

$$u \cdot v = \frac{1}{2} \operatorname{Re}(\overline{u}\, v + \overline{v}\, u), \tag{3.10}$$

where for $w \in \mathrm{Cl}_n$ we define its *real part* $\operatorname{Re} w$ to be the coefficient of 1 when w is expanded as an $\mathbb{R}$-linear combination of the basis monomials $e_{i_1} \cdots e_{i_r}$ with $1 \leqslant i_1 < \cdots < i_r \leqslant n$ and $0 \leqslant r$. It can be shown that for any $u, v \in \mathrm{Cl}_n$ and $w \in j_n\mathbb{R}^n$,

$$(wu) \cdot (wv) = |w|^2 (u \cdot v). \tag{3.11}$$

In particular, when $|w| = 1$ left multiplication by w defines an $\mathbb{R}$-linear transformation on Cl_n which is an isometry. The norm $|\ |$ gives rise to a metric on Cl_n. This makes the group of units $\mathrm{Cl}_n^\times$ into a topological group while the above embeddings of Cl_n into matrix rings are all continuous. This implies that $\mathrm{Cl}_n^\times$ is a matrix group. Unfortunately, they are not norm preserving. For example, $2 + e_1 e_2 e_3 \in \mathrm{Cl}_3$ has $|2 + e_1 e_2 e_3| = \sqrt{5}$, but the corresponding matrix in $\mathrm{M}_8(\mathbb{R})$ has norm $\sqrt{3}$. However, by defining for each $w \in \mathrm{Cl}_n$

$$\|w\| = \{|wx| : x \in \mathrm{Cl}_n,\ |x| = 1\},$$

we obtain another equivalent norm on Cl_n for which the above embedding $\mathrm{Cl}_n \longrightarrow \mathrm{M}_{2^n}(\mathbb{R})$ does preserve norms. For $w \in j_n\mathbb{R}^n$ we do have $\|w\| = |w|$ and more generally, for $w_1, \ldots, w_k \in j_n\mathbb{R}^n$,

$$\|w_1 \cdots w_k\| = |w_1 \cdots w_k| = |w_1| \cdots |w_k|.$$

For $x, y \in \mathrm{Cl}_n$,

$$\|xy\| \leqslant \|x\| \, \|y\|,$$

without equality in general.

6. The spinor groups

In this section we will describe the compact connected *spinor groups* $\mathrm{Spin}(n)$ which are groups of units in the Clifford algebras Cl_n. Moreover, there are surjective Lie homomorphisms $\mathrm{Spin}(n) \longrightarrow \mathrm{SO}(n)$ each of whose kernels have two elements.

We begin by using the injective linear transformation $j_n \colon \mathbb{R}^n \longrightarrow \mathrm{Cl}_n$ to identify $\mathbb{R}^n$ with a subspace of Cl_n, i.e.,

$$\sum_{r=1}^{n} x_r \mathbf{e}_r \longleftrightarrow j_n(\sum_{r=1}^{n} x_r \mathbf{e}_r) = \sum_{r=1}^{n} x_r e_r.$$

Notice that $\mathbb{R}^n \subseteq \mathrm{Cl}_n^-$, so for $x \in \mathbb{R}^n$, $u \in \mathrm{Cl}_n^+$ and $v \in \mathrm{Cl}_n^-$,

$$xu, ux \in \mathrm{Cl}_n^-, \quad xv, vx \in \mathrm{Cl}_n^+. \tag{3.12}$$

Inside of $\mathbb{R}^n \subseteq \mathrm{Cl}_n$ is the unit sphere

$$\mathbb{S}^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : |x| = 1\} = \{\sum_{r=1}^{n} x_r e_r : \sum_{r=1}^{n} x_r^2 = 1\}.$$

Lemma 3.27. *Let $u \in \mathbb{S}^{n-1} \subseteq \mathrm{Cl}_n$. Then u is a unit in Cl_n, $u \in \mathrm{Cl}_n^\times$.*

Proof. Since $u \in \mathbb{R}^n$,

$$(-u)u = u(-u) = -u^2 = -(-|u|^2) = 1,$$

so $(-u)$ is the inverse of u. Notice that $-u \in \mathbb{S}^{n-1}$. □

More generally, for $u_1, \ldots, u_k \in \mathbb{S}^{n-1}$ we have

$$(u_1 \cdots u_k)^{-1} = (-1)^k u_k \cdots u_1 = \overline{u_1 \cdots u_k}. \tag{3.13}$$

Definition 3.28. The *pinor group* $\mathrm{Pin}(n)$ is the subgroup of $\mathrm{Cl}_n^\times$ generated by the elements of $\mathbb{S}^{n-1}$,

$$\mathrm{Pin}(n) = \{u_1 \cdots u_k : k \geqslant 0,\ u_r \in \mathbb{S}^{n-1}\} \leqslant \mathrm{Cl}_n^\times.$$

Notice that $\mathrm{Pin}(n)$ is a topological group and is bounded as a subset of Cl_n with respect to the metric introduced in the last section. It is in fact a closed subgroup of $\mathrm{Cl}_n^\times$ and so is a matrix group; in fact it is even compact. We will show that $\mathrm{Pin}(n)$ acts on $\mathbb{R}^n$ in an interesting fashion.

We will require the following useful result.

Lemma 3.29. *Let $u, v \in \mathbb{R}^n \subseteq \mathrm{Cl}_n$. If $u \cdot v = 0$, then*

$$vu = -uv.$$

PROOF. Writing $u = \sum_{r=1}^{n} x_r e_r$ and $v = \sum_{s=1}^{n} y_s e_s$ with $x_r, y_s \in \mathbb{R}$, we obtain

$$\begin{aligned} vu &= \sum_{s=1}^{n} \sum_{r=1}^{n} y_s x_r e_s e_r \\ &= \sum_{r=1}^{n} y_r x_r e_r^2 + \sum_{r<s} (x_s y_r - x_r y_s) e_r e_s \\ &= -\sum_{r=1}^{n} y_r x_r - \sum_{r<s} (x_r y_s - x_s y_r) e_r e_s \\ &= -u.v - \sum_{r<s} (x_r y_s - x_s y_r) e_r e_s \\ &= -\sum_{r<s} (x_r y_s - x_s y_r) e_r e_s \\ &= v.u - \sum_{r<s} (x_r y_s - x_s y_r) e_r e_s \\ &= -\sum_{r=1}^{n} \sum_{s=1}^{n} x_r y_s e_r e_s \\ &= -uv. \end{aligned}$$

□

For $u \in \mathbb{S}^{n-1}$ and $x \in \mathbb{R}^n$,

$$\alpha(u) x \overline{u} = (-u)x(-u) = uxu.$$

If $u \cdot x = 0$, then by Lemma 3.29,

$$\alpha(u) x \overline{u} = -u^2 x = -(-1)x = x, \tag{3.14a}$$

since $u^2 = -|u|^2 = -1$. On the other hand, if $x = tu$ for some $t \in \mathbb{R}$, then

$$\alpha(u) x \overline{u} = t u^2 u = -tu. \tag{3.14b}$$

So in particular $\alpha(u) x \overline{u} \in \mathbb{R}^n$. This allows us to define a function

$$\rho_u \colon \mathbb{R}^n \longrightarrow \mathbb{R}^n; \quad \rho_u(x) = \alpha(u) x \overline{u} = uxu.$$

Similarly for $u \in \mathrm{Pin}(n)$, we can consider $\alpha(u) x \overline{u}$; if $u = u_1 \cdots u_r$ for $u_1, \ldots, u_r \in \mathbb{S}^{n-1}$, we have

$$\begin{aligned} \alpha(u) x \overline{u} &= \alpha(u_1 \cdots u_r) x \overline{u_1 \cdots u_r} \\ &= ((-1)^r u_1 \cdots u_r) x ((-1)^r u_r \cdots u_1) \\ &= \rho_{u_1} \circ \cdots \circ \rho_{u_r}(x) \in \mathbb{R}^n. \end{aligned} \tag{3.15}$$

So there is a linear transformation

$$\rho_u \colon \mathbb{R}^n \longrightarrow \mathbb{R}^n; \quad \rho_u(x) = \alpha(u) x \overline{u}.$$

PROPOSITION 3.30. *For $u \in \mathrm{Pin}(n)$, $\rho_u \colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an isometry, i.e., an element of* $\mathrm{O}(n)$.

PROOF. By Equation (3.15) it suffices to show this for $u \in \mathbb{S}^{n-1}$. Now Equations (3.14) show that geometrically $\rho_u \colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is reflection in the hyperplane orthogonal to u which is an isometry. □

Since each $\rho_u \in \mathrm{O}(n)$ we actually have a continuous homomorphism

$$\rho \colon \mathrm{Pin}(n) \longrightarrow \mathrm{O}(n); \quad \rho(u) = \rho_u.$$

PROPOSITION 3.31. $\rho \colon \mathrm{Pin}(n) \longrightarrow \mathrm{O}(n)$ *is surjective with kernel* $\ker \rho = \{1, -1\}$.

PROOF. The observation in the proof of Proposition 3.30 shows that reflection in the hyperplane orthogonal to $u \in \mathbb{S}^{n-1}$ has the form ρ_u. Surjectivity follows using the standard fact that every element of $\mathrm{O}(n)$ is a composition of reflections in hyperplanes.

Suppose that for some $u_1, \ldots, u_k \in \mathbb{S}^{n-1}$, $u = u_1 \cdots u_k \in \ker \rho$, i.e., $\rho_u = I_n$. Then

$$1 = \det \rho_u = \det(\rho_{u_1} \cdots \rho_{u_k}) = \det \rho_{u_1} \cdots \det \rho_{u_k}.$$

Each ρ_{u_r} is a reflection and so has $\det \rho_{u_r} = -1$. These facts imply k must be even, $u \in \mathrm{Cl}_n^+$, and then by Equation (3.13),

$$u^{-1} = u_k \cdots u_1 = \overline{u}.$$

So for any $x \in \mathbb{R}^n$ we have

$$x = \rho_u(x) = uxu^{-1},$$

which implies that

$$xu = ux.$$

For each $r = 1, \ldots, n$ we can write

$$u = a_r + e_r b_r = (a_r^+ + e_r b_r^-) + (a_r^- + e_r b_r^+),$$

where $a_r, b_r \in \mathrm{Cl}_n$ do not involve e_r in their expansions in terms of the monomial bases of Equation (3.9). On taking $x = e_r$ we obtain

$$e_r(a_r + e_r b_r) = (a_r + e_r b_r)e_r,$$

giving

$$\begin{aligned} a_r + e_r b_r &= -e_r(a_r + e_r b_r)e_r \\ &= -e_r a_r e_r - e_r^2 b_r e_r \\ &= -e_r^2 a_r - e_r b_r \\ &= a_r - e_r b_r \\ &= (a_r^+ - e_r b_r^-) + (a_r^- - e_r b_r^-) \\ &= a_r - e_r b_r, \end{aligned}$$

where we use the fact that for each $e_s \neq e_r$, $e_s e_r = -e_r e_s$. Thus we have $b_r = 0$ and so $u = a_r$ does not involve e_r. But this applies for all r, so $u = t1$ for some $t \in \mathbb{R}$. Since $\overline{u} = t1$,

$$t^2 1 = u\overline{u} = (-1)^k = 1,$$

by Equation (3.13) and the fact that k is even. This shows that $t = \pm 1$ and so $u = \pm 1$. □

For $n \geqslant 1$, the *spinor groups* are defined by

$$\mathrm{Spin}(n) = \rho^{-1}\,\mathrm{SO}(n) \leqslant \mathrm{Pin}(n).$$

THEOREM 3.32. $\mathrm{Spin}(n)$ *is a compact, path connected, closed normal subgroup of* $\mathrm{Pin}(n)$*, satisfying*

$$\mathrm{Spin}(n) = \mathrm{Pin}(n) \cap \mathrm{Cl}_n^+, \tag{3.16a}$$

$$\mathrm{Pin}(n) = \mathrm{Spin}(n) \cup e_r\,\mathrm{Spin}(n), \tag{3.16b}$$

for any $r = 1, \ldots, n$.

Furthermore, when $n \geqslant 3$ *the fundamental group of* $\mathrm{Spin}(n)$ *is trivial,* $\pi_1\,\mathrm{Spin}(n) = 1$.

PROOF. We only discuss connectivity. Recall that the sphere $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n \subseteq \mathrm{Cl}_n$ is path connected. Choose a base point $u_0 \in \mathbb{S}^{n-1}$. Now for an element $u = u_1 \cdots u_k \in \mathrm{Spin}(n)$ with $u_1, \ldots, u_k \in \mathbb{S}^{n-1}$, as noted in the proof of Proposition 3.31, we must have k even, say $k = 2m$. In fact, we might as well take m to be even since $u = u(-w)w$ for any $w \in \mathbb{S}^{n-1}$. Then there are continuous paths

$$p_r \colon [0,1] \longrightarrow \mathbb{S}^{n-1} \quad (r = 1, \ldots, 2m),$$

for which $p_r(0) = u_0$ and $p_r(1) = u_r$. Then

$$p \colon [0,1] \longrightarrow \mathbb{S}^{n-1} \quad p(t) = p_1(t) \cdots p_{2m}(t)$$

is a continuous path in $\mathrm{Pin}(n)$ with

$$p(0) = u_0^{2m} = (-1)^m = 1, \quad p(1) = u.$$

But $t \mapsto \rho(p(t))$ is a continuous path in $\mathrm{O}(n)$ with $\rho(p(0)) \in \mathrm{SO}(n)$, hence $\rho(p(t)) \in \mathrm{SO}(n)$ for all t. This shows that p is a path in $\mathrm{Spin}(n)$. So every element $u \in \mathrm{Spin}(n)$ can be connected to 1 and therefore $\mathrm{Spin}(n)$ is path connected.

The equations of (3.16) follow from details in proof of Proposition 3.31.

The final statement involves homotopy theory and is not proved here. It should be compared with the fact that for $n \geqslant 3$, $\pi_1 \mathrm{SO}(n) \cong \{1, -1\}$ and in fact the map is a *universal covering*. $\square$

The double covering maps $\rho \colon \mathrm{Spin}(n) \longrightarrow \mathrm{SO}(n)$ generalize the case of $\mathrm{SU}(2) \longrightarrow \mathrm{SO}(3)$ discussed in Section 5 of Chapter 2. In fact, around each element $u \in$ there is an open neighbourhood $N_u \subseteq \mathrm{Spin}(n)$ for which $\rho \colon N_u \longrightarrow \rho N_u$ is a homeomorphism, and actually a diffeomorphism. This implies the following.

PROPOSITION 3.33. *The derivative* $\mathrm{d}\,\rho \colon \mathfrak{spin}(n) \longrightarrow \mathfrak{so}(n)$ *is an isomorphism of* $\mathbb{R}$*-Lie algebras and*

$$\dim \mathrm{Spin}(n) = \dim \mathrm{SO}(n) = \binom{n}{2}.$$

7. The centres of spinor groups

Recall that for a group G the *centre* of G is

$$\mathrm{C}(G) = \{c \in G : \forall g \in G,\ gc = cg\}.$$

Then $\mathrm{C}(G) \lhd G$. It is well known that for groups $\mathrm{SO}(n)$ with $n \geqslant 3$ we have

PROPOSITION 3.34. *For* $n \geqslant 3$*,*

$$\mathrm{C}(\mathrm{SO}(n)) = \{tI_n : t = \pm 1,\ t^n = 1\} = \begin{cases} \{I_n\} & \text{if } n \text{ is odd,} \\ \{\pm I_n\} & \text{if } n \text{ is even.} \end{cases}$$

PROPOSITION 3.35. *For* $n \geqslant 3$*,*

$$\mathrm{C}(\mathrm{Spin}(n)) = \begin{cases} \{\pm 1\} & \text{if } n \text{ is odd,} \\ \{\pm 1, \pm e_1 \cdots e_n\} & \text{if } n \equiv 2 \bmod 4, \\ \{\pm 1, \pm e_1 \cdots e_n\} & \text{if } n \equiv 0 \bmod 4. \end{cases}$$

$$\cong \begin{cases} \mathbb{Z}/2 & \text{if } n \text{ is odd,} \\ \mathbb{Z}/4 & \text{if } n \equiv 2 \bmod 4, \\ \mathbb{Z}/2 \times \mathbb{Z}/2 & \text{if } n \equiv 0 \bmod 4. \end{cases}$$

PROOF. If $g \in \mathrm{C}(\mathrm{Spin}(n))$, then since $\rho\colon \mathrm{Spin}(n) \longrightarrow \mathrm{SO}(n)$, $\rho(g) \in \mathrm{C}(\mathrm{SO}(n))$. As $\pm 1 \in \mathrm{C}(\mathrm{Spin}(n))$, this gives $|\,\mathrm{C}(\mathrm{Spin}(n))| = 2|\,\mathrm{C}(\mathrm{SO}(n))|$ and indeed

$$\mathrm{C}(\mathrm{Spin}(n)) = \rho^{-1}\,\mathrm{C}(\mathrm{SO}(n)).$$

For n even,

$$(\pm e_1 \cdots e_n)^2 = e_1 \cdots e_n e_1 \cdots e_n = (-1)^{\binom{n}{2}} e_1^2 \cdots e_n^2 = (-1)^{\binom{n}{2}+n} = (-1)^{\binom{n+1}{2}}.$$

Since

$$\binom{n+1}{2} = \frac{(n+1)n}{2} \equiv \begin{cases} 0 \bmod 2 & \text{if } n \equiv 2 \bmod 4, \\ 1 \bmod 2 & \text{if } n \equiv 0 \bmod 4, \end{cases}$$

this implies

$$(\pm e_1 \cdots e_n)^2 = \begin{cases} 1 & \text{if } n \equiv 2 \bmod 4, \\ -1 & \text{if } n \equiv 0 \bmod 4. \end{cases}$$

Hence for n even, the multiplicative order of $\pm e_1 \cdots e_n$ is 1 or 2 depending on the congruence class of n modulo 4. This gives the stated groups. □

We remark that Spin(1) and Spin(2) are abelian.

8. Finite subgroups of spinor groups

Each orthogonal group $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ contains finite subgroups. For example, when $n = 2, 3$, these correspond to symmetry groups of compact plane figures and solids. Elements of $\mathrm{SO}(n)$ are often called *direct isometries,* while elements of $\mathrm{O}(n)^-$ are called *indirect isometries.* The case of $n = 3$ is explored in the Problem Set for this chapter. Here we make some remarks about the symmetric and alternating groups.

Recall that for each $n \geqslant 1$ the *symmetric group* S_n is the group of all permutations of the set $\mathbf{n} = 1, \ldots, n$. The corresponding *alternating group* $A_n \leqslant S_n$ is the subgroup consisting of all *even* permutations, i.e., the elements $\sigma \in S_n$ for which $\mathrm{sign}(\sigma) = 1$ where $\mathrm{sign}\colon S_n \longrightarrow \{\pm 1\}$ is the sign homomorphism.

For a field $\Bbbk$, we can make S_n act on $\Bbbk^n$ by linear transformations:

$$\sigma \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{\sigma^{-1}(1)} \\ x_{\sigma^{-1}(2)} \\ \vdots \\ x_{\sigma^{-1}(n)} \end{bmatrix}.$$

Notice that $\sigma(\mathbf{e}_r) = \mathbf{e}_{\sigma(r)}$. The matrix $[\sigma]$ of the linear transformation induced by σ with respect to the basis of $\mathbf{e}_r$'s has all its entries 0 or 1, with exactly one 1 in each row and column. For example, when $n = 3$,

$$[(1\;2\;3)] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad [(1\;3)] = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

When $\Bbbk = \mathbb{R}$ each of these matrices is orthogonal, while when $\Bbbk = \mathbb{C}$ it is unitary. For a given n we can view S_n as the subgroup of $\mathrm{O}(n)$ or $\mathrm{U}(n)$ consisting of all such matrices which are usually called *permutation matrices.*

Proposition 3.36. *For $\sigma \in S_n$,*

$$\operatorname{sign}(\sigma) = \det([\sigma]).$$

Hence we have

$$A_n = \begin{cases} \mathrm{SO}(n) \cap S_n & \text{if } \Bbbk = \mathbb{R}, \\ \mathrm{SU}(n) \cap S_n & \text{if } \Bbbk = \mathbb{C}. \end{cases}$$

Recall that if $n \geqslant 5$, A_n is a simple group.

As $\rho\colon \mathrm{Pin}(n) \longrightarrow \mathrm{O}(n)$ is onto, there are finite subgroups $\widetilde{S}_n = \rho^{-1}S_n \leqslant \mathrm{Pin}(n)$ and $\widetilde{A}_n = \rho^{-1}A_n \leqslant \mathrm{Spin}(n)$ for which there are surjective homomorphisms $\rho\colon \widetilde{S}_n \longrightarrow S_n$ and $\rho\colon \widetilde{A}_n \longrightarrow A_n$ whose kernels contain the two elements ± 1. Note that $|\widetilde{S}_n| = 2 \cdot n!$, while $|\widetilde{A}_n| = n!$. However, for $n \geqslant 4$, there are no homomorphisms $\tau\colon S_n \longrightarrow \widetilde{S}_n$, $\tau\colon A_n \longrightarrow \widetilde{A}_n$ for which $\rho \circ \tau = \mathrm{Id}$.

$$\begin{tikzcd} S_n \arrow[r, "\tau"] \arrow[dr, "\mathrm{Id}_{S_n}"'] & \widetilde{S}_n \arrow[d, "\rho"] \\ & S_n \end{tikzcd} \qquad \begin{tikzcd} A_n \arrow[r, "\tau"] \arrow[dr, "\mathrm{Id}_{A_n}"'] & \widetilde{A}_n \arrow[d, "\rho"] \\ & A_n \end{tikzcd}$$

Similar considerations apply to other finite subgroups of $\mathrm{O}(n)$.

In $\mathrm{Cl}_n^\times$ we have a subgroup E_n consisting of all the elements

$$\pm e_{i_1} \cdots e_{i_r} \quad (1 \leqslant i_1 < \cdots < i_r \leqslant n,\ 0 \leqslant r).$$

The order of this group is $|E_n| = 2^{n+1}$ and as it contains ± 1, its image under $\rho\colon \mathrm{Pin}(n) \longrightarrow \mathrm{O}(n)$ is $\overline{E}_n = \rho E_n$ of order $|\overline{E}_n| = 2^n$. In fact, $\{\pm 1\} = \mathrm{C}(E_n)$ is also the commutator subgroup since $e_i e_j e_i^{-1} e_j^{-1} = -1$ and so $\overline{E}_n$ is abelian. Every non-trivial element in $\overline{E}_n$ has order 2 since $e_i^2 = -1$, hence $\overline{E}_n \leqslant \mathrm{O}(n)$ is an *elementary* 2*-group,* i.e., it is isomorphic to $(\mathbb{Z}/2)^n$. Each element $\rho(e_r) \in \mathrm{O}(n)$ is a generalized permutation matrices with all its non-zero entries on the main diagonal. There is also a subgroup $\overline{E}{}_n^0 = \rho E_n^0 \leqslant \mathrm{SO}(n)$ of order 2^{n-1}, where

$$E_n^0 = E_n \cap \mathrm{Spin}(n).$$

In fact $\overline{E}{}_n^0$ is isomorphic to $(\mathbb{Z}/2)^{n-1}$. These groups E_n and E_n^0 are non-abelian and fit into exact sequences of the form

$$1 \to \mathbb{Z}/2 \longrightarrow E_n \longrightarrow (\mathbb{Z}/2)^n \to 1, \quad 1 \to \mathbb{Z}/2 \longrightarrow E_n^0 \longrightarrow (\mathbb{Z}/2)^{n-1} \to 1,$$

in which each kernel $\mathbb{Z}/2$ is equal to the centre of the corresponding group E_n or E_n^0. This means they are *extraspecial* 2*-groups.*

CHAPTER 4

Matrix groups as Lie groups

In this chapter we will discuss the basic ideas of *smooth manifolds* and *Lie groups*. Our main aim is to prove a theorem which identifies every real matrix of $\mathrm{GL}_n(\mathbb{R})$ is a Lie subgroup.

1. Smooth manifolds

DEFINITION 4.1. A continuous map $g\colon V_1 \longrightarrow V_2$ where each $V_k \subseteq \mathbb{R}^{m_k}$ is open, is called *smooth* if it is infinitely differentiable. A smooth map g is a *diffeomorphism* if it has an inverse g^{-1} which is also smooth.

Let M be a separable Hausdorff topological space.

DEFINITION 4.2. A homeomorphism $f\colon U \longrightarrow V$ where $U \subseteq M$ and $V \subseteq \mathbb{R}^n$ are open subsets, is called an *n-chart* for U.

If $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ is an open covering of M and $\mathcal{F} = \{f_\alpha\colon U_\alpha \longrightarrow V_\alpha\}$ is a collection of charts, then $\mathcal{F}$ is called an *atlas* for M if, whenever $U_\alpha \cap U_\beta \neq \emptyset$,

$$f_\beta \circ f_\alpha^{-1}\colon f_\alpha(U_\alpha \cap U_\beta) \longrightarrow f_\beta(U_\alpha \cap U_\beta)$$

is a diffeomorphism.

(4.1)
$$\begin{tikzcd} & U_\alpha \cap U_\beta \arrow[dr, "f_\beta"] & \\ f_\alpha(U_\alpha \cap U_\beta) \arrow[ur, "f_\alpha^{-1}"] \arrow[rr, "f_\beta\circ f_\alpha^{-1}"'] & & f_\beta(U_\alpha \cap U_\beta) \end{tikzcd}$$

We will sometimes denote an atlas by $(M, \mathcal{U}, \mathcal{F})$ and refer to it as a *smooth manifold of dimension n* or *smooth n-manifold.*

DEFINITION 4.3. Let $(M, \mathcal{U}, \mathcal{F})$ and $(M', \mathcal{U}', \mathcal{F}')$ be atlases on topological spaces M and M'. A *smooth map* $h\colon (M, \mathcal{U}, \mathcal{F}) \longrightarrow (M', \mathcal{U}', \mathcal{F}')$ is a continuous map $h\colon M \longrightarrow M'$ such that for each pair α, α' with $h(U_\alpha) \cap U'_{\alpha'} \neq \emptyset$, the composite

$$f'_{\alpha'} \circ h \circ f_\alpha^{-1}\colon f_\alpha(h^{-1}U'_{\alpha'}) \longrightarrow V'_{\alpha'}$$

is smooth.

(4.2)
$$\begin{tikzcd} f_\alpha(h^{-1}U'_{\alpha'}) \arrow[r, "f'_{\alpha'}\circ h\circ f_\alpha^{-1}"] \arrow[d, "f_\alpha^{-1}"'] & V'_{\alpha'} \arrow[d, "{f'_{\alpha'}}^{-1}"] \\ h^{-1}U'_{\alpha'} \arrow[r, "h"'] & h(U_\alpha) \cap U'_{\alpha'} \end{tikzcd}$$

2. Tangent spaces and derivatives

Let $(M, \mathcal{U}, \mathcal{F})$ be a smooth n-manifold and $p \in M$. Let $\gamma\colon (a, b) \longrightarrow M$ be a continuous curve with $a < 0 < b$.

DEFINITION 4.4. γ is *differentiable* at $t \in (a, b)$ if for every chart $f \colon U \longrightarrow V$ with $\gamma(t) \in U$, the curve $f \circ \gamma \colon (a, b) \longrightarrow V$ is differentiable at $t \in (a, b)$, i.e., $(f \circ \gamma)'(t)$ exists. γ is *smooth* at $t \in (a, b)$ if all the derivatives of $f \circ \gamma$ exists at t.

The curve γ is *differentiable* if it is differentiable at all points in (a, b). Similarly γ is *smooth* if it is smooth at all points in (a, b).

LEMMA 4.5. *Let $f_0 \colon U_0 \longrightarrow V_0$ be a chart with $\gamma(t) \in U_0$ and suppose that*

$$f_0 \circ \gamma \colon (a, b) \cap f_0^{-1} V_0 \longrightarrow V_0$$

is differentiable/smooth at t. Then for any chart $f \colon U \longrightarrow V$ with $\gamma(t) \in U$,

$$f \circ \gamma \colon (a, b) \cap f^{-1} V \longrightarrow V$$

is differentiable/smooth at t.

PROOF. This follows using the ideas of Definition 4.2. The smooth composite $f \circ \alpha$ is defined on a subinterval of (a, b) containing t and there is the usual *Chain* or *Function of a Function Rule* for the derivative of the composite

$$(f\gamma)'(t) = \mathrm{Jac}_{f \circ f_0^{-1}}(f_0\gamma(t))(f_0\gamma)'(t). \tag{4.3}$$

Here, for a differentiable function

$$h \colon W_1 \longrightarrow W_2; \quad h(\mathbf{x}) = \begin{bmatrix} h_1(\mathbf{x}) \\ \vdots \\ h_{m_2}(\mathbf{x}) \end{bmatrix}$$

with $W_1 \subseteq \mathbb{R}^{m_1}$ and $W_2 \subseteq \mathbb{R}^{m_2}$ open subsets, and $\mathbf{x} \in W_1$, the *Jacobian matrix* is

$$\mathrm{Jac}_h(\mathbf{x}) = \left[\frac{\partial h_i}{\partial x_j}(\mathbf{x})\right] \in \mathrm{M}_{m_2, m_1}(\mathbb{R}).$$

□

If $\gamma(0) = p$ and γ is differentiable at 0, then for any (and hence every) chart $f_0 \colon U_0 \longrightarrow V_0$ with $\gamma(0) \in U_0$, there is a derivative vector $\mathbf{v}_0 = (f\gamma)'(0) \in \mathbb{R}^n$. In passing to another chart $f \colon U \longrightarrow V$ with $\gamma(0) \in U$ by Equation (4.3) we have

$$(f\gamma)'(0) = \mathrm{Jac}_{f f_0^{-1}}(f_0\gamma(0))(f_0\gamma)'(0).$$

In order to define the notion of the *tangent space* $\mathrm{T}_p M$ to the manifold M at p, we consider all pairs of the form

$$((f\gamma)'(0), f \colon U \longrightarrow V)$$

where $\gamma(0) = p \in U$, and then impose an equivalence relation $\sim$ under which

$$((f_1\gamma)'(0), f_1 \colon U_1 \longrightarrow V_1) \sim ((f_2\gamma)'(0), f_2 \colon U_2 \longrightarrow V_2).$$

Since

$$(f_2\gamma)'(0) = \mathrm{Jac}_{f_2 f_1^{-1}}(f_1\gamma(0))(f_1\gamma)'(0),$$

we can also write this as

$$(\mathbf{v}, f_1 \colon U_1 \longrightarrow V_1) \sim (\mathrm{Jac}_{f_2 f_1^{-1}}(f_1(p))\mathbf{v}, f_2 \colon U_2 \longrightarrow V_2),$$

whenever there is a curve α in M for which

$$\gamma(0) = p, \quad (f_1\gamma)'(0) = \mathbf{v}.$$

The set of equivalence classes is $\mathrm{T}_p M$ and we will sometimes denote the equivalence class of $(\mathbf{v}, f\colon U \longrightarrow V)$ by $[\mathbf{v}, f\colon U \longrightarrow V]$.

PROPOSITION 4.6. *For $p \in M$, $\mathrm{T}_p M$ is an $\mathbb{R}$-vector space of dimension n.*

PROOF. For any chart $f\colon U \longrightarrow V$ with $p \in U$, we can identify the elements of $\mathrm{T}_p M$ with objects of the form $(\mathbf{v}, f\colon U \longrightarrow V)$. Every $\in \mathbb{R}^n$ arises as the derivative of a curve $\overline{\gamma}\colon (-\varepsilon, \varepsilon) \longrightarrow V$ for which $\overline{\gamma}(0) = f(p)$. For example for small enough ε, we could take

$$\overline{\gamma}(t) = f(p) + t\mathbf{v}.$$

There is an associated curve in M,

$$\gamma\colon (-\varepsilon, \varepsilon) \longrightarrow M; \quad \gamma(t) = f^{-1}\overline{\gamma}(t),$$

for which $\gamma(0) = p$. So using such a chart we can identify $\mathrm{T}_p M$ with $\mathbb{R}^n$ by

$$[\mathbf{v}, f\colon U \longrightarrow V] \longleftrightarrow \mathbf{v}.$$

The same argument as used to prove Proposition 2.9, shows that $\mathrm{T}_p M$ is a vector space and that the above correspondence is a linear isomorphism. $\square$

Let $h\colon (M, \mathcal{U}, \mathcal{F}) \longrightarrow (M', \mathcal{U}', \mathcal{F}')$ be a smooth map between manifolds of dimensions n, n'. We will use the notation of Definition 4.3. For $p \in M$, consider a pair of charts as in Diagram (4.2) with $p \in U_\alpha$ and $h(p) \in U'_{\alpha'}$. Since $h_{\alpha',\alpha} = f'_{\alpha'} \circ h \circ f_\alpha^{-1}$ is differentiable, the Jacobian matrix $\mathrm{Jac}_{h_{\alpha',\alpha}}(f_\alpha(p))$ has an associated $\mathbb{R}$-linear transformation

$$\mathrm{d}\, h_{\alpha',\alpha}\colon \mathbb{R}^n \longrightarrow \mathbb{R}^{n'}; \quad \mathrm{d}\, h_{\alpha',\alpha}(\mathbf{x}) = \mathrm{Jac}_{h_{\alpha',\alpha}}(f_\alpha(p))\mathbf{x}.$$

It is easy to verify that this passes to equivalence classes to give a well defined $\mathbb{R}$-linear transformation

$$\mathrm{d}\, h_p\colon \ \mathrm{T}_p M \longrightarrow \mathrm{T}_{h(p)} M'.$$

The following result summarises the properties of the derivative and should be compared with Proposition 2.16.

PROPOSITION 4.7. *Let $h\colon (M, \mathcal{U}, \mathcal{F}) \longrightarrow (M', \mathcal{U}', \mathcal{F}')$ and $g\colon (M', \mathcal{U}', \mathcal{F}') \longrightarrow (M'', \mathcal{U}'', \mathcal{F}'')$ be smooth maps between manifolds M, M', M'' of dimensions n, n', n''.*

a) *For each $p \in M$ there is an $\mathbb{R}$-linear transformation $\mathrm{d}\, h_p\colon \ \mathrm{T}_p M \longrightarrow \mathrm{T}_{h(p)} M'$.*
b) *For each $p \in M$,*

$$\mathrm{d}\, g_{h(p)} \circ \mathrm{d}\, h_p = \mathrm{d}(g \circ h)_p.$$

c) *For the identity map $\mathrm{Id}\colon M \longrightarrow M$ and $p \in M$,*

$$\mathrm{d}\,\mathrm{Id}_p = \mathrm{Id}_{\mathrm{T}_p M}.$$

DEFINITION 4.8. Let $(M, \mathcal{U}, \mathcal{F})$ be a manifold of dimension n. A subset $N \subseteq M$ is a *submanifold of dimension k* if for every $p \in N$ there is an open neighbourhood $U \subseteq M$ of p and an n-chart $f\colon U \longrightarrow V$ such that

$$p \in f^{-1}(V \cap \mathbb{R}^k) = N \cap U.$$

For such an N we can form k-charts of the form

$$f_0\colon N \cap U \longrightarrow V \cap \mathbb{R}^k; \quad f_0(x) = f(x).$$

We will denote this manifold by $(N, \mathcal{U}_N, \mathcal{F}_N)$. The following result is immediate.

PROPOSITION 4.9. *For a submanifold $N \subseteq M$ of dimension k, the inclusion function* $\mathrm{incl}\colon N \longrightarrow M$ *is smooth and for every $p \in N$, $\mathrm{d}\,\mathrm{incl}_p\colon \mathrm{T}_p N \longrightarrow \mathrm{T}_p M$ is an injection.*

The next result allows us to recognise submanifolds as inverse images of points under smooth mappings.

THEOREM 4.10 (Implicit Function Theorem for manifolds). *Let $h\colon (M, \mathcal{U}, \mathcal{F}) \longrightarrow (M', \mathcal{U}', \mathcal{F}')$ be a smooth map between manifolds of dimensions n, n'. Suppose that for some $q \in M'$, $\mathrm{d}\,h_p\colon \mathrm{T}_p M \longrightarrow \mathrm{T}_{h(p)} M'$ is surjective for every $p \in N = h^{-1}q$. Then $N \subseteq M$ is submanifold of dimension $n - n'$ and the tangent space at $p \in N$ is given by $\mathrm{T}_p N = \ker \mathrm{d}\,h_p$.*

PROOF. This follows from the Implicit Function Theorem of Calculus. □

Another important application of the Implicit Function Theorem is to the following version of the Inverse Function Theorem.

THEOREM 4.11 (Inverse Function Theorem for manifolds). *Let $h\colon (M, \mathcal{U}, \mathcal{F}) \longrightarrow (M', \mathcal{U}', \mathcal{F}')$ be a smooth map between manifolds of dimensions n, n'. Suppose that for some $p \in M$, $\mathrm{d}\,h_p\colon \mathrm{T}_p M \longrightarrow \mathrm{T}_{h(p)} M'$ is an isomorphism. Then there is an open neighbourhood $U \subseteq M$ of p and an open neighbourhood $V \subseteq M'$ of $h(p)$ such that $hU = V$ and the restriction of h to the map $h_1\colon U \longrightarrow V$ is a diffeomorphism. In particular, the derivative $\mathrm{d}\,h_p\colon \mathrm{T}_p \longrightarrow \mathrm{T}_{h(p)}$ is an $\mathbb{R}$-linear isomorphism and $n = n'$.*

When this occurs we say that h is locally a diffeomorphism at p.

EXAMPLE 4.12. Consider the exponential function $\exp\colon \mathrm{M}_n(\mathbb{R}) \longrightarrow \mathrm{GL}_n(\mathbb{R})$. Then by Proposition 1.36,

$$\mathrm{d}\exp_O(X) = X.$$

Hence exp is locally a diffeomorphism at O.

3. Lie groups

The following should be compared with Definition 1.14.

DEFINITION 4.13. Let G be a smooth manifold which is also a topological group with multiplication map $\mathrm{mult}\colon G \times G \longrightarrow G$ and inverse map $\mathrm{inv}\colon G \longrightarrow G$ and view $G \times G$ as the product manifold. Then G is a *Lie group* if $\mathrm{mult}, \mathrm{inv}$ are smooth maps.

DEFINITION 4.14. Let G be a Lie group. A closed subgroup $H \leqslant G$ that is also a submanifold is called a *Lie subgroup* of G. It is then automatic that the restrictions to H of the multiplication and inverse maps on G are smooth, hence H is also a Lie group.

For a Lie group G, at each $g \in G$ there is a tangent space $\mathrm{T}_g G$ and when G is a matrix group this agrees with the tangent space defined in Chapter 2. We adopt the notation $\mathfrak{g} = \mathrm{T}_1 G$ for the tangent space at the identity of G. A smooth homomorphism of Lie groups $G \longrightarrow H$ has the properties of a Lie homomorphism as in Definition 2.15.

For a Lie group G, let $g \in G$. There are following three functions are of great importance.

$$\mathrm{L}_g\colon G \longrightarrow G;\quad \mathrm{L}_g(x) = gx. \qquad \text{(Left multiplication)}$$

$$\mathrm{R}_g\colon G \longrightarrow G;\quad \mathrm{R}_g(x) = xg. \qquad \text{(Right multiplication)}$$

$$\chi_g\colon G \longrightarrow G;\quad \chi_g(x) = gxg^{-1}. \qquad \text{(Conjugation)}$$

PROPOSITION 4.15. *For $g \in G$, the maps L_g, R_g, χ_g are all diffeomorphisms with inverses*

$$\mathrm{L}_g^{-1} = \mathrm{L}_{g^{-1}}, \quad \mathrm{R}_g^{-1} = \mathrm{R}_{g^{-1}}, \quad \chi_g^{-1} = \chi_{g^{-1}}.$$

PROOF. Charts for $G \times G$ have the form

$$\varphi_1 \times \varphi_2 \colon U_1 \times U_2 \longrightarrow V_1 \times V_2,$$

where $\varphi_k \colon U_k \longrightarrow V_k$ are charts for G. Now suppose that $\mu U_1 \times U_2 \subseteq W \subseteq G$ where there is a chart $\theta \colon W \longrightarrow Z$. By assumption, the composition

$$\theta \circ \mu \circ (\varphi_1 \times \varphi_2)^{-1} = \theta \circ \mu \circ (\varphi_1^{-1} \times \varphi_2^{-1}) \colon V_1 \times V_2 \longrightarrow Z.$$

is smooth. Then $\mathrm{L}_g(x) = \mu(g, x)$, so if $g \in U_1$ and $x \in U_2$, we have

$$\mathrm{L}_g(x) = \theta^{-1}(\theta \circ \mathrm{L}_g \circ \varphi_2^{-1}) \circ \varphi_2(x).$$

But then it is clear that

$$\theta \circ \varphi_2^{-1} \colon V_2 \longrightarrow Z$$

is smooth since it is obtained from $\theta \circ \mu \circ (\varphi_1 \times \varphi_2)^{-1}$, but treating the first variable as a constant.

A similar argument deals with R_g. For χ_g, notice that

$$\chi_g = \mathrm{L}_g \circ \mathrm{R}_g = \mathrm{R}_g \circ \mathrm{L}_g,$$

and a composite of smooth maps is smooth. □

The derivatives of these maps at the identity $1 \in G$ are worth studying. Since L_g and R_g are diffeomorphisms with inverses $\mathrm{L}_{g^{-1}}$ and $\mathrm{R}_{g^{-1}}$,

$$\mathrm{d}(\mathrm{L}_g)_1,\ \mathrm{d}(\mathrm{R}_g)_1 \colon \mathfrak{g} = \mathrm{T}_1\, G \longrightarrow \mathrm{T}_g\, G$$

are $\mathbb{R}$-linear isomorphisms. We can use this to identify every tangent space of G with $\mathfrak{g}$. The conjugation map χ_g fixes 1, so it induces an $\mathbb{R}$-linear isomorphism

$$\mathrm{Ad}_g = \mathrm{d}(\chi_g)_1 \colon \mathfrak{g} \longrightarrow \mathfrak{g}.$$

This is the *adjoint action* of $g \in G$ on $\mathfrak{g}$. For G a matrix group this is the same as defined in Chapter 2. There is also a natural Lie bracket $[\ ,\ ]$ defined on $\mathfrak{g}$, making it into an $\mathbb{R}$-Lie algebra. The construction follows that for matrix groups. The following Lie group analogue of Theorem 2.18 holds.

THEOREM 4.16. *Let G, H be Lie groups and $\varphi \colon G \longrightarrow H$ a Lie homomorphism. Then the derivative is a homomorphism of Lie algebras. In particular, if $G \leqslant H$ is a Lie subgroup, the inclusion map* $\mathrm{incl} \colon G \longrightarrow H$ *induces an injection of Lie algebras* $\mathrm{d}\,\mathrm{incl} \colon \mathfrak{g} \longrightarrow \mathfrak{h}$.

4. Some examples of Lie groups

EXAMPLE 4.17. For $\Bbbk = \mathbb{R}, \mathbb{C}$, $\mathrm{GL}_n(\Bbbk)$ is a Lie group.

PROOF. This follows from Proposition 1.13(a) which shows that $\mathrm{GL}_n(\Bbbk) \subseteq \mathrm{M}_n(\Bbbk)$ is an open subset where as usual $\mathrm{M}_n(\Bbbk)$ we identify with $\Bbbk^{n^2}$. For charts we take the open sets $U \subseteq \mathrm{GL}_n(\Bbbk)$ and the identity function $\mathrm{Id} \colon U \longrightarrow U$. The tangent space at each point $A \in \mathrm{GL}_n(\Bbbk)$ is just $\mathrm{M}_n(\Bbbk)$. So the notions of tangent space and dimension of Sections 1,2 and of Chapter 2 agree here. The multiplication and inverse maps are obviously smooth as they are defined by polynomial and rational functions between open subsets of $\mathrm{M}_n(\Bbbk)$. □

EXAMPLE 4.18. For $\Bbbk = \mathbb{R}, \mathbb{C}$, $\mathrm{SL}_n(\Bbbk)$ is a Lie group.

PROOF. Following Proposition 1.13(b), we have

$$\mathrm{SL}_n(\Bbbk) = \det^{-1} 1 \subseteq \mathrm{GL}_n(\Bbbk)$$

where $\det\colon \mathrm{GL}_n(\Bbbk) \longrightarrow \Bbbk$ is continuous. $\Bbbk$ is a smooth manifold of dimension $\dim_{\mathbb{R}} \Bbbk$ with tangent space $\mathrm{T}_r \mathbb{R} = \mathbb{R}$ at each $r \in \mathbb{R}$ and det is smooth. In order to apply Theorem 4.10, we will first show that the derivative $\mathrm{d}\det_A\colon \mathrm{M}_n(\Bbbk) \longrightarrow \mathbb{R}$ is surjective for every $A \in \mathrm{GL}_n(\Bbbk)$. To do this, consider a smooth curve $\alpha\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{GL}_n(\Bbbk)$ with $\alpha(0) = A$. We calculate the derivative on $\alpha'(0)$ using the formula

$$\mathrm{d}\det_A(\alpha'(0)) = \frac{\mathrm{d}\det\alpha(t)}{\mathrm{d}\,t}_{|t=0}.$$

The modified curve

$$\alpha_0\colon (-\varepsilon, \varepsilon) \longrightarrow \mathrm{GL}_n(\Bbbk); \quad \alpha_0(t) = A^{-1}\alpha(t)$$

satisfies $\alpha_0(0) = I$ and Lemma 2.19 implies

$$\mathrm{d}\det_I(\alpha_0'(0)) = \frac{\mathrm{d}\det\alpha_0(t)}{\mathrm{d}\,t}_{|t=0} = \operatorname{tr}\alpha_0'(0).$$

Hence we have

$$\mathrm{d}\det_A(\alpha'(0)) = \frac{\mathrm{d}\det(A\alpha_0(t))}{\mathrm{d}\,t}_{|t=0} = \det A\frac{\mathrm{d}\det(\alpha_0(t))}{\mathrm{d}\,t}_{|t=0} = \det A \operatorname{tr}\alpha_0'(0).$$

So $\mathrm{d}\det_A$ is the $\Bbbk$-linear transformation

$$\mathrm{d}\det_A\colon \mathrm{M}_n(\Bbbk) \longrightarrow \Bbbk; \quad \mathrm{d}\det_A(X) = \det A \operatorname{tr}(A^{-1}X).$$

The kernel of this is $\ker \mathrm{d}\det_A = A\mathfrak{sl}_n(\Bbbk)$ and it is also surjective since tr is. In particular this is true for $A \in \mathrm{SL}_n(\Bbbk)$. By Theorem 4.10, $\mathrm{SL}_n(\Bbbk) \longrightarrow \mathrm{GL}_n(\Bbbk)$ is a submanifold and so is a Lie subgroup. Again we find that the two notions of tangent space and dimension agree. □

There is a useful general principle at work in this last proof. Although we state the following two results for matrix groups, it is worth noting that they still apply when $\mathrm{GL}_n(\mathbb{R})$ is replaced by an arbitrary Lie group.

PROPOSITION 4.19 (Left Translation Trick). *Let $F\colon \mathrm{GL}_n(\mathbb{R}) \longrightarrow M$ be a smooth function and suppose that $B \in \mathrm{GL}_n(\mathbb{R})$ satisfies $F(BC) = F(C)$ for all $C \in \mathrm{GL}_n(\mathbb{R})$. Let $A \in \mathrm{GL}_n(\mathbb{R})$ with $\mathrm{d}\,F_A$ surjective. Then $\mathrm{d}\,F_{BA}$ is surjective.*

PROOF. Left multiplication by $B \in G$, $\mathrm{L}_B\colon \mathrm{GL}_n(R) \longrightarrow \mathrm{GL}_n(R)$, is a diffeomorphism, and its derivative at $A \in \mathrm{GL}_n(R)$ is

$$\mathrm{d}(\mathrm{L}_B)\colon \mathrm{M}_n(R) \longrightarrow \mathrm{M}_n(R); \quad \mathrm{d}\,\mathrm{L}_B(X) = BX.$$

By assumption, $F \circ \mathrm{L}_B = F$ as a function on $\mathrm{GL}_n(R)$. Then

$$\begin{aligned}
\mathrm{d}\,F_{BA}(X) &= \mathrm{d}\,F_{BA}(B(B^{-1}X)) \\
&= \mathrm{d}\,F_{BA} \circ \mathrm{d}(\mathrm{L}_B)_A(B^{-1}X) \\
&= \mathrm{d}(F \circ \mathrm{L}_B)_A(B^{-1}X) \\
&= \mathrm{d}\,F_A(B^{-1}X).
\end{aligned}$$

Since left multiplication by B^{-1} on $\mathrm{M}_n(\mathbb{R})$ is surjective, this proves the result. □

PROPOSITION 4.20 (Identity Check Trick). *Let $G \leqslant \mathrm{GL}_n(\mathbb{R})$ be a matrix subgroup, M a smooth manifold and $F\colon \mathrm{GL}_n(\mathbb{R}) \longrightarrow M$ a smooth function with $F^{-1}q = G$ for some $q \in M$. Suppose that for every $B \in G$, $F(BC) = F(C)$ for all $C \in \mathrm{GL}_n(\mathbb{R})$. If $\mathrm{d}\,F_I$ is surjective then $\mathrm{d}\,F_A$ is surjective for all $A \in G$ and $\ker \mathrm{d}\,F_A = A\mathfrak{g}$.*

EXAMPLE 4.21. $\mathrm{O}(n)$ is a Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$.

PROOF. Recall from Chapter 2 that we can specify $\mathrm{O}(n) \subseteq \mathrm{GL}_n(\mathbb{R})$ as the solution set of a family of polynomial equations in n^2 variables arising from the matrix equation $A^T A = I$. In fact, the following $n + \binom{n}{2} = \binom{n+1}{2}$ equations in the entries of the matrix $A = [a_{ij}]$ are sufficient:

$$\sum_{k=1}^{n} a_{kr}^2 - 1 = 0 \quad (1 \leqslant r \leqslant n), \qquad \sum_{k=1}^{n} a_{kr} a_{ks} = 0 \quad (1 \leqslant r < s \leqslant n).$$

We combine the left hand sides of these in some order to give a function $F\colon \mathrm{GL}_n(\mathbb{R}) \longrightarrow \mathbb{R}^{\binom{n+1}{2}}$, for example

$$F([a_{ij}]) = \begin{bmatrix} \sum_{k=1}^{n} a_{k1}^2 - 1 \\ \vdots \\ \sum_{k=1}^{n} a_{kn}^2 - 1 \\ \sum_{k=1}^{n} a_{k1} a_{k2} \\ \vdots \\ \sum_{k=1}^{n} a_{k1} a_{kn} \\ \vdots \\ \sum_{k=1}^{n} a_{k(n-1)} a_{kn} \end{bmatrix}.$$

We need to investigate the derivative $\mathrm{d}\, F_A\colon \mathrm{M}_n(\mathbb{R}) \longrightarrow \mathbb{R}^{\binom{n+1}{2}}$.

By the Identity Check Trick 4.20, to show that $\mathrm{d}\, F_A$ is surjective for all $A \in \mathrm{O}(n)$, it is sufficient to check the case $A = I$. The Jacobian matrix of F at $A = [a_{ij}] = I$ is the $\binom{n+1}{2} \times n^2$ matrix

$$\mathrm{d}\, F_I = \begin{bmatrix} 2 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 2 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\ \vdots & & & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 & 0 \end{bmatrix}$$

where in the top block of n rows, the r th row has a 2 corresponding to the variable a_{rr} and in the bottom block, each row has a 1 in each column corresponding to one of the pair a_{rs}, a_{sr} with $r < s$. The rank of this matrix is $n + \binom{n}{2} = \binom{n+1}{2}$, so $\mathrm{d}\, F_I$ is surjective. It is also true that

$$\ker \mathrm{d}\, F_I = \text{Sk-Sym}_n(\mathbb{R}) = \mathfrak{o}(n).$$

Hence $\mathrm{O}(n) \leqslant \mathrm{GL}_n(\mathbb{R})$ is a Lie subgroup and at each element, the tangent space and dimension agree with those obtained using the definitions of Chapter 2. $\square$

This example is typical of what happens for any matrix group that is a Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$. We summarise the situation in the following, whose proof involves a careful comparison between the ideas introduced in Chapter 2 and the definitions involving manifolds.

THEOREM 4.22. *Let $G \leqslant \mathrm{GL}_n(\mathbb{R})$ be a matrix group which is also a submanifold, hence a Lie subgroup. Then the tangent space to G at I agrees with the Lie algebra $\mathfrak{g}$ and the dimension of the smooth manifold G is $\dim G$; more generally, $\mathrm{T}_A\, G = A\mathfrak{g}$.*

In the next sections, our goal will be to prove the following important result.

THEOREM 4.23. *Let $G \leqslant \mathrm{GL}_n(\mathbb{R})$ be a matrix subgroup. Then G is a Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$.*

The following more general result also holds but we will not give a proof.

THEOREM 4.24. *Let $G \leqslant H$ be a closed subgroup of a Lie group H. Then G is a Lie subgroup of H.*

5. Some useful formula in matrix groups

Let $G \leqslant \mathrm{GL}_n(\mathbb{R})$ be a closed matrix subgroup. We will use Proposition 1.35. Choose r so that $0 < r \leqslant 1/2$ and if $A, B \in \mathrm{N}_{\mathrm{M}_n(\mathbb{R})}(O; r)$ then $\exp(A)\exp(B) \in \exp(\mathrm{N}_{\mathrm{M}_n(\mathbb{R})}(O; 1/2))$. Since exp is injective on $\mathrm{N}_{\mathrm{M}_n(\mathbb{R})}(O; r)$, there is a unique $C \in \mathrm{M}_n(\mathbb{R})$ for which

$$\exp(A)\exp(B) = \exp(C). \tag{4.4}$$

We also set

$$S = C - A - B - \frac{1}{2}[A, B] \in \mathrm{M}_n(\mathbb{R}). \tag{4.5}$$

PROPOSITION 4.25. *$\|S\|$ satisfies*

$$\|S\| \leqslant 65(\|A\| + \|B\|)^3.$$

PROOF. For $X \in \mathrm{M}_n(\mathbb{R})$ we have

$$\exp(X) = I + X + R_1(X),$$

where the remainder term $R_1(X)$ is given by

$$R_1(X) = \sum_{k \leqslant 2} \frac{1}{k!} X^k.$$

Hence,

$$\|R_1(X)\| \leqslant \|X\|^2 \sum_{k \leqslant 2} \frac{1}{k!} \|X\|^{k-2},$$

so if $\|X\| \leqslant 1$,

$$\|R_1(X)\| \leqslant \|X\|^2 \left(\sum_{k \leqslant 2} \frac{1}{k!} \right) = \|X\|^2 (e - 2) < \|X\|^2,$$

Since $\|C\| < 1/2$,

$$\|R_1(C)\| < \|C\|^2. \tag{4.6}$$

Similarly,

$$\exp(C) = \exp(A)\exp(B) = I + A + B + R_1(A, B),$$

where

$$R_1(A, B) = \sum_{k \geqslant 2} \frac{1}{k!} \left(\sum_{r=0}^{k} \binom{k}{r} A^r B^{k-r} \right)$$

giving

$$\begin{aligned}
\|R_1(A, B)\| &\leqslant \sum_{k \geqslant 2} \frac{1}{k!} \left(\sum_{r=0}^{k} \binom{k}{r} \|A\|^r \|B\|^{k-r} \right) \\
&= \sum_{k \geqslant 2} \frac{(\|A\| + \|B\|)^k}{k!} \\
&= (\|A\| + \|B\|)^2 \sum_{k \geqslant 2} \frac{(\|A\| + \|B\|)^{k-2}}{k!} \\
&\leqslant (\|A\| + \|B\|)^2
\end{aligned}$$

since $\|A\| + \|B\| < 1$.

Combining the two ways of writing $\exp(C)$, we have

$$C = A + B + R_1(A, B) - R_1(C) \tag{4.7}$$

and so

$$\begin{aligned}\|C\| &\leqslant \|A\| + \|B\| + \|R_1(A, B)\| + \|R_1(C)\| \\ &< \|A\| + \|B\| + (\|A\| + \|B\|)^2 + \|C\|^2 \\ &\leqslant 2(\|A\| + \|B\|) + \frac{1}{2}\|C\|^2,\end{aligned}$$

since $\|A\|, \|B\|, \|C\| \leqslant 1/2$. Finally this gives

$$\|C\| \leqslant 4(\|A\| + \|B\|).$$

Equation (4.7) also gives

$$\begin{aligned}\|C - A - B\| &\leqslant \|R_1(A, B)\| + \|R_1(C)\| \\ &\leqslant (\|A\| + \|B\|)^2 + (4(\|A\| + \|B\|))^2,\end{aligned}$$

giving

$$\|C - A - B\| = 17(\|A\| + \|B\|)^2. \tag{4.8}$$

□

Now we will refine these estimates further. Write

$$\exp(C) = I + C + \frac{1}{2}C^2 + R_2(C)$$

where

$$R_2(C) = \sum_{k \geqslant 3} \frac{1}{k!} C^k$$

which satisfies the estimate

$$\|R_2(C)\| \leqslant \frac{1}{3}\|C\|^3$$

since $\|C\| \leqslant 1$. With the aid of Equation (4.5) we obtain

$$\begin{aligned}\exp(C) &= I + A + B + \frac{1}{2}[A, B] + S + \frac{1}{2}C^2 + R_2(C) \\ &= I + A + B + \frac{1}{2}[A, B] + \frac{1}{2}(A + B)^2 + T \\ &= I + A + B + \frac{1}{2}(A^2 + 2AB + B^2) + T, \qquad (4.9)\end{aligned}$$

where

$$T = S + \frac{1}{2}(C^2 - (A + B)^2) + R_2(C). \tag{4.10}$$

Also,

$$\exp(A)\exp(B) = I + A + B + \frac{1}{2}(A^2 + 2AB + B^2) + R_2(A, B) \tag{4.11}$$

where

$$R_2(A, B) = \sum_{k \geqslant 3} \frac{1}{k!} \left(\sum_{r=0}^{k} \binom{k}{r} A^r B^{k-r} \right),$$

which satisfies

$$\|R_2(A, B)\| \leqslant \frac{1}{3}(\|A\| + \|B\|)^3$$

since $\|A\| + \|B\| \leqslant 1$.

Comparing Equations (4.9) and (4.11) and using (4.4) we see that

$$S = R_2(A, B) + \frac{1}{2}((A + B)^2 - C^2) - R_2(C).$$

Taking norms we have

$$\begin{aligned}
\|S\| &\leqslant \|R_2(A, B)\| + \frac{1}{2}\|(A + B)(A + B - C) - (A + B - C)C\| + \|R_2(C)\| \\
&\leqslant \frac{1}{3}(\|A\| + \|B\|)^3 + \frac{1}{2}(\|A\| + \|B\| + \|C\|)\|A + B - C\| + \frac{1}{3}\|C\|^3 \\
&\leqslant \frac{1}{3}(\|A\| + \|B\|)^3 + \frac{5}{2}(\|A\| + \|B\|) \cdot 17(\|A\| + \|B\|)^2 + \frac{1}{3}(4\|A\| + \|B\|)^3 \\
&\leqslant 65(\|A\| + \|B\|)^3,
\end{aligned}$$

yielding the estimate

$$\|S\| \leqslant 65(\|A\| + \|B\|)^3. \tag{4.12}$$

THEOREM 4.26. *If $U, V \in \mathrm{M}_n(\mathbb{R})$, then the following identities are satisfied.*

[Trotter Product Formula]

$$\exp(U + V) = \lim_{r\to\infty} (\exp((1/r)U) \exp((1/r)V))^r .$$

[Commutator Formula] $\exp([U, V]) = \lim_{r\to\infty} (\exp((1/r)U) \exp((1/r)V) \exp(-(1/r)U) \exp(-(1/r)V))^{r^2}$.

PROOF. For large r we may take $A = \dfrac{1}{r}U$ and $B = \dfrac{1}{r}V$ and apply Equation (4.5) to give

$$\exp((1/r)U) \exp((1/r)V) = \exp(C_r)$$

with

$$\|C_r - (1/r)(U + V)\| \leqslant \frac{17(\|U\| + \|V\|)^2}{r^2}.$$

As $r \to \infty$,

$$\|rC_r - (U + V)\| = \frac{17(\|U\| + \|V\|)^2}{r} \to 0,$$

hence $rC_r \to (U + V)$. Since $\exp(rC_r) = \exp(C_r)^r$, the Trotter Formula follows by continuity of exp.

We also have

$$C_r = \frac{1}{r}(U + V) + \frac{1}{2r^2}[U, V] + S_r$$

where

$$\|S_r\| \leqslant 65\frac{(\|U\| + \|V\|)^3}{r^3}.$$

Similarly, replacing U, V by $-U, -V$ we have

$$\exp((-1/r)U)) \exp((-1/r)V)) = \exp(C'_r),$$

where

$$C'_r = \frac{1}{r}(U + V) + \frac{1}{2r^2}[U, V] + S'_r$$

and

$$\|S'_r\| \leqslant 65\frac{(\|U\| + \|V\|)^3}{r^3}.$$

Combining these we obtain

$$\begin{aligned}
\exp((1/r)U)) \exp((1/r)V)) \exp((-1/r)U)) \exp((-1/r)V)) &= \exp(C_r) \exp(C'_r) \\
&= \exp(E_r),
\end{aligned}$$

where

$$
\begin{aligned}
E_r &= C_r + C'_r + \frac{1}{2}[C_r, C'_r] + T_r \\
&= \frac{1}{r^2} + \frac{1}{2}[C_r, C'_r] + S_r + S'_r + T_r.
\end{aligned} \tag{4.13}
$$

Here T_r is defined from Equation (4.5) by setting $C_r = A$, $C'_r = B$ and $T_r = S$.

A tedious computation now shows that

$$
\begin{aligned}
[C_r, C'_r] &= \left[\frac{1}{r}(U+V) + \frac{1}{2r^2}[U,V] + S_r, \frac{-1}{r}(U+V) + \frac{1}{2r^2}[U,V] + S'_r\right] \\
&= \frac{1}{r^3}[U+V,[U,V]] + \frac{1}{r}[U+V, S_r + S'_r] + \frac{1}{2r^2}[[U,V], S'_r - S_r] + [S_r, S'_r].
\end{aligned}
$$

By Proposition 4.25, all four of these terms has norm bounded by an expression of the form (constant)$/r^3$, so the same is true of $[C_r, C'_r]$. Proposition 4.25 also implies that S_r, S'_r, T_r have similarly bounded norms. Setting

$$Q_r = r^2 E_r - [U,V],$$

we obtain

$$\|Q_r\| = \left\|E_r - \frac{1}{r^2}[U,V]\right\| \leqslant \frac{\text{(constant)}}{r^3} \to 0$$

as $r \to \infty$, so

$$\exp(E_r)^{r^2} = \exp([U,V] + Q_r) \to \exp([U,V]).$$

The Commutator Formula now follows using continuity of exp. $\square$

Another proof of Lemma 2.20. As an application of the Trotter formula, we will reprove the formula of Lemma 2.20:

$$\det\exp(A) = \exp(\operatorname{tr} A).$$

The case $n = 1$ is immediate, so assume that $n > 1$.

If $U, V \in \mathrm{M}_n(\mathbb{C})$ then by the Trotter formula together with the fact that det is continuous and multiplicative,

$$
\begin{aligned}
\det\exp(U+V) &= \det\left(\lim_{r\to\infty}\left(\exp((1/r)U)\exp((1/r)V)\right)^r\right) \\
&= \lim_{r\to\infty}\det\left(\exp((1/r)U)\exp((1/r)V)\right)^r \\
&= \lim_{r\to\infty}\det\exp((1/r)U)^r\det\exp((1/r)V)^r \\
&= \lim_{r\to\infty}\det\left(\exp((1/r)U)^r\right)\det\left(\exp((1/r)V)^r\right) \\
&= \lim_{r\to\infty}\det\exp(U)\det\exp(V) \\
&= \det\exp(U)\det\exp(V).
\end{aligned}
$$

More generally, given $U_1 \ldots, U_k \in \mathrm{M}_n(\mathbb{C})$ we have

$$\det\exp(U_1 + \cdots + U_k) = \det\exp(U_1)\cdots\det\exp(U_k). \tag{4.14}$$

So if $A = A_1 + \cdots + A_k$ where the A_j satisfy

$$\det\exp(A_j) = \exp(\operatorname{tr} A_j) \quad (j = 1, \ldots, k),$$

we have

$$\begin{aligned}\det\exp(A) &= \det\exp(A_1 + \cdots + A_k)\\ &= \exp(\operatorname{tr} A_1)\cdots\exp(\operatorname{tr} A_k)\\ &= \exp(\operatorname{tr} A_1 + \cdots + \operatorname{tr} A_k)\\ &= \exp(\operatorname{tr} A).\end{aligned}$$

So it suffices to show that every matrix A has this form.

Recall that $A = [a_{ij}]$ can be expressed as

$$A = \sum_{\substack{1\leqslant r\leqslant n\\ 1\leqslant s\leqslant n}} a_{rs}E^{rs},$$

where E^{rs} is the matrix having 1 in the (r,s) place and 0 everywhere else, i.e.,

$${E^{rs}}_{ij} = \delta_{ir}\delta_{js}.$$

For $z \in \mathbb{C}$,

$$\begin{aligned}\det\exp(zE^{rs}) &= \det\left(\sum_{k\geqslant 0}\frac{1}{k!}(z^k(E^{rs})^k\right)\\ &= \begin{cases}\det((e^z-1)E^{rr}+I_n) & \text{if } r=s,\\ \det I_n & \text{if } r\neq s,\end{cases}\\ &= \begin{cases}e^z & \text{if } r=s,\\ 1 & \text{if } r\neq s.\end{cases}\end{aligned}$$

On the other hand,

$$\operatorname{tr} zE^{rs} = \begin{cases}z & \text{if } r=s,\\ 0 & \text{if } r\neq s.\end{cases}$$

Thus

$$\exp(\operatorname{tr} zE^{rs}) = \begin{cases}e^z & \text{if } r=s,\\ 1 & \text{if } r\neq s,\end{cases}$$

giving the desired equation

$$\det\exp(zE^{rs}) = \exp(\operatorname{tr} zE^{rs}).$$

□

6. Matrix groups are Lie groups

Our goal in this section is to prove Theorem 4.23. Let $G \leqslant \mathrm{GL}_n(\mathbb{R})$ be a matrix subgroup. Recall that the Lie algebra $\mathfrak{g} = \mathrm{T}_I\, G$ is an $\mathbb{R}$-Lie subalgebra of $\mathfrak{gl}_n(\mathbb{R}) = \mathrm{M}_n(\mathbb{R})$. Let

$$\widetilde{\mathfrak{g}} = \{A \in \mathrm{M}_n(\mathbb{R}) : \forall t \in \mathbb{R},\ \exp(tA) \in G\}.$$

THEOREM 4.27. *$\widetilde{\mathfrak{g}}$ is an $\mathbb{R}$-Lie subalgebra of $\mathrm{M}_n(\mathbb{R})$.*

PROOF. By definition $\widetilde{\mathfrak{g}}$ is closed under multiplication by real scalars. If $U, V \in \widetilde{\mathfrak{g}}$ and $r \geqslant 1$, then the following are in G:

$$\begin{aligned}
&(\exp((1/r)U)\exp((1/r)V)),\\
&(\exp((1/r)U)\exp((1/r)V))^r,\\
&(\exp((1/r)U)\exp((1/r)V)\exp(-(1/r)U)\exp(-(1/r)V))^{r^2},\\
&(\exp((1/r)U)\exp((1/r)V)\exp(-(1/r)U)\exp(-(1/r)V))^{r^2}.
\end{aligned}$$

By Theorem 4.26, for $t \in \mathbb{R}$,

$$\begin{aligned}
\exp(tU + tV) &= \lim_{r\to\infty} (\exp((1/r)tU)\exp((1/r)tV))^r,\\
\exp(t[U,V]) &= \exp([tU,V])\\
&= \lim_{r\to\infty} (\exp((1/r)tU)\exp((1/r)V)\exp(-(1/r)tU)\exp(-(1/r)V))^{r^2},
\end{aligned}$$

and as these are both limits of elements of the closed subgroup G they are also in G.

Hence $\widetilde{\mathfrak{g}}$ is a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{R}) = \mathrm{M}_n(\mathbb{R})$. $\square$

PROPOSITION 4.28. *For a matrix subgroup $G \leqslant \mathrm{GL}_n(\mathbb{R})$, $\widetilde{\mathfrak{g}}$ is an $\mathbb{R}$-Lie subalgebra of $\mathfrak{g}$.*

PROOF. Let $U \in \widetilde{\mathfrak{g}}$. The curve

$$\gamma\colon \mathbb{R} \longrightarrow G; \quad \gamma(t) = \exp(tU),$$

has $\gamma(0) = I$ and $\gamma'(0) = U$. Hence $U \in \mathfrak{g}$. $\square$

REMARK 4.29. Eventually we will see that $\widetilde{\mathfrak{g}} = \mathfrak{g}$.

Later we will require the following technical result.

LEMMA 4.30. *Let $\{A_n \in \exp^{-1} G\}_{n\geqslant 1}$ and $\{s_n \in \mathbb{R}\}_{n\geqslant 1}$ be sequences for which $\|A_n\| \to 0$ and $s_n A_n \to A \in \mathrm{M}_n(\mathbb{R})$ as $n \to \infty$. Then $A \in \widetilde{\mathfrak{g}}$.*

PROOF. Let $t \in \mathbb{R}$. For each n, choose an integer $m_n \in \mathbb{Z}$ so that $|ts_n - m_n| \leqslant 1$. Then

$$\begin{aligned}
\|m_n A_n - tA\| &\leqslant \|(m_n - ts_n)A_n\| + \|ts_n A_n - tA\|\\
&= |m_n - ts_n|\,\|A_n\| + \|ts_n A_n - tA\|\\
&\leqslant \|A_n\| + \|ts_n A_n - tA\| \to 0
\end{aligned}$$

as $n \to \infty$, showing that $m_n A_n \to tA$. Since

$$\exp(m_n A_n) = \exp(A_n)^{m_n} \in G,$$

and G is closed in $\mathrm{GL}_n(\mathbb{R})$, we have

$$\exp(tA) = \lim_{n\to\infty} \exp(m_n A_n) \in G.$$

Thus every real scalar multiple tA is in $\exp^{-1} G$, showing that $A \in \widetilde{\mathfrak{g}}$. $\square$

Choose a complementary $\mathbb{R}$-subspace $\mathfrak{w}$ to $\widetilde{\mathfrak{g}}$ in $\mathfrak{gl}_n(\mathbb{R}) = \mathrm{M}_n(\mathbb{R})$, i.e., any vector subspace such that

$$\widetilde{\mathfrak{g}} + \mathfrak{w} = \mathrm{M}_n(\mathbb{R}),$$
$$\dim_{\mathbb{R}} \widetilde{\mathfrak{g}} + \dim_{\mathbb{R}} \mathfrak{w} = \dim_{\mathbb{R}} \mathrm{M}_n(\mathbb{R}) = n^2.$$

The second condition is equivalent to $\widetilde{\mathfrak{g}} \cap \mathfrak{w} = 0$. This gives a direct sum decomposition of $\mathrm{M}_n(\mathbb{R})$, so every element $X \in \mathrm{M}_n(\mathbb{R})$ has a unique expression of the form

$$X = U + V, \quad (U \in \widetilde{\mathfrak{g}},\ V \in \mathfrak{w}).$$

Consider the map

$$\Phi\colon\ \mathrm{M}_n(\mathbb{R}) \longrightarrow \mathrm{GL}_n(\mathbb{R}); \qquad \Phi(U+V) = \exp(U)\exp(V) \quad (U \in \widetilde{\mathfrak{g}},\ V \in \mathfrak{w}).$$

Φ is a smooth function which maps 0 to I. Notice that the factor $\exp(U)$ is in G. Consider the derivative at O,

$$\mathrm{d}\,\Phi_O\colon\ \mathrm{M}_n(\mathbb{R}) \longrightarrow \mathfrak{gl}_n(\mathbb{R}) = \mathrm{M}_n(\mathbb{R}).$$

To determine $\mathrm{d}\,\Phi_O(A+B)$, where $A \in \widetilde{\mathfrak{g}}$ and $B \in \mathfrak{w}$, we differentiate the curve $t \mapsto \Phi(t(A+B))$ at $t = 0$. Assuming that A, B are small enough and following the notation of Equations (4.4),(4.5) for small $t \in \mathbb{R}$, there is a unique $C(t)$ depending on t for which

$$\Phi(t(A+B)) = \exp(C(t)).$$

Proposition 4.25 gives

$$\|(C(t) - tA - tB) - \frac{t^2}{2}[A,B]\| \leqslant 65|t|^3(\|A\| + \|B\|)^3.$$

From this we obtain

$$\begin{aligned}\|(C(t) - tA - tB\| &\leqslant \frac{t^2}{2}\|[A,B]\| + 65|t|^3(\|A\| + \|B\|)^3 \\ &= \frac{t^2}{2}\left(\|[A,B]\| + 130|t|(\|A\| + \|B\|)^3\right)\end{aligned}$$

and so

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\Phi(t(A+B))_{|t=0} = \frac{\mathrm{d}}{\mathrm{d}\,t}\exp(C(t))_{|t=0} = A + B.$$

By linearity of the derivative, for small A, B,

$$\mathrm{d}\,\Phi_O(A+B) = A + B,$$

so $\mathrm{d}\,\Phi_O$ is the identity function on $\mathrm{M}_n(\mathbb{R})$. By the Inverse Function Theorem 4.11, Φ is a diffeomorphism onto its image when restricted to a small open neighbourhood of O, and we might as well take this to be an open disc $\mathrm{N}_{\mathrm{M}_n(\Bbbk)}(O;\delta)$ for some $\delta > 0$; hence the restriction of Φ to

$$\Phi_1\colon \mathrm{N}_{\mathrm{M}_n(\Bbbk)}(O;\delta) \longrightarrow \Phi\mathrm{N}_{\mathrm{M}_n(\Bbbk)}(O;\delta)$$

is a diffeomorphism.

Now we must show that Φ maps some open subset (which we could assume to be an open disc) of $\mathrm{N}_{\mathrm{M}_n(\Bbbk)}(O;\delta)\cap\widetilde{\mathfrak{g}}$ containing O onto an open neighbourhood of I in G. Suppose not; then there is a sequence of elements $U_n \in G$ with $U_n \to I$ as $n \to \infty$ but $U_n \notin \Phi\widetilde{\mathfrak{g}}$. For large enough n, $U_n \in \Phi\mathrm{N}_{\mathrm{M}_n(\Bbbk)}(O;\delta)$, hence there are unique elements $A_n \in \widetilde{\mathfrak{g}}$ and $B_n \in \mathfrak{w}$ with $\Phi(A_n + B_n) = U_n$; notice that $B_n \neq O$ since otherwise $U_n \in \Phi\widetilde{\mathfrak{g}}$. As Φ_1 is a diffeomorphism, $A_n + B_n \to O$ and this implies that $A_n \to O$ and $B_n \to O$. By definition of Φ,

$$\exp(B_n) = \exp(A_n)^{-1}U_n \in G,$$

hence $B_n \in \exp^{-1} G$. Consider the elements $\overline{B}_n = (1/\|B_n\|)B_n$ of unit norm. Each $\overline{B}_n$ is in the unit sphere in $\mathrm{M}_n(\mathbb{R})$, which is compact hence there is a convergent subsequence of $\{\overline{B}_n\}$. By renumbering this subsequence, we can assume that $\overline{B}_n \to B$ as $n \to \infty$, where $\|B\| = 1$. Applying Lemma 4.30 to the sequences $\{B_n\}$ and $\{1/\|B_n\|\}$ we find that $B \in \widetilde{\mathfrak{g}}$. But each B_n (and hence $\overline{B}_n$) is in $\mathfrak{w}$, so B must be too. Thus $B \in \widetilde{\mathfrak{g}} \cap \mathfrak{w}$, contradicting the fact $B \neq O$.

So there must be an open disc

$$\mathrm{N}_{\widetilde{\mathfrak{g}}}(O;\delta_1) = \mathrm{N}_{\mathrm{M}_n(\mathbb{R})}(O;\delta_1) \cap \widetilde{\mathfrak{g}}$$

which is mapped by Φ onto an open neighbourhood of I in G. So the restriction of Φ to this open disc is a local diffeomorphism at O. The inverse map gives a chart for $\mathrm{GL}_n(\mathbb{R})$ at I and moreover $\mathrm{N}_{\widetilde{\mathfrak{g}}}(O;\delta_1)$ is then a submanifold of $\mathrm{N}_{\mathrm{M}_n(\mathbb{R})}(O;\delta_1)$.

We can use left translation to move this chart to a new chart at any other point $U \in G$, by considering $\mathrm{L}_U \circ \Phi$. We leave the details as an exercise.

So we have shown that $G \leqslant \mathrm{GL}_n(\mathbb{R})$ is a Lie subgroup, proving Theorem 4.23. Notice that the dimension of G as a manifold is $\dim_{\mathbb{R}} \widetilde{\mathfrak{g}}$. By Proposition 4.28, $\widetilde{\mathfrak{g}} \subseteq \mathfrak{g}$ so $\dim_{\mathbb{R}} \widetilde{\mathfrak{g}} \leqslant \dim_{\mathbb{R}} \mathfrak{g}$. But by Theorem 4.22, these dimensions are in fact equal, hence $\widetilde{\mathfrak{g}} = \mathfrak{g}$.

We have established a fundamental result that we now reformulate. The proof of the second part is similar to our proof of the first with minor adjustments required for the general case.

Theorem 4.31. *A subgroup of* $\mathrm{GL}_n(\mathbb{R})$ *is a closed Lie subgroup if and only if it is a matrix subgroup, i.e., a closed subgroup.*

More generally, a subgroup of an arbitrary Lie group G is a closed Lie subgroup if and only if it is a closed subgroup.

7. Not all Lie groups are matrix groups

For completeness we describe the simplest example of a Lie group which is not a matrix group. In fact there are finitely many related examples of such *Heisenberg groups* Heis_n and the example we will discuss Heis_3 is particularly important in Quantum Physics.

For $n \geqslant 3$, the *Heisenberg group* Heis_n is defined as follows. Recall the group of $n \times n$ real unipotent matrices $\mathrm{SUT}_n(\mathbb{R})$, whose elements have the form

$$\begin{bmatrix} 1 & a_{12} & \cdots & \cdots & \cdots & a_{1n} \\ 0 & 1 & a_{21} & \ddots & \ddots & a_{2n} \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 & a_{n-2\,n-1} & \vdots \\ \vdots & \vdots & \ddots & 0 & 1 & a_{n-1\,n} \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix},$$

with $a_{ij} \in \mathbb{R}$. The Lie algebra $\mathfrak{sut}_n(\mathbb{R})$ of $\mathrm{SUT}_n(\mathbb{R})$ consists of the matrices of the form

$$\begin{bmatrix} 0 & t_{12} & \cdots & \cdots & \cdots & t_{1n} \\ 0 & 0 & t_{21} & \ddots & \ddots & t_{2n} \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & t_{n-2\,n-1} & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 & t_{n-1\,n} \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

with $t_{ij} \in \mathbb{R}$. SUT_n is a matrix subgroup of $\mathrm{GL}_n(\mathbb{R})$ with $\dim \mathrm{SUT}_n = \binom{n}{2}$. It is a nice algebraic exercise to show that the following hold in general.

Proposition 4.32. *For $n \geqslant 3$, the centre* $\mathrm{C}(\mathrm{SUT}_n)$ *of* SUT_n *consists of all the matrices* $[a_{ij}] \in \mathrm{Heis}_n$ *with $a_{ij} = 0$ except when $i = 1$ and $j = n$. Furthermore,* $\mathrm{C}(\mathrm{SUT}_n)$ *is contained in the commutator subgroup of* SUT_n.

Notice that there is an isomorphism of Lie groups $\mathbb{R} \cong \mathrm{C}(\mathrm{SUT}_n)$. Under this isomorphism, the subgroup of integers $\mathbb{Z} \subseteq \mathbb{R}$ corresponds to the matrices with $a_{1\,n} \in \mathbb{Z}$ and these form a discrete normal (in fact central) subgroup $\mathrm{Z}_n \triangleleft \mathrm{SUT}_n$. We can form the quotient group

$$\mathrm{Heis}_n = \mathrm{SUT}_n / \mathrm{Z}_n.$$

This has the quotient space topology and as Z_n is a discrete subgroup, the quotient map $q\colon \mathrm{SUT}_n \longrightarrow \mathrm{Heis}_n$ is a local homeomorphism. This can be used to show that Heis_n is also a Lie group since charts for SUT_n defined on small open sets will give rise to charts for Heis_n. The Lie algebra of Heis_n is the same as that of SUT_n, i.e., $\mathfrak{heis}_n = \mathfrak{sut}_n$.

PROPOSITION 4.33. *For $n \geqslant 3$, the centre $\mathrm{C}(\mathrm{Heis}_n)$ of Heis_n consists of the image under q of $\mathrm{C}(\mathrm{SUT}_n)$. Furthermore, $\mathrm{C}(\mathrm{Heis}_n)$ is contained in the commutator subgroup of Heis_n.*

Notice that $\mathrm{C}(\mathrm{Heis}_n) = \mathrm{C}(\mathrm{SUT}_n)/\mathrm{Z}_n$ is isomorphic to the circle group

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$$

with the correspondence coming from the map

$$\mathbb{R} \longrightarrow \mathbb{T}; \quad t \longmapsto e^{2\pi i t}.$$

When $n = 3$, there is a surjective Lie homomorphism

$$p\colon \mathrm{SUT}_3 \longrightarrow \mathbb{R}^2; \quad \begin{bmatrix} 1 & x & t \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \end{bmatrix}$$

whose kernel is $\ker p = \mathrm{C}(\mathrm{SUT}_3)$. Since $\mathrm{Z}_3 \leqslant \ker p$, there is an induced surjective Lie homomorphism $\overline{p}\colon \mathrm{Heis}_3 \longrightarrow \mathbb{R}^2$ for which $\overline{p} \circ q = p$. In this case the isomorphism $\mathrm{C}(\mathrm{Heis}_n) \cong \mathbb{T}$ is given by

$$\begin{bmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathrm{Z}_3 \longleftrightarrow e^{2\pi i t}$$

From now on we will write $[x, y, e^{2\pi i t}]$ for the element

$$\begin{bmatrix} 1 & x & t \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \mathrm{Z}_3 \in \mathrm{Heis}_3 .$$

Thus a general element of Heis_3 has the form $[x, y, z]$ with $x, y \in \mathbb{R}$ and $z \in \mathbb{T}$. The identity element is $1 = [0, 0, 1]$. The element

$$\begin{bmatrix} 1 & x & t \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}$$

of the Lie algebra $\mathfrak{heis}_3$ will be denoted (x, y, t).

PROPOSITION 4.34. *Multiplication, inverses and commutators in Heis_3 are given by*

$$[x_1, y_1, z_1][x_2, y_2, z_2] = [x_1 + x_2, y_1 + y_2, z_1 z_2 e^{2\pi i x_1 y_2}],$$

$$[x, y, z]^{-1} = [-x, -y, z^{-1} e^{2\pi i x y}]$$

$$[x_1, y_1, z_1][x_2, y_2, z_2][x_1, y_1, z_1]^{-1}[x_2, y_2, z_2]^{-1} = [0, 0, e^{2\pi i (x_1 y_2 - y_1 x_2)}].$$

The Lie bracket in $\mathfrak{heis}_3$ is given by

$$[(x_1, y_1, t_1), (x_2, y_2, t_2)] = (0, 0, x_1 y_2 - y_1 x_2).$$

The Lie algebra $\mathfrak{heis}_3$ is often called a *Heisenberg (Lie) algebra* and occurs throughout Quantum Physics. It is essentially the same as the Lie algebra of operators on differentiable functions $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ spanned by $\mathbf{1}, \mathbf{q}$ given by

$$\mathbf{1}f(x) = f(x), \quad \mathbf{p}f(x) = \frac{\mathrm{d}\, f(x)}{\mathrm{d}\, x}, \quad \mathbf{q}f(x) = xf(x).$$

The non-trivial commutator involving these three operators is given by the *canonical commutation relation*

$$[\mathbf{p}, \mathbf{q}] = \mathbf{pq} - \mathbf{qp} = \mathbf{1}.$$

In $\mathfrak{heis}_3$ he elements $(1,0,0), (1,0,0), (0,0,1)$ a basis with the only non-trivial commutator

$$[(1,0,0), (1,0,0)] = (0,0,1).$$

THEOREM 4.35. *There are no continuous homomorphisms* $\varphi\colon \mathrm{Heis}_3 \longrightarrow \mathrm{GL}_n(\mathbb{C})$ *with trivial kernel* $\ker\varphi = 1$.

PROOF. Suppose that $\varphi\colon \mathrm{Heis}_3 \longrightarrow \mathrm{GL}_n(\mathbb{C})$ is a continuous homomorphism with trivial kernel and suppose that n is minimal with this property. For each $g \in \mathrm{Heis}_3$, the matrix $\varphi(g)$ acts on vectors in $\mathbb{C}^n$.

We will identify $\mathrm{C}(\mathrm{Heis}_3)$ with the circle $\mathbb{T}$ as above. Then $\mathbb{T}$ has a *topological generator* z_0; this is an element whose powers form a cyclic subgroup $\langle z_0 \rangle \leqslant \mathbb{T}$ whose closure is $\mathbb{T}$. Proposition 7.7 will provide a more general version of this phenomenon. For now we point out that for any irrational number $r \in \mathbb{R}$, the following is true: for any real number $s \in \mathbb{R}$ and any $\varepsilon > 0$, there are integers $p, q \in \mathbb{Z}$ such that

$$|s - pr - q| < \varepsilon.$$

This implies that $e^{2\pi i r}$ is a topological generator of $\mathbb{T}$ since its powers are dense.

Let λ be an eigenvalue for the matrix $\varphi(z_0)$, with eigenvector $\mathbf{v}$. If necessary replacing z_0 with z_0^{-1}, we may assume that $\lambda \geqslant 1$. If $\|\lambda\| > 1$, then

$$\varphi(z_0^k)\mathbf{v} = \varphi(z_0)^k\mathbf{v} = \lambda^k\mathbf{v}$$

and so

$$\|\varphi(z_0^k)\| \geqslant \|\lambda\|^k.$$

Thus $\|\varphi(z_0^k)\| \to \infty$ as $k \to \infty$, implying that $\varphi\mathbb{T}$ is unbounded. But φ is continuous and $\mathbb{T}$ is compact hence $\varphi\mathbb{T}$ is bounded. So in fact $\|\lambda\| = 1$.

Since φ is a homomorphism and $z_0 \in \mathrm{C}(\mathrm{Heis}_3)$, for any $g \in \mathrm{Heis}_3$ we have

$$\varphi(z_0)\varphi(g)\mathbf{v} = \varphi(z_0 g)\mathbf{v} = \varphi(g z_0)\mathbf{v} = \varphi(g)\varphi(z_0)\mathbf{v} = \lambda\varphi(g)\mathbf{v},$$

which shows that $\varphi(g)$ is another eigenvector of $\varphi(z_0)$ for the eigenvalue λ. If we set

$$V_\lambda = \{\mathbf{v} \in \mathbb{C}^n : \exists k \geqslant 1 \text{ s.t. } (\varphi(z_0) - \lambda I_n)^k\mathbf{v} = \mathbf{0}\},$$

then $V_\lambda \subseteq \mathbb{C}^n$ is a vector subspace which is also closed under the actions of all the matrices $\varphi(g)$ with $g \in \mathrm{Heis}_3$. Choose $k_0 \geqslant 1$ to be the largest number for which there is a vector $\mathbf{v}_0 \in V_\lambda$ satisfying

$$(\varphi(z_0) - \lambda I_n)^{k_0}\mathbf{v}_0 = \mathbf{0}, \quad (\varphi(z_0) - \lambda I_n)^{k_0 - 1}\mathbf{v}_0 \neq \mathbf{0}.$$

If $k_0 > 1$, there are vectors $\mathbf{u}, \mathbf{v} \in V_\lambda$ for which

$$\varphi(z_0)\mathbf{u} = \lambda\mathbf{u} + \mathbf{v}, \quad \varphi(z_0)\mathbf{v} = \lambda\mathbf{v}.$$

Then

$$\varphi(z_0^k)\mathbf{u} = \varphi(z_0)^k\mathbf{u} = \lambda^k\mathbf{u} + k\lambda^{k-1}\mathbf{v}$$

and since $|\lambda| = 1$,

$$\|\varphi(z_0^k)\| = \|\varphi(z_0)^k\| \geqslant |\lambda\mathbf{u} + k\mathbf{v}| \to \infty$$

as $k \to \infty$. This also contradicts the fact that $\varphi\mathbb{T}$ is bounded. So $k_0 = 1$ and V_λ is just the eigenspace for the eigenvalue λ. This argument actually proves the following important general result, which in particular applies to finite groups viewed as zero-dimensional compact Lie groups.

Proposition 4.36. *Let G be a compact Lie group and $\rho\colon G \longrightarrow \mathrm{GL}_n(\mathbb{C})$ a continuous homomorphism. Then for any $g \in G$, $\rho(g)$ is diagonalizable.*

On choosing a basis for V_λ, we obtain a continuous homomorphism $\theta\colon \mathrm{Heis}_3 \longrightarrow \mathrm{GL}_d(\mathbb{C})$ for which $\theta(z_0) = \lambda I_d$. By continuity, every element of $\mathbb{T}$ also has the form (scalar)I_d. By minimality of n, we must have $d = n$ and we can assume $\varphi(z_0) = \lambda I_n$.

By the equation for commutators in Proposition 4.34, every element $z \in \mathbb{T} \leqslant \mathrm{Heis}_3$ is a commutator $z = ghg^{-1}h^{-1}$ in Heis_3, hence

$$\det\varphi(z) = \varphi(ghg^{-1}h^{-1}) = 1,$$

since det and φ are homomorphisms. So for every $z \in \mathbb{T}$, $\varphi(z) = \mu(z)I_d$ and $\mu(z)^d = 1$, where the function $\mu\colon \mathbb{T} \longrightarrow \mathbb{C}^\times$ is continuous. But $\mathbb{T}$ is path connected, so $\mu(z) = 1$ for every $z \in \mathbb{T}$. Hence for each $z \in \mathbb{T}$, the only eigenvalue of $\varphi(z)$ is 1. This shows that $\mathbb{T} \leqslant \ker\varphi$, contradicting the assumption that $\ker\varphi$ is trivial. $\square$

A modification of this argument works for each of the Heisenberg groups Heis_n ($n \geqslant 3$), showing that none of them is a matrix group.

CHAPTER 5

Homogeneous spaces

1. Homogeneous spaces as manifolds

Let G be a Lie group of dimension $\dim G = n$ and $H \leqslant G$ a closed subgroup, which is therefore a Lie subgroup of dimension $\dim H = k$. The set of left cosets

$$G/H = \{gH : g \in G\}$$

has an associated quotient map

$$\pi\colon G \longrightarrow G/H; \quad \pi(g) = gH.$$

We give G/H a topology by requiring that a subset $W \subseteq G/H$ is open if and only if $\pi^{-1}W \subseteq G$ is open; this is called the *quotient topology* on G/H.

LEMMA 5.1. *The projection map $\pi\colon G \longrightarrow G/H$ is an open mapping and G/H is a topological space which is separable and Hausdorff.*

PROOF. For $U \subseteq G$,

$$\pi^{-1}(\pi U) = \bigcup_{h\in H} Uh,$$

where

$$Uh = \{uh \in G : u \in U\} \subseteq G.$$

If $U \subseteq G$ is open, then each Uh ($h \in H$) is open, implying that $\pi U \subseteq G$ is also open.

G/H is separable since a countable basis of G is mapped by π to a countable collection of open subsets of G/H that is also a basis.

To see that G/H is Hausdorff, consider the continuous map

$$\theta\colon G \times G \longrightarrow G; \quad \theta(x, y) = x^{-1}y.$$

Then

$$\theta^{-1}H = \{(x, y) \in G \times G : xH = yH\},$$

and this is a closed subset since $H \subseteq G$ is closed. Hence,

$$\{(x, y) \in G \times G : xH = yH\} \subseteq G \times G$$

is open. By definition of the product topology, this means that whenever $x, y \in G$ with $xH \neq yH$, there are open subsets $U, V \subseteq G$ with $x \in U$, $y \in V$, $U \neq V$ and $\pi U \cap \pi V = \emptyset$. Since $\pi U, \pi V \subseteq G/H$ are open, this shows that G/H is Hausdorff. $\square$

The quotient map $\pi\colon G \longrightarrow G/H$ has an important property which characterises it.

PROPOSITION 5.2 (Universal Property of the Quotient Topology). *For any topological space X, a function $f\colon G/H \longrightarrow X$ is continuous if and only if $f \circ \pi\colon G \longrightarrow X$ is continuous.*

We would like to make G/H into a smooth manifold so that $\pi\colon G \longrightarrow G/H$ is smooth. Unfortunately, the construction of an atlas is rather complicated so we merely state a general result then consider some examples where the smooth structure comes from an existing manifold which is diffeomorphic to a quotient.

Theorem 5.3. *G/H can be given the structure of a smooth manifold of dimension*

$$\dim G/H = \dim G - \dim H$$

so that the projection map $\pi\colon G \longrightarrow G/H$ is smooth and at each $g \in G$,

$$\ker(\mathrm{d}\,\pi\colon\ \mathrm{T}_g\, G \longrightarrow \mathrm{T}_{gH}\, G/H) = \mathrm{d}\, \mathrm{L}_g \mathfrak{h}.$$

There is an atlas for G/H consisting of charts of the form $\theta\colon W \longrightarrow \theta W \subseteq \mathbb{R}^{n-k}$ for which there is a diffeomorphism $\Theta\colon W \times H \longrightarrow \pi^{-1}W$ satisfying the conditions

$$\Theta(w, h_1h_2) = \Theta(w, h_1)h_2, \quad \pi(\Theta(w,h)) = w \qquad (w \in W,\ h, h_1, h_2 \in H).$$

$$\begin{tikzcd} W \times H \arrow[rr, "\Theta"] \arrow[dr, "\mathrm{proj}_1"'] & & \pi^{-1}W \arrow[dl, "\pi"] \\ & W & \end{tikzcd}$$

The projection π looks like $\mathrm{proj}_1\colon \pi^{-1}W \longrightarrow W$, the projection onto W, when restricted to $\pi^{-1}W$. For such a chart, the map Θ is said to provide a *local trivialisation of π over W*. An atlas consisting of such charts and local trivialisations $(\theta\colon W \longrightarrow \theta W, \Theta)$ provides a *local trivalisation of π*. This is related to the important notion of a *principal H-bundle* over G/H.

Notice that given such an atlas, an atlas for G can be obtained by taking each pair $(\theta\colon W \longrightarrow \theta W, \Theta)$ and combining the map θ with a chart $\psi\colon U \longrightarrow \psi U \subseteq \mathbb{R}^k$ for H to get a chart

$$(\theta \times \psi) \circ \Theta^{-1}\colon \Theta(W \times U) \longrightarrow \theta W \times \psi U \subseteq \mathbb{R}^{n-k} \times \mathbb{R}^k = \mathbb{R}^n.$$

Such a manifold G/H is called a *homogeneous space* since each left translation map L_g on G gives rise to a diffeomorphism

$$\overline{\mathrm{L}}_g\colon G/H \longrightarrow G/H; \quad \overline{\mathrm{L}}_g(xH) = gxH,$$

for which $\pi \circ \mathrm{L}_g = \overline{\mathrm{L}}_g \circ \pi$.

$$\begin{tikzcd} G \arrow[r, "\mathrm{L}_g"] \arrow[d, "\pi"'] & G \arrow[d, "\pi"'] \\ G/H \arrow[r, "\overline{\mathrm{L}}_g"] & G/H \end{tikzcd}$$

So each point gH has a neighbourhood diffeomorphic under $\overline{\mathrm{L}}_g^{-1}$ to a neighbourhood of $1H$; so locally G/H is unchanged as gH is varied. This is the basic insight in Felix Klein's view of a Geometry which is characterised as a homogeneous space G/H for some group of transformations G and subgroup H.

2. Homogeneous spaces as orbits

Just as in ordinary group theory, group actions have orbits equivalent to sets of cosets G/H, so homogeneous spaces also arise as orbits associated to smooth groups actions of G on a manifolds.

Theorem 5.4. *Suppose that a Lie group G acts smoothly on a manifold M. If the element $x \in M$ has stabilizer $\mathrm{Stab}_G(x) \leqslant G$ and the orbit $\mathrm{Orb}_G(x) \subseteq M$ is a closed submanifold, then the function*

$$f\colon G/\operatorname{Stab}_G(x) \longrightarrow \mathrm{Orb}_G(x); \quad f(g \operatorname{Stab}_G(x)) = gx$$

is a diffeomorphism.

Example 5.5. For $n \geqslant 1$, $\mathrm{O}(n)$ acts smoothly on $\mathbb{R}^n$ by matrix multiplication. For any nonzero vector $\mathbf{v} \in \mathbb{R}^n$, the orbit $\mathrm{Orb}_{\mathrm{O}(n)}(\mathbf{v}) \subseteq \mathbb{R}^n$ is diffeomorphic to $\mathrm{O}(n)/\,\mathrm{O}(n-1)$.

PROOF. First observe that when $\mathbf{v}$ is the standard basis vector $\mathbf{e}_n$, for $A \in \mathrm{O}(n)$, $A\mathbf{e}_n = \mathbf{e}_n$ if and only if $\mathbf{e}_n$ is the last column of A, while all the other columns of A are orthogonal to $\mathbf{e}_n$. Since the columns of A must be an orthonormal set of vectors, this means that each of the first $(n-1)$ columns of A has the form

$$\begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \\ 0 \end{bmatrix}$$

where the matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} \\ a_{21} & a_{22} & \cdots & a_{2\,n-1} \\ \vdots & & \ddots & \\ a_{n-1\,1} & a_{n-1\,2} & \cdots & a_{n-1\,n-1} \end{bmatrix}$$

is orthogonal and hence in $\mathrm{O}(n-1)$. We identify $\mathrm{O}(n-1)$ with the subset of $\mathrm{O}(n)$ consisting of matrices of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & 0 \\ a_{21} & a_{22} & \cdots & a_{2\,n-1} & 0 \\ \vdots & & \ddots & & 0 \\ a_{n-1\,1} & a_{n-1\,2} & \cdots & a_{n-1\,n-1} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

and then have $\mathrm{Stab}_{\mathrm{O}(n)}(\mathbf{e}_n) = \mathrm{O}(n-1)$. The orbit of $\mathbf{e}_n$ is the whole unit sphere $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ since given a unit vector $\mathbf{u}$ we can extend it to an orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}, \mathbf{u}_n = \mathbf{u}$ which form the columns of an orthogonal matrix $U \in \mathrm{O}(n)$ for which $U\mathbf{e}_n = \mathbf{u}$. So we have a diffeomorphism

$$\mathrm{O}(n)/\,\mathrm{Stab}_{\mathrm{O}(n)}(\mathbf{e}_n) = \mathrm{O}(n)/\,\mathrm{O}(n-1) \longrightarrow \mathrm{Orb}_{\mathrm{O}(n)}(\mathbf{e}_n) = \mathbb{S}^{n-1}.$$

Now for a general nonzero vector $\mathbf{v}$ notice that $\mathrm{Stab}_{\mathrm{O}(n)}(\mathbf{v}) = \mathrm{Stab}_{\mathrm{O}(n)}(\hat{\mathbf{v}})$ where $\hat{\mathbf{v}} = (1/|\mathbf{v}|)\mathbf{v}$ and

$$\mathrm{Orb}_{\mathrm{O}(n)}(\mathbf{v}) = \mathbb{S}^{n-1}(|\mathbf{v}|),$$

the sphere of radius $|\mathbf{v}|$. If we choose any $P \in \mathrm{O}(n)$ with $\hat{\mathbf{v}} = P\mathbf{e}_n$, we have

$$\mathrm{Stab}_{\mathrm{O}(n)}(\mathbf{v}) = P\,\mathrm{Stab}_{\mathrm{O}(n)}(\mathbf{e}_n)P^{-1}$$

and so there is a diffeomorphism

$$\mathrm{Orb}_{\mathrm{O}(n)}(\mathbf{v}) \longrightarrow \mathrm{O}(n)/P\,\mathrm{O}(n-1)P^{-1} \xrightarrow{\chi_{P^{-1}}} \mathrm{O}(n)/\,\mathrm{O}(n-1).$$

□

A similar result holds for $\mathrm{SO}(n)$ and the homogeneous space $\mathrm{SO}(n)/\,\mathrm{SO}(n-1)$. For the unitary and special unitary groups we can obtain the homogeneous spaces $\mathrm{U}(n)/\,\mathrm{U}(n-1)$ and $\mathrm{SU}(n)/\,\mathrm{SU}(n-1)$ as orbits of non-zero vectors in $\mathbb{C}^n$ on which these groups act by matrix multiplication; these are all diffeomorphic to $\mathbb{S}^{2n-1}$. The action of the quaternionic symplectic group $\mathrm{Sp}(n)$ on $\mathbb{H}^n$ leads to orbits of non-zero vectors diffeomorphic to $\mathrm{Sp}(n)/\,\mathrm{Sp}(n-1)$ and $\mathbb{S}^{4n-1}$.

3. Projective spaces

More exotic orbit spaces are obtained as follows. Let $\Bbbk = \mathbb{R}, \mathbb{C}, \mathbb{H}$ and set $d = \dim_{\mathbb{R}} \Bbbk$. Consider $\Bbbk^{n+1}$ as a right $\Bbbk$-vector space. Then there is an action of the group of units $\Bbbk^\times$ on the subset of non-zero vectors $\Bbbk_0^{n+1} = \Bbbk^{n+1} - \{\mathbf{0}\}$:

$$z \cdot \mathbf{x} = \mathbf{x} z^{-1}.$$

The set of orbits is denoted $\Bbbk\mathrm{P}^n$ and is called *n-dimensional $\Bbbk$-projective space.* Projective spaces An element of $\Bbbk\mathrm{P}^n$ written $[\mathbf{x}]$ is a set of the form

$$[\mathbf{x}] = \{\mathbf{x} z^{-1} : z \in \Bbbk^\times\} \subseteq \Bbbk_0^{n+1}.$$

Notice that $[\mathbf{x}] = [\mathbf{y}]$ if and only if there is a $z \in \Bbbk^\times$ for which $\mathbf{y} = \mathbf{x} z^{-1}$.

REMARK 5.6. Because of this we can identify elements $\Bbbk\mathrm{P}^n$ with $\Bbbk$-*lines* in $\Bbbk^{n+1}$ (i.e., 1-dimensional $\Bbbk$-vector subspaces). $\Bbbk\mathrm{P}^n$ is often viewed as the set of all such lines, particularly in the study of Projective Geometry.

There is a quotient map

$$q_n \colon \Bbbk_0^{n+1} \longrightarrow \Bbbk\mathrm{P}^n; \quad q_n(\mathbf{x}) = [\mathbf{x}],$$

and we give $\Bbbk\mathrm{P}^n$ the quotient topology which is Hausdorff and separable.

PROPOSITION 5.7. *$\Bbbk\mathrm{P}^n$ is a smooth manifold of dimension $\dim \Bbbk\mathrm{P}^n = n \dim_{\mathbb{R}} \Bbbk$. Moreover, the quotient map $q_n \colon \Bbbk_0^{n+1} \longrightarrow \Bbbk\mathrm{P}^n$ is smooth with surjective derivative at every point in $\Bbbk_0^{n+1}$.*

PROOF. For $r = 1, 2, \ldots, n$, set $\Bbbk\mathrm{P}^n_r = \{[\mathbf{x}] : x_r \neq 0\}$, where as usual we write $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{bmatrix}$. Then $\Bbbk\mathrm{P}^n_r \subseteq \Bbbk\mathrm{P}^n$ is open. There is a function

$$\sigma_r \colon \Bbbk\mathrm{P}^n_r \longrightarrow \Bbbk^n; \quad \sigma_r([\mathbf{x}]) = \begin{bmatrix} x_1 x_r^{-1} \\ x_2 x_r^{-1} \\ \vdots \\ x_{r-1} x_r^{-1} \\ x_{r+1} x_r^{-1} \\ \vdots \\ x_{n+1} x_r^{-1} \end{bmatrix}$$

which is a continuous bijection that is actually a homeomorphism. Whenever $r \neq s$, the induced map

$$\sigma_s^{-1} \circ \sigma_r \colon \sigma_r^{-1} \Bbbk\mathrm{P}^n_r \cap \Bbbk\mathrm{P}^n_s \longrightarrow \sigma_s^{-1} \Bbbk\mathrm{P}^n_r \cap \Bbbk\mathrm{P}^n_s$$

is given by

$$\sigma_s^{-1} \circ \sigma_r(\mathbf{x}) = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{s-1} \\ y_{s+1} \\ \vdots \\ y_{n+1} \end{bmatrix}$$

where

$$y_j = \begin{cases} x_j x_s^{-1} & \text{if } j \neq r, s, \\ x_s^{-1} & \text{if } j = r. \end{cases}$$

These $(n+1)$ charts form the standard atlas for n-dimensional projective space over $\Bbbk$. $\square$

An alternative description of $\Bbbk\mathrm{P}^n$ is given by considering the action of the subgroup

$$\Bbbk_1^\times = \{z \in \Bbbk^\times : |z| = 1\} \leqslant \Bbbk^\times$$

on the unit sphere $\mathbb{S}^{(n+1)d-1} \subseteq \Bbbk_0^{n+1}$. Notice that every element $[\mathbf{x}] \in \Bbbk\mathrm{P}^n$ contains elements of $\mathbb{S}^n$. Furthermore, if $\mathbf{x}, \mathbf{y} \in \Bbbk_0^{n+1}$ have unit length $|\mathbf{x}| = |\mathbf{y}| = 1$, then $[\mathbf{x}] = [\mathbf{y}]$ if and only if $\mathbf{y} = \mathbf{x}z^{-1}$ for some $z \in \Bbbk_1^\times$. This means we can also view $\Bbbk\mathrm{P}^n$ as the orbit space of this action of $\Bbbk_1^\times$ on $\mathbb{S}^{(n+1)d-1}$, and we also write the quotient map as $q_n \colon \mathbb{S}^{(n+1)d-1} \longrightarrow \Bbbk\mathrm{P}^n$; this map is also smooth.

PROPOSITION 5.8. *The quotient space given by the map $q_n \colon \mathbb{S}^{(n+1)d-1} \longrightarrow \Bbbk\mathrm{P}^n$ is compact Hausdorff.*

PROOF. This follows from the standard fact that the image of a compact space under a continuous mapping is compact. $\square$

Consider the action of $\mathrm{O}(n+1)$ on the unit sphere $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$. Then for $A \in \mathrm{O}(n+1)$, $z = \pm 1$ and $\mathbf{x} \in \mathbb{S}^n$, we have

$$A(\mathbf{x}z^{-1}) = (A\mathbf{x})z^{-1}.$$

Hence there is an induced action of $\mathrm{O}(n+1)$ on $\mathbb{R}\mathrm{P}^n$ given by

$$A \cdot [\mathbf{x}] = [A\mathbf{x}].$$

This action is transitive and also the matrices $\pm I_{n+1}$ fix every point of $\mathbb{R}\mathrm{P}^n$. There is also an action of $\mathrm{SO}(n+1)$ on $\mathbb{R}\mathrm{P}^n$; notice that $-I_{n+1} \in \mathrm{SO}(n+1)$ only if n is odd.

Similarly, $\mathrm{U}(n+1)$ and $\mathrm{SU}(n+1)$ act on $\mathbb{C}\mathrm{P}^n$ with scalar matrices wI_{n+1} ($w \in \mathbb{C}_1^\times$) fixing every element. Notice that if $wI_{n+1} \in \mathrm{SU}(n+1)$ then $w^{n+1} = 1$, so there are exactly $(n+1)$ such values.

Finally, $\mathrm{Sp}(n+1)$ acts on $\mathbb{H}\mathrm{P}^n$ and the matrices $\pm I_{n+1}$ fix everything.

There are some important new quotient Lie groups associated to these actions, the *projective unitary, special unitary and quaternionic symplectic groups*

$$\begin{aligned} \mathrm{PU}(n+1) &= \mathrm{U}(n+1)/\{wI_{n+1} : w \in \mathbb{C}_1^\times\}, \\ \mathrm{PSU}(n+1) &= \mathrm{SU}(n+1)/\{wI_{n+1} : w^{n+1} = 1\}, \\ \mathrm{PSp}(n+1) &= \mathrm{Sp}(n+1)/\{\pm I_{n+1}\}. \end{aligned}$$

Projective spaces are themselves homogeneous spaces. Consider the subgroup of $\mathrm{O}(n+1)$ consisting of elements of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & 0 \\ a_{21} & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-1\,1} & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \pm 1 \end{bmatrix}$$

We denote this subgroup of $\mathrm{O}(n+1)$ by $\mathrm{O}(n)\times\mathrm{O}(1)$. There is a subgroup $\widetilde{\mathrm{O}(n)} \leqslant \mathrm{SO}(n+1)$ whose elements have the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & 0 \\ a_{21} & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-1\,1} & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & w \end{bmatrix}$$

where

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} \\ a_{21} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ a_{n-1\,1} & \ddots & \ddots & a_{n-1\,n-1} \end{bmatrix} \in \mathrm{O}(n), \quad w = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} \\ a_{21} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ a_{n-1\,1} & \ddots & \ddots & a_{n-1\,n-1} \end{bmatrix}.$$

Similarly, there is a subgroup $\mathrm{U}(n)\times\mathrm{U}(1) \leqslant \mathrm{U}(n+1)$ whose elements have the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & 0 \\ a_{21} & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-1\,1} & \ddots & \ddots & a_{n-1\,n-1} & 0 \\ 0 & 0 & \cdots & 0 & w \end{bmatrix}$$

and $\widetilde{\mathrm{U}}(n) \leqslant \mathrm{SU}(n+1)$ with elements

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & 0 \\ a_{21} & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-1\,1} & \ddots & \ddots & a_{n-1\,n-1} & 0 \\ 0 & 0 & \cdots & 0 & w \end{bmatrix}$$

where

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} \\ a_{21} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ a_{n-1\,1} & \ddots & \ddots & a_{n-1\,n-1} \end{bmatrix} \in \mathrm{U}(n), \quad w = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} \\ a_{21} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ a_{n-1\,1} & \ddots & \ddots & a_{n-1\,n-1} \end{bmatrix}^{-1}.$$

Finally we have $\mathrm{Sp}(n)\times\mathrm{Sp}(1) \in \mathrm{Sp}(n+1)$ consisting of matrices of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & 0 \\ a_{21} & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-1\,1} & \ddots & \ddots & a_{n-1\,n-1} & 0 \\ 0 & 0 & \cdots & 0 & w \end{bmatrix}.$$

PROPOSITION 5.9. *There are diffeomorphisms between*

- $\mathbb{R}\mathrm{P}^n$ *and* $\mathrm{O}(n+1)/\mathrm{O}(n)\times\mathrm{O}(1)$, $\mathrm{SO}(n+1)/\widetilde{\mathrm{O}(n)}$;
- $\mathbb{C}\mathrm{P}^n$ *and* $\mathrm{U}(n+1)/\mathrm{U}(n)\times\mathrm{U}(1)$, $\mathrm{SU}(n+1)/\widetilde{\mathrm{U}(n)}$;

- $\mathbb{HP}^n$ *and* $\mathrm{Sp}(n+1)/\mathrm{Sp}(n)\times\mathrm{Sp}(1)$.

There are similar homogeneous space of the general and special linear groups giving these projective spaces. We illustrate this with one example.

$\mathrm{SL}_2(\mathbb{C})$ contains the matrix subgroup P consisting of its lower triangular matrices

$$\begin{bmatrix} u & 0 \\ w & v \end{bmatrix} \in \mathrm{SL}_2(\mathbb{C}).$$

This is often called a *parabolic subgroup*.

PROPOSITION 5.10. $\mathbb{CP}^1$ *is diffeomorphic to* $\mathrm{SL}_2(\mathbb{C})/\mathrm{P}$.

PROOF. There is smooth map

$$\psi\colon \mathrm{SL}_2(\mathbb{C}) \longrightarrow \mathbb{CP}^1; \quad \psi(A) = [A\mathbf{e}_2].$$

Notice that for $B = \begin{bmatrix} u & 0 \\ w & v \end{bmatrix} \in \mathrm{P}$,

$$\begin{bmatrix} u & 0 \\ w & v \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ v \end{bmatrix},$$

hence $[(AB)\mathbf{e}_2] = [A\mathbf{e}_2]$ for any $A \in \mathrm{SL}_2(\mathbb{C})$. This means that $\psi(A)$ only depends on the coset $A\mathrm{P} \in \mathrm{SL}_2(\mathbb{C})/\mathrm{P}$. It is easy to see that is onto and that the induced map $\mathrm{SL}_2(\mathbb{C})/\mathrm{P} \longrightarrow \mathbb{CP}^1$ is injective. □

4. Grassmannians

There are some important families of homogeneous spaces directly generalizing projective spaces. These are the real, complex and quaternionic *Grassmannians* which we now define.

Let $\mathrm{O}(k)\times\mathrm{O}(n-k) \leqslant \mathrm{O}(n)$ be closed the subgroup whose elements have the form

$$\begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix} \quad (A \in \mathrm{O}(k),\ B \in \mathrm{O}(n-k))$$

Similarly there are closed subgroups $\mathrm{U}(k)\times\mathrm{U}(n-k) \leqslant \mathrm{U}(n)$ and $\mathrm{Sp}(k)\times\mathrm{Sp}(n-k) \leqslant \mathrm{Sp}(n)$ with elements

$$\mathrm{U}(k)\times\mathrm{U}(n-k): \quad \begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix} \quad (A \in \mathrm{U}(k),\ B \in \mathrm{U}(n-k));$$
$$\mathrm{Sp}(k)\times\mathrm{Sp}(n-k): \quad \begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix} \quad (A \in \mathrm{Sp}(k),\ B \in \mathrm{Sp}(n-k)).$$

The associated homogeneous spaces are the Grassmannians

$$\mathrm{Gr}_{k,n}(\mathbb{R}) = \mathrm{O}(n)/\mathrm{O}(k)\times\mathrm{O}(n-k);$$
$$\mathrm{Gr}_{k,n}(\mathbb{C}) = \mathrm{U}(n)/\mathrm{U}(k)\times\mathrm{U}(n-k);$$
$$\mathrm{Gr}_{k,n}(\mathbb{H}) = \mathrm{Sp}(n)/\mathrm{Sp}(k)\times\mathrm{Sp}(n-k).$$

PROPOSITION 5.11. *For* $\Bbbk = \mathbb{R},\mathbb{C},\mathbb{H}$*, the Grassmannian* $\mathrm{Gr}_{k,n}(\Bbbk)$ *can be viewed as the set of all* k*-dimensional* $\Bbbk$*-vector subspaces in* $\Bbbk^n$.

PROOF. We describe the case $\Bbbk = \mathbb{R}$, the others being similar.

Associated to element $W \in \mathrm{O}(n)$ is the subspace spanned by the first k columns of W, say $\mathbf{w}_1,\ldots,\mathbf{w}_k$; we will denote this subspace by $\langle \mathbf{w}_1,\ldots,\mathbf{w}_k\rangle$. As the columns of W are an orthonormal set, they are linearly independent, hence $\dim_{\mathbb{R}} \langle \mathbf{w}_1,\ldots,\mathbf{w}_k\rangle = k$. Notice that the remaining $(n-k)$ columns give rise to another subspace $\langle \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n \rangle$ of dimension $\dim_{\mathbb{R}} \langle \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n \rangle = n - k$. In fact these are mutually orthogonal in the sense that

$$\begin{aligned}
\langle \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n \rangle &= \langle \mathbf{w}_1, \ldots, \mathbf{w}_k \rangle^{\perp} \\
&= \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{w}_r = 0, r = 1, \ldots, k\}, \\
\langle \mathbf{w}_1, \ldots, \mathbf{w}_k \rangle &= \langle \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n \rangle^{\perp} \\
&= \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{w}_r = 0, r = k+1, \ldots, n\}.
\end{aligned}$$

For a matrix

$$\begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix} \in \mathrm{O}(k) \times \mathrm{O}(n-k),$$

the columns in the product

$$W' = W \begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix}$$

span subspaces $\langle \mathbf{w}'_1, \ldots, \mathbf{w}'_k \rangle$ and $\left\langle \mathbf{w}'_{k+1}, \ldots, \mathbf{w}'_n \right\rangle$. But note that $\mathbf{w}'_1, \ldots, \mathbf{w}'_k$ are orthonormal and also linear combinations of $\mathbf{w}_1, \ldots, \mathbf{w}_k$; similarly, $\mathbf{w}'_{k+1}, \ldots, \mathbf{w}'_n$ are linear combinations of $\mathbf{w}_{k+1}, \ldots, \mathbf{w}_n$. Hence

$$\langle \mathbf{w}'_1, \ldots, \mathbf{w}'_k \rangle = \langle \mathbf{w}_1, \ldots, \mathbf{w}_k \rangle, \quad \left\langle \mathbf{w}'_{k+1}, \ldots, \mathbf{w}'_n \right\rangle = \langle \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n \rangle.$$

So there is a well defined function

$$\mathrm{O}(n)/\,\mathrm{O}(k) \times \mathrm{O}(n-k) \longrightarrow k\text{-dimensional vector subpaces of } \mathbb{R}^n$$

which sends the coset of W to the subspace $\langle \mathbf{w}_1, \ldots, \mathbf{w}_k \rangle$. This is actually a bijection.

Notice also that there is another bijection

$$\mathrm{O}(n)/\,\mathrm{O}(k) \times \mathrm{O}(n-k) \longrightarrow (n-k)\text{-dimensional vector subpaces of } \mathbb{R}^n$$

which sends the coset of W to the subspace $\langle \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n \rangle$. This corresponds to a diffeomorphism $\mathrm{Gr}_{k,n}(\mathbb{R}) \longrightarrow \mathrm{Gr}_{n-k,n}(\mathbb{R})$ which in turn corresponds to the obvious isomorphism $\mathrm{O}(k) \times \mathrm{O}(n-k) \longrightarrow \mathrm{O}(n-k) \times \mathrm{O}(k)$ induced by conjugation by a suitable element $P \in \mathrm{O}(n)$. $\square$

CHAPTER 6

Connectivity of matrix groups

1. Connectivity of manifolds

DEFINITION 6.1. A topological space X is *connected* if whenever $X = U \cup V$ with $U, V \neq \emptyset$, then $U \cap V \neq \emptyset$.

DEFINITION 6.2. A topological space X is *path connected* if whenever $x, y \in X$, there is a continuous path $p\colon [0,1] \longrightarrow X$ with $p(0) = x$ and $p(1) = y$.

X is *locally path connected* if every point is contained in a path connected open neighbourhood.

The following result is fundamental to Real Analysis.

PROPOSITION 6.3. *Every interval* $[a,b], [a,b), (a,b], (a,b) \subseteq \mathbb{R}$ *is path connected and connected. In particular,* $\mathbb{R}$ *is path connected and connected.*

PROPOSITION 6.4. *If* X *is a path connected topological space then* X *is connected.*

PROOF. Suppose X is not connected. Then $X = U \cup V$ where $U, V \subseteq X$ are non-empty and $U \cap V = \emptyset$. Let $x \in U$ and $y \in V$. By path connectedness of there X, is a continuous map $p\colon [0,1] \longrightarrow X$ with $p(0) = x$ and $p(1) = y$. Then $[0,1] = p^{-1}U \cup p^{-1}V$ expresses $[0,1]$ as a union of open subsets with no common elements. But this contradicts the connectivity of $[0,1]$. So X must be connected. $\square$

PROPOSITION 6.5. *Let* X *be a connected topological space which is locally path connected. Then* X *is path connected.*

PROOF. Let $x \in X$, and set

$$X_x = \{y \in X : \exists\, p\colon [0,1] \longrightarrow X \text{ continuous such that } p(0) = x \text{ and } p(1) = y\}.$$

Then for each $y \in X_x$, there is a path connected open neighbourhood U_y. But for each point $z \in U_y$ there is a continuous path from to z via y, hence $U_y \subseteq X_x$. This shows that

$$X_x = \bigcup_{y \in X_x} U_y \subseteq X$$

is open in X. Similarly, if $w \in X - X_x$, then $X_w \subseteq X - X_x$ and this is also open. But then so is

$$X - X_x = \bigcup_{w \in X - X_x} X_w.$$

Hence $X = X_x \cup (X - X_x)$, and so by connectivity, $X_x = \emptyset$ or $X - X_x = \emptyset$. So X is path connected. $\square$

PROPOSITION 6.6. *If the topological spaces* X *and* Y *are path connected then their product* $X \times Y$ *is path connected.*

COROLLARY 6.7. *For* $n \geqslant 1$, $\mathbb{R}^n$ *is path connected and connected.*

It is also useful to record the following standard results.

Proposition 6.8. i) *Let $n \geqslant 2$. The unit sphere $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ is path connected. In $\mathbb{S}^0 = \{\pm 1\} \subseteq \mathbb{R}$, the subsets $\{1\}$ and $\{-1\}$ are path connected. The set of non-zero vectors $\mathbb{R}^n_0 \subseteq \mathbb{R}^n$ is path connected.*
ii) *For $n \geqslant 1$, the sets of non-zero complex and quaternionic vectors $\mathbb{C}^n_0 \subseteq \mathbb{C}^n$ and $\mathbb{H}^n_0 \subseteq \mathbb{H}^n$ are path connected.*

Proposition 6.9. *Every manifold is locally path connected. Hence every connected manifold is path connected.*

Proof. Every point is contained in an open neighbourhood homeomorphic to some open subset of $\mathbb{R}^n$ which can be taken to be an open disc which is path connected. The second statement now follows from Proposition 6.5. $\square$

Theorem 6.10. *Let M be a connected manifold and $N \subseteq M$ a non-empty submanifold which is also a closed subset. If $\dim N = \dim M$ then $N = M$.*

Proof. Since $N \subseteq M$ is closed, $M - N \subseteq M$ is open. But $N \subseteq M$ is also open since every element is contained in an open subset of M contained in N; hence $M - N \subseteq M$ is closed. Since M is connected, $M - N = \emptyset$. $\square$

Proposition 6.11. *Let G be a Lie group and $H \leqslant G$ a closed subgroup. If G/H and H are connected, then so is G.*

Proof. First we remark on the following: for any $g \in G$, left translation map $\mathrm{L}_g \colon H \longrightarrow gH$ provides a homeomorphism between these spaces, hence gH is connected since H is.

Suppose that G is not connected, and let $U, V \subseteq G$ be nonempty open subsets for which $U \cap V = \emptyset$ and $U \cup V = G$. By Lemma 5.1 the projection $\pi \colon G \longrightarrow G/H$ is a surjective open mapping, so $\pi U, \pi V \subseteq G/H$ are open subsets for which $\pi U \cup \pi V = G/H$. As G/H is connected, there is an element gH say in $\pi U \cap \pi V$. In G we have

$$gH = (gH \cap U) \cup (gH \cap V),$$

where $(gH \cap U), (gH \cap V) \subseteq gH$ are open subsets in the subspace topology on gH since U, V are open in G. By connectivity of gH, this can only happen if $gH \cap U = \emptyset$ or $gH \cap V = \emptyset$, since these are subsets of U, V which have no common elements. As

$$\pi^{-1} gH = \{gh : h \in H\},$$

this is false, so $(gH \cap U) \cap (gH \cap V) \neq \emptyset$ which implies that $U \cap V \neq \emptyset$. This contradicts the original assumption on U, V. $\square$

This result together with Proposition 6.9 gives a useful criterion for path connectedness of a Lie group which may need to be applied repeatedly to show a particular example is path connected. Recall that by Theorem 4.31, a closed subgroup of a Lie group is a submanifold.

Proposition 6.12. *Let G be a Lie group and $H \leqslant G$ a closed subgroup. If G/H and H are connected, then G is path connected.*

2. Examples of path connected matrix groups

In this section we apply Proposition 6.12 to show that many familiar matrix groups are path connected.

Example 6.13. For $n \geqslant 1$, $\mathrm{SL}_n(\mathbb{R})$ is path connected.

PROOF. For the real case, we proceed by induction on n. Notice that $\mathrm{SL}_1(\mathbb{R}) = \{1\}$, which is certainly connected. Now suppose that $\mathrm{SL}_{n-1}(\mathbb{R})$ is path connected for some $n \geqslant 2$.

Recall that $\mathrm{SL}_n(\mathbb{R})$ acts continuously on $\mathbb{R}^n$ by matrix multiplication. Consider the continuous function

$$f\colon \mathrm{SL}_n(\mathbb{R}) \longrightarrow \mathbb{R}^n; \quad f(A) = A\mathbf{e}_n.$$

The image of f is $\operatorname{im} f = \mathbb{R}^n_0 = \mathbb{R}^n - \{0\}$ since every vector $\mathbf{v} \in \mathbb{R}^n_0$ can be extended to a basis

$$\mathbf{v}_1 \ldots, \mathbf{v}_{n-1}, \mathbf{v}_n = \mathbf{v}$$

of $\mathbb{R}^n$, and we can multiply $\mathbf{v}_1$ by a suitable scalar to ensure that the matrix $A_{\mathbf{v}}$ with these vectors as its columns has determinant 1. Then $A_{\mathbf{v}}\mathbf{e}_n = \mathbf{v}$.

Notice that $P\mathbf{e}_n = \mathbf{e}_n$ if and only if

$$P = \begin{bmatrix} Q & \mathbf{0} \\ \mathbf{w} & 1 \end{bmatrix},$$

where Q is $(n-1)\times(n-1)$ with $\det Q = 1$, is the $(n-1)\times 1$ zero vector and $\mathbf{w}$ is an arbitrary $1\times(n-1)$ vector. The set of all such matrices is the stabilizer of $\mathbf{e}_n$, $\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$, which is a closed subgroup of $\mathrm{SL}_n(\mathbb{R})$. More generally, $A\mathbf{e}_n = \mathbf{v}$ if and only if

$$A = A_{\mathbf{v}}P \quad \text{for some } P \in \mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n).$$

So the homogeneous space $\mathrm{SL}_n(\mathbb{R})/\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$ is homeomorphic to $\mathbb{R}^n_0$.

Since $n \geqslant 2$, it is well known that $\mathbb{R}^n_0$ is path connected, hence is connected. This implies that $\mathrm{SL}_n(\mathbb{R})/\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$ is connected.

The subgroup $\mathrm{SL}_{n-1}(\mathbb{R}) \leqslant \mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$ is closed and the well defined map

$$\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)/\mathrm{SL}_{n-1}(\mathbb{R}) \longrightarrow \mathbb{R}^{n-1}; \quad \begin{bmatrix} Q & \mathbf{0} \\ \mathbf{w} & 1 \end{bmatrix}\mathrm{SL}_{n-1}(\mathbb{R}) \longmapsto (\mathbf{w}Q^{-1})^T$$

is a homeomorphism so the homogeneous space $\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)/\mathrm{SL}_{n-1}(\mathbb{R})$ is homeomorphic to $\mathbb{R}^{n-1}$. Hence by Corollary 6.7 together with the inductive assumption, $\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$ is path connected. We can combine this with the connectivity of $\mathbb{R}^n_0$ to deduce that $\mathrm{SL}_n(\mathbb{R})$ is path connected, demonstrating the inductive step. □

EXAMPLE 6.14. For $n \geqslant 1$, $\mathrm{GL}^+_n(\mathbb{R})$ is path connected.

PROOF. Since $\mathrm{SL}_n(\mathbb{R}) \leqslant \mathrm{GL}^+_n(\mathbb{R})$, it suffices to show that $\mathrm{GL}^+_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{R})$ is path connected. But for this we can use the determinant to define a continuous map

$$\det\colon \mathrm{GL}^+_n(\mathbb{R}) \longrightarrow \mathbb{R}^+ = (0, \infty),$$

which is surjective onto a path connected space. The homogeneous space $\mathrm{GL}^+_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{R})$ is then diffeomorphic to $\mathbb{R}^+$ and hence is path connected. So $\mathrm{GL}^+_n(\mathbb{R})$ is path connected. □

This shows that

$$\mathrm{GL}_n(\mathbb{R}) = \mathrm{GL}^+_n(\mathbb{R}) \cup \mathrm{GL}^-_n(\mathbb{R})$$

is the decomposition of $\mathrm{GL}_n(\mathbb{R})$ into two path connected components.

EXAMPLE 6.15. For $n \geqslant 1$, $\mathrm{SO}(n)$ is path connected. Hence

$$\mathrm{O}(n) = \mathrm{SO}(n) \cup \mathrm{O}(n)^-$$

is the decomposition of $\mathrm{O}(n)$ into two path connected components.

PROOF. For $n = 1$, $\mathrm{SO}(1) = \{1\}$. So we will assume that $n \geqslant 2$ and proceed by induction on n. SO assume that $\mathrm{SO}(n-1)$ is path connected.

Consider the continuous action of on $\mathbb{R}^n$ by left multiplication. The stabilizer of $\mathbf{e}_n$ is $\mathrm{SO}(n-1) \leqslant \mathrm{SO}(n)$ thought of as the closed subgroup of matrices of the form

$$\begin{bmatrix} P & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

with $P \in \mathrm{SO}(n-1)$ and $\mathbf{0}$ the $(n-1) \times 1$ zero matrix. The orbit of $\mathbf{e}_n$ is the unit sphere $\mathbb{S}^{n-1}$ which is path connected. Since the orbit space is also diffeomorphic to $\mathrm{SO}(n)/\mathrm{SO}(n-1)$ we have the inductive step. $\square$

EXAMPLE 6.16. For $n \geqslant 1$, $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are path connected.

PROOF. For $n = 1$, $\mathrm{U}(1)$ is the unit circle in $\mathbb{C}$ while $\mathrm{SU}(1) = \{1\}$, so both of these are path connected. Assume that $\mathrm{U}(n-1)$ and $\mathrm{SU}(n-1)$ are path connected for some $n \geqslant 2$.

Then $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ act on $\mathbb{C}^n$ by matrix multiplication and by arguments of Chapter 5,

$$\mathrm{Stab}_{\mathrm{U}(n)}(\mathbf{e}_n) = \mathrm{U}(n-1), \quad \mathrm{Stab}_{\mathrm{SU}(n)}(\mathbf{e}_n) = \mathrm{SU}(n-1).$$

We also have

$$\mathrm{Orb}_{\mathrm{U}(n)}(\mathbf{e}_n) = \mathrm{Orb}_{\mathrm{SU}(n)}(\mathbf{e}_n) = \mathbb{S}^{2n-1},$$

where $\mathbb{S}^{2n-1} \subseteq \mathbb{C}^n \cong \mathbb{R}^{2n}$ denotes the unit sphere consisting of unit vectors. Since $\mathbb{S}^{2n-1}$ is path connected, we can deduce that $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are too, which gives the inductive step. $\square$

3. The path components of a Lie group

Let G be a Lie group. We say that two elements $x, y \in G$ are *connected by a path in* G if there is a continuous path $p\colon [0,1] \longrightarrow G$ with $p(0) = x$ and $p(1) = y$; we will then write $x \underset{G}{\sim} y$.

LEMMA 6.17. $\underset{G}{\sim}$ *is an equivalence relation on* G.

For $g \in G$, we can consider the equivalence class of g, the *path component* of g in G,

$$G_g = \{x \in G : x \underset{G}{\sim} g\}.$$

PROPOSITION 6.18. *The path component of the identity is a clopen normal subgroup of* G, $G_1 \lhd G$*; hence it is a closed Lie subgroup of dimension* $\dim G$.

The path component G_g *agrees with the coset of* g *with respect to* G_1, $G_g = gG_1 = G_1 g$ *and is a closed submanifold of* G.

PROOF. By Proposition 6.9, G_g contains an open neighbourhood of g in G. This shows that every component is actually a submanifold of G with dimension equal to $\dim G$. The argument used in the proof of Proposition 6.5 shows that each is G_g actually clopen in G.

Let $x, y \in G_1$. Then there are continuous paths $p, q\colon [0,1] \longrightarrow G$ with $p(0) = 1 = q(0)$, $p(0) = x$ and $q(0) = y$. The product path

$$r\colon [0,1] \longrightarrow G; \quad r(t) = p(t)q(t)$$

has $r(0) = 1$ and $r(1) = xy$. So $G_1 \leqslant G$. For $g \in G$, the path

$$s\colon [0,1] \longrightarrow G; \quad s(t) = gp(t)g^{-1}$$

has $s(0) = 1$ and $s(1) = gxg^{-1}$; hence $G_1 \lhd G$. If $z \in gG_1 = G_1 g$, then $g^{-1}z \in G_1$ and so there is a continuous path $h\colon [0,1] \longrightarrow G$ with $h(0) = 1$ and $h(1) = g^{-1}z$. Then the path

$$gh\colon [0,1] \longrightarrow G; \quad gh(t) = g(h(t))$$

has $gh(0) = g$ and $gh(1) = z$. So each coset gG_1 is path connected, hence $gG_1 \subseteq G_g$. To show equality, suppose that g is connected by a path $k\colon [0,1] \longrightarrow G$ in G to $w \in G_g$. Then the path $g^{-1}k$ connects 1 to $g^{-1}w$, so $g^{-1}w \in G_1$, giving $w \in gG_1$. This shows that $G_g \subseteq gG_1$. $\square$

The quotient group G/G_1 is the *group of path components* of G, which we will denote by $\pi_0 G$.

Example 6.19. We have the following groups of path components:

$$\pi_0\,\mathrm{SO}(n) = \pi_0\mathrm{SL}_n(\mathbb{R}) = \pi_0\,\mathrm{SU}(n) = \pi_0\,\mathrm{U}(n) = \pi_0\mathrm{SL}_n(\mathbb{C}) = \pi_0\mathrm{GL}_n(\mathbb{C}) = \{1\},$$
$$\pi_0\,\mathrm{O}(n) \cong \pi_0\mathrm{GL}_n(\mathbb{R}) \cong \{\pm 1\}.$$

Example 6.20. Let

$$T = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} : \theta \in \mathbb{R} \right\} \leqslant \mathrm{SO}(3).$$

and let $G = \mathrm{N}_{\mathrm{SO}(3)}(T) \leqslant \mathrm{SO}(3)$ be its normalizer. Then T and G are Lie subgroups of $\mathrm{SO}(3)$ and $\pi_0 G \cong \{\pm 1\}$.

Proof. A straightforward computation shows that

$$\mathrm{N}_{\mathrm{SO}(3)}(T) = T \cup \left\{ \begin{bmatrix} -\cos\theta & \sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{bmatrix} : \theta \in \mathbb{R} \right\} = T \cup \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} T.$$

Notice that T is isomorphic to the unit circle,

$$T \cong \mathbb{T}; \qquad \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \longleftrightarrow e^{\theta i}.$$

This implies that T is path connected and abelian since $\mathbb{T}$ is. The function

$$\varphi\colon G \longrightarrow \mathbb{R}^\times; \quad \varphi([a_{ij}]) = a_{3\,3}$$

is continuous with

$$\varphi^{-1}\mathbb{R}^+ = T, \quad \varphi^{-1}\mathbb{R}^- = T \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

hence these are clopen subsets. This shows that the path components of G are

$$G_I = T, \qquad \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} T.$$

Hence $\pi_0 G \cong \{\pm 1\}$.

Notice that $\mathrm{N}_{\mathrm{SO}(3)}(T)$ acts by conjugation on T and in fact every element of $T \triangleleft \mathrm{N}_{\mathrm{SO}(3)}(T)$ acts trivially since T is abelian. Hence $\pi_0 G$ acts on T with the action of the non-trivial coset given by conjugation by the matrix $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$,

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}^{-1}$$
$$= \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}$$

which corresponds to the inversion homomorphism on the unit circle $\mathbb{T} \cong T$. □

EXAMPLE 6.21. Let $T = \{x1+yi : x, y \in \mathbb{R},\ x^2+y^2 = 1\} \leqslant \mathrm{Sp}(1)$, the group of unit quaternions. Let $G = \mathrm{N}_{\mathrm{Sp}(1)}(T) \leqslant \mathrm{Sp}(1)$ be its normalizer. Then T and G are Lie subgroups of $\mathrm{Sp}(1)$ and $\pi_0 G \cong \{\pm 1\}$.

PROOF. By a straightforward calculation,

$$G = T \cup \{xj - yk : x, y \in \mathbb{R},\ x^2 + y^2 = 1\} = T \cup jT.$$

T is isomorphic to the unit circle so is path connected and abelian. The function

$$\theta \colon G \longrightarrow \mathbb{R}; \quad \theta(t1 + xi + yj + zk) = y^2 + z^2,$$

is continuous and

$$\theta^{-1}0 = T, \quad \theta^{-1}1 = jT.$$

Hence the path components of G are T, jT. So $\pi_0 G = \cong \{\pm 1\}$.

The conjugation action of G on T has every element of T acting trivially, so $\pi_0 G$ acts on T. The action of the non-trivial coset is given by conjugation with j,

$$j(x1 + yi)j^{-1} = x1 - yi,$$

corresponding to the inversion map on the unit circle $\mathbb{T} \cong T$. □

The significance of such examples will become clearer when we discuss maximal tori and their normalizers in Chapter 7.

4. Another connectivity result

The following result will be used in Chapter 7.

PROPOSITION 6.22. *Let G be a connected Lie group and $H \leqslant G$ a subgroup which contains an open neighbourhood of* 1 *in G. Then $H = G$.*

PROOF. Let $U \subseteq H$ be an open neighbourhood of 1 in G. Since the inverse map $\mathrm{inv} \colon G \longrightarrow G$ is a homeomorphism and maps H into itself, by replacing U with $U \cap \mathrm{inv}\, U$ if necessary, we can assume that inv maps the open neighbourhood into itself, i.e., $\mathrm{inv}\, U = U$.

For $k \geqslant 1$, consider

$$U^k = \{u_1 \cdots u_k \in G : u_j \in U\} \subseteq H.$$

Notice that $\mathrm{inv}\, U^k = U^k$. Also, $U^k \subseteq G$ is open since for $u_1, \ldots, u_k \in U$,

$$u_1 \cdots u_k \in \mathrm{L}_{u_1 \cdots u_{k-1}} U \subseteq U^k$$

where $\mathrm{L}_{u_1\cdots u_{k-1}}U = \mathrm{L}^{-1}_{(u_1\cdots u_{k-1})^{-1}}U$ is an open subset of G. Then

$$V = \bigcup_{k\geqslant 1} U^k \subseteq H$$

satisfies $\operatorname{inv} V = V$.

V is closed in G since given $g \in G - V$, for the open set $gV \subseteq G$, if $x \in gV \cap V$ there are $u_1, \dots, u_r, v_1 \dots, v_s \in U$ such that

$$gu_1 \cdots u_r = v_1 \cdots v_s,$$

implying $g = v_1 \cdots v_s u_1^{-1} \cdots u_r^{-1} \in V$, contradicting the assumption on g.

So V is a nonempty clopen subset of G, which is connected. Hence $G - V = \emptyset$, and therefore $V = G$, which also implies that $H = G$. $\square$

CHAPTER 7

Compact connected Lie groups and their maximal tori

In this chapter we will describe some results on the structure of compact connected Lie groups, focusing on the important notion of a *maximal torus* which is central to the classification of simple compact connected Lie Groups. From Chapter 6 we know that many familiar examples of compact matrix groups are path connected.

Although we state results for arbitrary Lie groups we will often only give proofs for matrix groups. However, there is no loss in generality in assuming this because of the following important result which we will not prove (the proof uses ideas of Haar measure and integration on such compact Lie groups).

THEOREM 7.1. *Let G be a compact Lie group. Then there are injective Lie homomorphisms $G \longrightarrow \mathrm{O}(m)$ and $G \longrightarrow \mathrm{U}(n)$ for some m, n. Hence G is a matrix group.*

1. Tori

The *circle group*

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\} \leqslant \mathbb{C}^{\times}$$

is a matrix group since $\mathbb{C}^{\times} = \mathrm{GL}_1(\mathbb{C})$. For each $r \geqslant 1$, the *standard torus of rank r* is

$$\mathbb{T}^r = \{\mathrm{diag}(z_1, \ldots, z_r) : \forall k,\ |z_k| = 1\} \leqslant \mathrm{GL}_r(\mathbb{C}).$$

This is a matrix group of dimension r. More generally, a *torus of rank r* is a Lie group isomorphic to $\mathbb{T}^r$. We will often view elements of $\mathbb{T}^r$ as sequences of complex numbers $(z_1, \ldots, z_r)$ with $|z_k| = 1$, this corresponds to the identification

$$\mathbb{T}^r \cong \mathbb{T} \times \cdots \times \mathbb{T} \leqslant (\mathbb{C}^{\times})^r \quad (r \text{ factors}).$$

Such a torus is a compact path connected abelian Lie group.

Now let G be Lie group and $T \leqslant G$ a closed subgroup which is a torus. Then T is *maximal in G* if the only torus $T' \leqslant G$ for which $T \leqslant T'$ is T itself. Here are some examples.

For $\theta \in [0, 2\pi)$, let

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in \mathrm{SO}(2).$$

More generally, for each $n \geqslant 1$, and $\theta_i \in [0, 2\pi)$ $(i = 1, \ldots, n)$, let

$$R_{2n}(\theta_1, \ldots, \theta_n) = \begin{bmatrix} R(\theta_1) & O & \cdots & \cdots & \cdots & O \\ O & R(\theta_2) & O & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ O & \cdots & \cdots & \cdots & O & R(\theta_n) \end{bmatrix} \in \mathrm{SO}(2n),$$

$$R_{2n+1}(\theta_1, \ldots, \theta_n) = \begin{bmatrix} R(\theta_1) & O & \cdots & \cdots & \cdots & \cdots & O \\ O & R(\theta_2) & O & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & O & R(\theta_n) & O \\ O & \cdots & \cdots & \cdots & \cdots & O & 1 \end{bmatrix} \in \mathrm{SO}(2n+1),$$

where each entry marked O is an appropriately sized block so that these are matrices of size $2n \times 2n$ and $(2n+1) \times (2n+1)$ respectively.

By identifying $\mathbb{C}$ with $\mathbb{R}^2$ as real vector spaces using the bases $\{1, i\}$ and $\{e_1, e_2\}$, we obtain an isomorphism

$$\mathrm{U}(1) \longrightarrow \mathrm{SO}(2), \quad e^{\theta i} \longmapsto R_2(\theta).$$

Proposition 7.2. *Each of the following is a maximal torus in the stated group.*

$$\{R_{2n}(\theta_1, \ldots, \theta_n) : \forall k,\ \theta_k \in [0, 2\pi)\} \leqslant \mathrm{SO}(2n).$$
$$\{R_{2n+1}(\theta_1, \ldots, \theta_n) : \forall k,\ \theta_k \in [0, 2\pi)\} \leqslant \mathrm{SO}(2n+1).$$
$$\{\mathrm{diag}(z_1, \ldots, z_n) : \forall k,\ |z_k| = 1\} \leqslant \mathrm{U}(n).$$
$$\{\mathrm{diag}(z_1, \ldots, z_n) : \forall k,\ |z_k| = 1,\ z_1 \cdots z_n = 1\} \leqslant \mathrm{SU}(n).$$
$$\{\mathrm{diag}(z_1, \ldots, z_n) : \forall k\ z_k \in \mathbb{C},\ |z_k| = 1\} \leqslant \mathrm{Sp}(n).$$

The maximal tori listed will be referred to as the standard maximal tori for these groups.

Proposition 7.3. *Let T be a torus. Then T is compact, path-connected and abelian.*

Proof. Since the circle $\mathbb{T}$ is compact and abelian the same is true for $\mathbb{T}^r$ and hence for any torus. If $(z_1, \ldots, z_r) \in \mathbb{T}^r$, let $z_k = e^{\theta_k i}$. Then there is a continuous path

$$p \colon [0, 1] \longrightarrow \mathbb{T}^r; \quad p(t) = (e^{t\theta_1 i}, \ldots, e^{t\theta_r i}),$$

with $p(0) = (1, \ldots, 1)$ and $p(1) = (z_1, \ldots, z_r)$. So $\mathbb{T}^r$ and hence any torus is path connected □

Theorem 7.4. *Let H be a compact Lie group. Then H is a torus if and only if it is connected and abelian.*

Proof. We know that H is a compact Lie group. Every torus is path connected and abelian by Proposition 7.3. So we need to show that when H is connected and abelian it is a torus since by Proposition 6.9 it would be path connected.

Suppose that $\dim H = r$ and let $\mathfrak{h}$ be the Lie algebra of H; then $\dim \mathfrak{h} = r$. From the definition of the Lie bracket in the proof of Theorem 2.14, for $X, Y \in \mathfrak{h}$,

$$[X, Y] = \frac{\mathrm{d}}{\mathrm{d}\,s}_{|s=0} \frac{\mathrm{d}}{\mathrm{d}\,t}_{|t=0} \exp(sX)\exp(tY)\exp(-sX) = 0$$

since $\exp(sX), \exp(tY) \in H$ and so $\exp(sX)\exp(tY)\exp(-sX) = \exp(tY)$ because H is abelian. Thus all Lie brackets in $\mathfrak{h}$ are zero. Consider the exponential map $\exp \colon \mathfrak{h} \longrightarrow H$. For $X, Y \in \mathfrak{h}$, Propositions

1.33 and 1.32 give

$$\exp(X)\exp(Y) = \exp(X+Y), \quad \exp(-X) = \exp(X)^{-1}.$$

So $\exp \mathfrak{h} = \operatorname{im} \exp \subseteq H$ is a subgroup. By Proposition 1.35 $\exp \mathfrak{h}$ is a subgroup containing a neighbourhood of 1, hence by Proposition 6.22, $\exp \mathfrak{h} = H$.

As exp is a continuous homomorphism, its kernel $K = \ker \exp$ must be discrete since otherwise $\dim \exp(\mathfrak{h}) < r$. This means that $K \subseteq \mathfrak{h}$ is a free abelian subgroup with basis $\{v_1, \ldots, v_s\}$ for some $s \leqslant r$. Extending this to an $\mathbb{R}$-basis $\{v_1, \ldots, v_s, v_{s+1}, \ldots, v_r\}$ of $\mathfrak{h}$ we obtain isomorphisms of Lie groups

$$\exp(\mathfrak{h}) \cong \mathfrak{h}/K \cong \mathbb{R}^s/\mathbb{Z}^s \times \mathbb{R}^{r-s}.$$

But the right hand term is only compact if $s = r$, hence K contains a basis of $\mathfrak{h}$ and

$$\mathbb{R}^r/\mathbb{Z}^r \cong \mathfrak{h}/K \cong H.$$

Since $\mathbb{T} \cong \mathbb{R}/\mathbb{Z}$, this gives H the structure of a torus. □

Notice that in this proof and that of Theorem 7.4, we made use of the following fact.

PROPOSITION 7.5. *Let T be a torus of rank r. Then the exponential map* $\exp\colon \mathfrak{t} \longrightarrow T$ *is a surjective homomorphism of Lie groups, whose kernel is a discrete subgroup isomorphic to $\mathbb{Z}^r$. Hence there is an isomorphism of Lie groups $\mathbb{R}^r/\mathbb{Z}^r \cong T$.*

In the proof Theorem 4.35, we met the idea of a topological generator of the circle group. It turns out that all tori have such generators.

DEFINITION 7.6. Let G be a Lie group. Then an element $g \in G$ is a *topological generator* or just a *generator* of G if the cyclic subgroup $\langle g \rangle \leqslant G$ is dense in G, i.e., $\overline{\langle g \rangle} = G$.

PROPOSITION 7.7. *Every torus T has a generator.*

PROOF. Without loss of generality we can assume $T = \mathbb{R}^r/\mathbb{Z}^r$ and will write elements in the form $[x_1, \ldots, x_r] = (x_1, \ldots, x_r) + \mathbb{Z}^r$. The group operation is then addition. Let $U_1, U_2, U_3, \ldots$ be a countable base for the topology on T.

A *cube of side* $\varepsilon > 0$ in T is a subset of the form

$$C([u_1, \ldots, u_r], \varepsilon) = \{[x_1, \ldots, x_r] \in T : |x_k - u_k| < \varepsilon/2 \ \forall k\},$$

for some $[u_1, \ldots, u_r] \in T$. Such a cube is the image of a cube in $\mathbb{R}^r$ under the quotient map $\mathbb{R}^r \longrightarrow T$.

Let $C_0 \subseteq T$ be a cube of side $\varepsilon > 0$. Suppose that we have a decreasing sequence of cubes C_k of side ε_k,

$$C_0 \supseteq C_1 \supseteq \cdots \supseteq C_m,$$

where for each $0 \leqslant k \leqslant m$, there is an integer N_k satisfying $N_k\varepsilon_k > 1$ and $N_kC_k \subseteq U_k$. Now choose an integer N_{m+1} large enough to guarantee that $N_{m+1}C_m = T$. Now choose a small cube $C_{m+1} \subseteq C_m$ of side ε_{m+1} so that $N_{m+1}C_{m+1} \subseteq U_{m+1}$. Then if $\mathbf{z} = [z_1, \ldots, z_r] \in \bigcap_{k \geqslant 1} C_k$, we have $N_k\mathbf{z} \in C_k$ for each k, hence the powers of $\mathbf{z}$ are dense in T, so $\mathbf{z}$ is a generator of T. □

2. Maximal tori in compact Lie groups

We now begin to study the structure of compact Lie groups in terms of their maximal tori. Throughout the section, let G be a compact connected Lie group and $T \leqslant G$ a maximal torus.

THEOREM 7.8. *If $g \in G$, there is an $x \in G$ such that $g \in xTx^{-1}$, i.e., g is conjugate to an element of T. Equivalently,*

$$G = \bigcup_{x \in G} xTx^{-1}.$$

Proof. The proof this uses the powerful *Lefschetz Fixed Point Theorem* from Algebraic Topology and we only give a sketch indicating how this is used.

The quotient space G/T is a compact space and each element $g \in G$ gives rise to a continuous map

$$\mu_g \colon G/T \longrightarrow G/T; \quad \mu_g(xT) = (gx)T = gxT.$$

Since G is path connected, there is a continuous map

$$p\colon [0,1] \times G/T \longrightarrow G/T;$$

for which $p(0, xT) = xT$ and $p(1, xT) = gxT$, i.e., p is a *homotopy* $\mathrm{Id}_{G/T} \simeq \mu_g$.

The *Lefschetz Fixed Point Theorem* asserts that μ_g has a fixed point provided the *Euler characteristic* $\chi(G/T)$ is non-zero. Indeed it can be shown that $\chi(G/T) \neq 0$, so this tells us that there is an $x \in G$ such that $gxT = xT$, or equivalently $g \in xTx^{-1}$. □

Theorem 7.9. *If $T, T' \leqslant G$ are maximal tori then they are conjugate in G, i.e., there is a $y \in G$ such that $T' = yTy^{-1}$.*

Proof. By Proposition 7.7, T' has a generator t say. By Theorem 7.8, there is a $y \in G$ such that $t \in$, so $T' \leqslant yTy^{-1}$ As T' is a maximal torus and yTy^{-1} is a torus, we must have $T' = yTy^{-1}$. □

The next result gives some important special cases related to the examples of Proposition 7.2. Notice that if $A \in \mathrm{SO}(m)$, $A^{-1} = A^T$, while if $B \in \mathrm{U}(m)$, $B^{-1} = B^*$.

Theorem 7.10 (Principle Axis Theorem). *In each of the following matrix groups every element is conjugate to one of the stated form.*

- $\mathrm{SO}(2n)$: $R_{2n}(\theta_1, \ldots, \theta_n)$, $\forall k\ \theta_k \in [0, 2\pi)$;
- $\mathrm{SO}(2n+1)$: $R_{2n+1}(\theta_1, \ldots, \theta_n)$, $\forall k\ \theta_k \in [0, 2\pi)$;
- $\mathrm{U}(n)$: $\mathrm{diag}(z_1, \ldots, z_n)$, $\forall k\ z_k \in \mathbb{C}$, $|z_k| = 1$;
- $\mathrm{SU}(n)$: $\mathrm{diag}(z_1, \ldots, z_n)$, $\forall k\ z_k \in \mathbb{C}$, $|z_k| = 1$, $z_1 \cdots z_n = 1$
- $\mathrm{Sp}(n)$: $\mathrm{diag}(z_1, \ldots, z_n)$, $\forall k\ z_k \in \mathbb{C}$, $|z_k| = 1$.

We can also deduce a results on the Lie algebra $\mathfrak{g}$ of such a compact, connected matrix group G. Recall that for each $g \in G$, there is a linear transformation

$$\mathrm{Ad}_g \colon G \longrightarrow \mathfrak{g}; \quad \mathrm{Ad}_g(t) = gtg^{-1}.$$

Proposition 7.11. *Suppose that $g \in G$ and $H, H' \leqslant G$ are Lie subgroups with $gHg^{-1} = H'$. Then $\mathrm{Ad}_g \mathfrak{h} = \mathfrak{h}'$.*

Proof. By definition, for $x \in \mathfrak{h}$ there is a curve $\gamma\colon (-\varepsilon, \varepsilon) \longrightarrow H$ with $\gamma(0) = 1$ and $\gamma'(0) = x$. Then

$$\mathrm{Ad}_g(t) = \frac{\mathrm{d}}{\mathrm{d}\,t} g\gamma(t)g^{-1}_{|t=0} \in \mathfrak{h}',$$

since $t \mapsto g\gamma(t)g^{-1}$ is a curve in H'. □

If $x, y \in \mathfrak{g}$ and $y = \mathrm{Ad}_g(x)$ we will say that x is *conjugate in G* to y. This defines an equivalence relation on $\mathfrak{g}$.

For $t \in \mathbb{R}$, let

$$R'(t) = \begin{bmatrix} 0 & -t \\ t & 0 \end{bmatrix}$$

and

$$R'_{2n}(t_1,\ldots,t_n) = \begin{bmatrix} R'(t_1) & O & \cdots & \cdots & \cdots & O \\ O & R'(t_2) & O & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ O & \cdots & \cdots & \cdots & O & R'(t_n) \end{bmatrix} \in \mathfrak{so}(2n),$$

$$R'_{2n+1}(t_1,\ldots,t_n) = \begin{bmatrix} R'(t_1) & O & \cdots & \cdots & \cdots & \cdots & O \\ O & R'(t_2) & O & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & O & R'(t_n) & O \\ O & \cdots & \cdots & \cdots & \cdots & O & 1 \end{bmatrix} \in \mathfrak{so}(2n+1).$$

THEOREM 7.12 (Principle Axis Theorem for Lie algebras). *For each of the following Lie algebras, every element $x \in \mathfrak{g}$ is conjugate in G to one of the stated form.*

- $\mathfrak{so}(2n)$: $R'_{2n}(t_1,\ldots,t_n)$, $\forall k\ \theta_k \in [0,2\pi)$;
- $\mathfrak{so}(2n+1)$: $R'_{2n+1}(t_1,\ldots,t_n)$, $\forall k\ \theta_k \in [0,2\pi)$;
- $\mathfrak{u}(n)$: $\mathrm{diag}(t_1 i,\ldots,t_n i)$, $\forall k\ t_k \in \mathbb{R}$;
- $\mathfrak{su}(n)$: $\mathrm{diag}(t_1 i,\ldots,t_n i)$, $\forall k\ t_k \in \mathbb{R}$, $t_1+\cdots+t_n=1$;
- $\mathfrak{sp}(n)$: $\mathrm{diag}(t_1 i,\ldots,t_n i)$, $\forall k\ t_k \in \mathbb{R}$.

We can now give an important result which we have already seen is true for many familiar examples.

THEOREM 7.13. *Let G be a compact, connected Lie group. Then the exponential map $\exp\colon \mathfrak{g} \longrightarrow G$ is surjective.*

PROOF. Let $T \leqslant G$ be a maximal torus. By Theorem 7.8, every element $g \in G$ is conjugate to an element $xgx^{-1} \in T$. By Proposition 7.5, $xgx^{-1} = \exp(t)$ for some $t \in \mathfrak{t}$, hence

$$g = x^{-1}\exp(t)x = \exp(\mathrm{Ad}_x(t)),$$

where $\mathrm{Ad}_x(t) \in \mathfrak{g}$. So $g \in \exp\mathfrak{g}$. Therefore $\exp\mathfrak{g} = G$. □

3. The normalizer and Weyl group of a maximal torus

Given Theorem 7.8, we can continue to develop the general theory for a compact connected Lie group G.

PROPOSITION 7.14. *Let $A \leqslant G$ be a compact abelian Lie group and suppose that $A_1 \leqslant A$ is the connected component of the identity element. If A/A_1 is cyclic then A has a generator and hence A is contained in a torus in G.*

PROOF. Let $d = |A/A_1|$. As A_1 is connected and abelian, it is a torus by Theorem 7.4, hence it has a generator a_0 by Proposition 7.7. Let $g \in A$ be an element of A for which the coset gA_1 generates A/A_1. Notice that $g^d \in A_1$ and therefore $a_0 g^{-d} \in A_1$. Now choose $b \in A_1$ so that $a_0 g^{-d} = b^d$. Then $a_0 = (gb)^d$, so the powers $(gb)^{kd}$ are dense in A_1. More generally, the powers of the from $(gb)^{kd+r}$ are dense in the coset $g^r A_1$. Hence the powers of gb are dense in A, which shows that this element is a generator of A.

Let $T \leqslant G$ be a maximal torus. By Theorem 7.8, any generator u of A is contained in a maximal torus xTx^{-1} conjugate to T. Hence $\langle u\rangle$ and its closure A are contained in xTx^{-1} which completes the proof of the Proposition. □

PROPOSITION 7.15. *Let $A \leqslant G$ be a connected abelian subgroup and let $g \in G$ commute with all the elements of A. Then there is a torus $T \leqslant G$ containing the subgroup $\langle A, g\rangle \leqslant G$ generated by A and g.*

PROOF. By replacing A by its closure which is also connected, we can assume that A is closed in G, hence compact and so a torus, by Theorem 7.4. Now consider the abelian subgroup $\langle A, g\rangle \leqslant G$ generated by A and g, whose closure $B \leqslant G$ is again compact and abelian. If the connected component of the identity is $B_1 \leqslant B$ then B_1 has finitely many cosets by compactness, and these is of the form $g^r B_1$ $(r = 0, 1, \ldots, d-1)$ for some d. By Proposition 7.14, $\langle A, g\rangle$ is contained in a torus. □

THEOREM 7.16. *Let $T \leqslant G$ be a maximal torus and let $T \leqslant A \leqslant G$ where A is abelian. Then $A = T$. Equivalently, every maximal torus is a maximal abelian subgroup.*

PROOF. For each element $g \in A$, Proposition 7.15 implies that there is a torus containing $\langle T, g\rangle$, but by the maximality of T this must equal T. Hence $A = T$. □

We have now established that every maximal torus is also a maximal abelian subgroup, and that any two maximal tori are conjugate in G.

Recall that for a subgroup $H \leqslant G$, the *normalizer* of H in G is

$$\mathrm{N}_G(H) = \{g \in G : gHg^{-1} = H\}.$$

Then $\mathrm{N}_G(H) \leqslant G$ is a closed subgroup of G, hence compact. It also contains H and its closure in G as normal subgroups. There is a continuous left action of $\mathrm{N}_G(H)$ on H by conjugation, i.e., for $g \in \mathrm{N}_G(H)$ and $h \in H$, the action is given by

$$g \cdot h = ghg^{-1}.$$

If $H = T$ is a maximal torus in G, the quotient group $\mathrm{N}_G(T)/T$ acts on T since T acts trivially on itself by conjugation. Notice that the connected component of the identity in $\mathrm{N}_G(T)$ contains T, in fact it agrees with T by the following Lemma.

LEMMA 7.17. *Let $T \leqslant G$ be a torus and let $Q \leqslant \mathrm{N}_G(T)$ be a connected subgroup acting on T by conjugation. Then Q acts trivially, i.e., for $g \in Q$ and $x \in T$,*

$$g \cdot x = gxg^{-1} = x.$$

PROOF. Recall that $T \cong \mathbb{R}^r/\mathbb{Z}^r$ as Lie groups. By Proposition 7.5, the exponential map is a surjective group homomorphism $\exp\colon \mathfrak{t} \longrightarrow T$ whose kernel is a discrete subgroup. In fact, there is a commutative diagram

$$\begin{array}{ccccc} \ker\exp & \longrightarrow & \mathfrak{t} & \longrightarrow & T \\ \cong\big\downarrow & & \cong\big\downarrow & & \cong\big\downarrow \\ \mathbb{Z}^r & \longrightarrow & \mathbb{R}^r & \longrightarrow & \mathbb{R}^r/\mathbb{Z}^r \end{array}$$

in which all the maps are the evident ones.

Now a Lie group automorphism $\alpha\colon T \longrightarrow T$ lifts to homomorphism $\widetilde{\alpha}\colon \mathfrak{t} \longrightarrow \mathfrak{t}$ restricting to an isomorphism $\widetilde{\alpha}_0\colon \ker\exp \longrightarrow \ker\exp$. Indeed, since each element of $\ker\exp \cong \mathbb{Z}^r$ is uniquely divisible in $\mathfrak{t} \cong \mathbb{R}^r$, continuity implies that $\widetilde{\alpha}_0$ determines $\widetilde{\alpha}$ on $\mathfrak{t}$. But the automorphism group $\mathrm{Aut}(\ker\exp) \cong \mathrm{Aut}(\mathbb{Z}^r)$ of $\ker\exp \cong \mathbb{Z}^r$ is a discrete group.

From this we see that the action of Q on T by conjugation is determined by its restriction to the action on $\ker\exp$. As Q is connected, every element of Q gives rise to the identity automorphism of the discrete group $\mathrm{Aut}(\ker\exp)$. Hence the action of Q on T is trivial. □

This result shows that $\mathrm{N}_G(T)_1$, the connected component of the identity in $\mathrm{N}_G(T)$, acts trivially on the torus T. In fact, if $g \in \mathrm{N}_G(T)$ acts trivially on T then it commutes with all the elements of T, so by Theorem 7.16 g is in T. Thus T consists of all the elements of G with this property, i.e.,

$$T = \{g \in G : gxg^{-1} = x \ \forall x \in T\}. \tag{7.1}$$

In particular, we have $\mathrm{N}_G(T)_1 = T$.

The *Weyl group* of the maximal torus T in G is the quotient group

$$\mathrm{W}_G(T) = \mathrm{N}_G(T)/T = \pi_0 \, \mathrm{N}_G(T)$$

which is also the group of path components $\pi_0 \, \mathrm{N}_G(T)$. The Weyl group $\mathrm{W}_G(T)$ acts on T by conjugation, i.e., according to the formula

$$gT \cdot x = gxg^{-1}.$$

THEOREM 7.18. *Let $T \leqslant G$ be a maximal torus. Then the Weyl group $\mathrm{W}_G(T)$ is finite and acts faithfully on T, i.e., the coset $gT \in \mathrm{N}_G(T)/T$ acts trivially on T if and only if $g \in T$.*

PROOF. $\mathrm{N}_G(T)$ has finitely many cosets of T since it is closed, hence compact, so each coset is clopen. The faithfulness of the action follows from Equation (7.1). $\square$

PROPOSITION 7.19. *Let $T \leqslant G$ be a maximal torus and $x, y \in T$. If x, y are conjugate in G then they are conjugate in $\mathrm{N}_G(T)$, hence there is an element $w \in \mathrm{W}_G(T)$ for which $y = w \cdot x$.*

PROOF. Suppose that $y = gxg^{-1}$. Then the centralizer $\mathrm{C}_G(y) \leqslant G$ of y is a closed subgroup containing T. It also contains the maximal torus gTg^{-1} since every element of this commutes with y. Let $H = \mathrm{C}_G(y)_1$, the connected component of the identity in $\mathrm{C}_G(y)$; this is a closed subgroup of G since it is closed in $\mathrm{C}_G(y)$. Then as T, gTg^{-1} are connected subgroups of $\mathrm{C}_G(y)$ they are both contained in H. So T, gTg^{-1} are tori in H and must be maximal since a torus in H containing one of these would be a torus in G where they are already maximal.

By Theorem 7.8 applied to the compact connected Lie group H, gTg^{-1} is conjugate to T in H, so for some $h \in H$ we have $gTg^{-1} = hTh^{-1}$ which gives

$$(h^{-1}g)T(h^{-1}g)^{-1} = T.$$

Thus $h^{-1}g \in \mathrm{N}_G(T)$ and

$$(h^{-1}g)x(h^{-1}g)^{-1} = h^{-1}yh = y.$$

Now setting $w = h^{-1}gT \in \mathrm{W}_G(T)$ we obtain the desired result. $\square$

Bibliography

[1] J. F. Adams, Lectures on Lie Groups, University of Chicago Press (1969).

[2] J. F. Adams, Lectures on Exceptional Lie Groups, University of Chicago Press (1996).

[3] R. Carter, G. Segal, I. Macdonald, Lectures on Lie groups and Lie algebras, Cambridge University Press, (1995).

[4] M. L. Curtis, Matrix Groups, Springer-Verlag (1984).

[5] R. Howe, Very basic Lie theory, Amer. Math. Monthly **90** (1983) 600–623; correction: Amer. Math. Monthly **91** (1984) 247.

[6] I. R. Porteous, Topological geometry, Van Nostrand Reinhold Co. (1969).

[7] J-P. Serre, Complex Semisimple Lie Algebras, Springer-Verlag (1987).

[8] S. Sternberg, Group Theory and Physics, Cambridge University Press (1994).

[9] F. W. Warner, Foundations of Differentiable Manifolds and Lie Groups, Springer-Verlag (1983).

Index

$\Bbbk$-lines, 76
$\pm$-grading, 48
n-manifold, 55

adjoint action, 28, 59
affine group, 7
affine transformation, 7
algebra, 37
algebra homomorphism, 37
anti-homomorphism of rings, 43
atlas, 55

Borel subgroup, 7

canonical automorphism, 47
canonical commutation relation, 71
Cauchy sequence, 4
central subfield, 38
centre, 52
chart, 55
circle group, 89
Clifford algebra, 45
clopen, 6
commutator, 20
commute, 20
compact, 6
compact matrix group, 6
complete, 4
complex dimension, 20
complex matrix group, 10
conjugate, 92
conjugate in G, 92
conjugation map, 58
connected, 81
connected by a path, 84
continuous group action, 12
continuous homomorphism of matrix groups, 11
continuous isomorphism of matrix groups, 11
continuous map, 2
coordinate function, 3

derivation property, 23
derivative, 22
diffeomorphism, 55
differentiable, 56
differentiable curve, 19, 56
differentiable homomorphism, 22
differentiable map, 21
dimension, 20
dimension, complex, 20
direct isometry, 53
direct product, 37
division algebra, 37

elementary 2-group,, 54
Euler characteristic, 92
exponential function, 14
extraspecial 2-group, 54

general linear group, 1
generalized permutation matrix, 46
generator, 91
Grassmannian, 79
group action, 12
group of path components, 85

Heisenberg algebra, 71
Heisenberg group, 69
Heisenberg Lie algebra, 71
hermitian conjugate, 9
homogeneous space, 74
homomorphism of algebras, 37
homomorphism of Lie algebras, 22
homomorphism of skew rings, 43

indirect isometry, 53
invariant, 29
isometry, 8
isomorphism of algebras, 37
isomorphism of Lie algebras, 28

Jacobi identity, 20
Jacobian matrix, 56
Jordan Canonical Form, 25

Kronecker symbol, 8

Lefschetz Fixed Point Theorem, 92
left multiplication map, 58
left regular representation, 38
Lie algebra, 20
Lie bracket, 20
Lie group, 58

Lie homomorphism, 22
Lie subalgebra, 21
Lie subgroup, 58
local trivialisation, 74
locally path connected, 81
logarithmic function, 14
Lorentz group, 9
lower triangular matrices, 32

manifold, 55
matrix group, 5
matrix subgroup, 5
maximal torus, 89
maximal torus in G, 89

norm, 1
normalizer, 94

one parameter group, 18
one parameter semigroup, 18
one parameter subgroup, 18
open disc, 2
operator norm, 2
orbit, 12
orthogonal, 30
orthogonal group, 8
orthogonal matrix, 8

parabolic subgroup, 79
path component, 84
path connected, 81
permutation matrix, 53
pinor group, 49
positive definite, 29
principal H-bundle, 74
product topology, 13
projective quaternionic symplectic group, 77
projective special unitary group, 77
projective unitary group, 77
pure quaternion, 43

quaternionic conjugate, 44
quaternionic conjugation, 43
quaternionic symplectic group, 44
quaternions, 41, 42
quotient topology, 12, 73

real Clifford algebras, 45
real inner product, 28
real part, 48
real quaternion, 43
reduced determinant, 41
right multiplication map, 58

simple algebra, 38
skew field, 37
skew symmetric, 9
smooth, 56
smooth n-manifold, 55
smooth curve, 56
smooth manifold, 55
smooth map, 55
special linear group, 1
special orthogonal group, 8
special unitary group, 9
spinor group, 49
stabilizer, 12
standard torus, 89
strictly upper triangular, 26
submanifold, 57
supremum norm, 2
symmetric bilinear form, 29
symplectic geometry, 9
symplectic group, 9

tangent space, 19, 56
topological generator, 71, 91
topological group, 5
torus, 89
torus of rank r, 89
torus, maximal, 89
translation subgroup, 7

unimodular group, 1
unipotent, 7
unitary group, 9
Universal Property of Clifford algebras, 46
upper triangular, 7

Weyl group, 95

Problem sets

Throughout these problem sets, $\Bbbk$ will denote one of the fields $\mathbb{R}, \mathbb{C}$ and treat vectors in $\Bbbk^n$ as column vectors. All other notation follows the notes.

Problems on Chapter 1

1-1. Determine $\|A\|$ for each of the following matrices A, where $t, u, v \in \mathbb{R}$.

$$\begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix}, \quad \begin{bmatrix} u & 1 \\ 0 & u \end{bmatrix}, \quad \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}, \quad \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}.$$

What can be said when $u, v \in \mathbb{C}$?

1-2. Let $A \in \mathrm{M}_n(\mathbb{C})$.
a) If $B \in \mathrm{U}(n)$, show that $\|BAB^{-1}\| = \|A\|$.
b) For a general element $C \in \mathrm{GL}_n(C)$, what can be said about $\|CAC^{-1}\|$?

1-3. [*This problem relates to Remark 1.3*] Let $A \in \mathrm{M}_n(\mathbb{C})$.
a) Show that $\|A\|$ satisfies

$$\|A\|^2 = \sup\{\mathbf{x}^* A^* A\mathbf{x} : \mathbf{x} \in \mathbb{C}^n,\ |\mathbf{x}| = 1\} = \max\{\mathbf{x}^* A^* A\mathbf{x} : \mathbf{x} \in \mathbb{C}^n,\ |\mathbf{x}| = 1\}.$$

b) Show that the eigenvalues of A^*A are non-negative real numbers. Deduce that if $\lambda \in \mathbb{R}$ is the largest eigenvalue of A^*A then $\|A\| = \sqrt{\lambda}$ and for any unit eigenvector $\mathbf{v} \in \mathbb{C}^n$ of A^*A for the eigenvalue λ, $\|A\| = |A\mathbf{v}|$.
c) When A is real, show that $\|A\| = |A\mathbf{w}|$ for some unit vector $\mathbf{w} \in \mathbb{R}^n$.

1-4. If $\{A_r\}_{r\geqslant 0}$ is a sequence of matrices $A_r \in \mathrm{M}_n(\Bbbk)$, prove the following.

a) If $\displaystyle\lim_{r\to\infty} \frac{\|A_{r+1}\|}{\|A_r\|} < 1$, the series $\sum_{r=0}^{\infty} A_r$ converges in $\mathrm{M}_n(\Bbbk)$.

b) If $\displaystyle\lim_{r\to\infty} \frac{\|A_{r+1}\|}{\|A_r\|} > 1$, the series $\sum_{r=0}^{\infty} A_r$ diverges in $\mathrm{M}_n(\Bbbk)$.

c) Develop other convergence tests for $\sum_{r=0}^{\infty} A_r$.

1-5. Suppose that $A \in \mathrm{M}_n(\Bbbk)$ and $\|A\| < 1$.
a) Show that the series

$$\sum_{r=0}^{\infty} A^r = I + A + A^2 + A^3 + \cdots$$

converges in $\mathrm{M}_n(\Bbbk)$.
b) Show that $(I - A)$ is invertible and give a formula for $(I - A)^{-1}$.
c) If A nilpotent (i.e., $A^k = O$ for k large) determine $(I - A)^{-1}$ and $\exp(A)$.

1-6. a) Show that the set of all $n \times n$ real orthogonal matrices $\mathrm{O}(n) \subseteq \mathrm{M}_n(\mathbb{R})$ is compact.
b) Show that the set of all $n \times n$ unitary matrices $\mathrm{U}(n) \subseteq \mathrm{M}_n(\mathbb{C})$ is compact.

c) Show that $\mathrm{GL}_n(\Bbbk)$ and $\mathrm{SL}_n(\Bbbk)$ are not compact if $n \geqslant 2$.

d) Investigate which of the other matrix groups of Section 4 are compact.

1-7. Using Example 1.21, for $n \geqslant 1$ show that

a) $\mathrm{O}(n)$ is a matrix subgroup of $\mathrm{O}(n+1)$;

b) $\mathrm{SO}(n)$ is a matrix subgroup of $\mathrm{SO}(n+1)$;

c) $\mathrm{U}(n)$ is a matrix subgroup of $\mathrm{U}(n+1)$;

d) $\mathrm{SU}(n)$ is a matrix subgroup of $\mathrm{SU}(n+1)$.

1-8. For $t \in \mathbb{R}$, determine the matrices

$$\exp\left(\begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix}\right), \quad \exp\left(\begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix}\right).$$

1-9. Let $\Bbbk = \mathbb{R}, \mathbb{C}$, and $A \in \mathrm{M}_n(\Bbbk)$.

a) Show that for $B \in \mathrm{GL}_n(\Bbbk)$,

$$\exp(BAB^{-1}) = B\exp(A)B^{-1}.$$

b) If A is diagonalisable, say with $A = C\operatorname{diag}(\lambda_1, \ldots, \lambda_n)C^{-1}$, for $C \in \mathrm{GL}_n(\Bbbk)$, determine $\exp(A)$.

c) Use this to find the matrices

$$\exp\left(\begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix}\right), \quad \exp\left(\begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix}\right).$$

1-10. If $S \in \mathrm{M}_n(\mathbb{R})$ be skew symmetric (i.e., $S^T = -S$), show that $\exp(S)$ is orthogonal, i.e., $\exp(S)^T = \exp(S)^{-1}$.

More generally, if $S \in \mathrm{M}_n(\mathbb{C})$ is skew hermitian (i.e., $S^* = -S$), show that $\exp(S)$ is unitary, i.e., $\exp(S)^* = \exp(S)^{-1}$.

1-11. Let

$$G = \{A \in \mathrm{GL}_n(\mathbb{R}) : \det A \in \mathbb{Q}\} \leqslant \mathrm{GL}_n(\mathbb{R}).$$

a) Show that G not a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$.

b) Find the closure $\overline{G}$ of G in $\mathrm{GL}_n(\mathbb{R})$.

1-12. For $\Bbbk = \mathbb{R}, \mathbb{C}$ and $n \geqslant 1$, let $N \in \mathrm{M}_n(\Bbbk)$.

a) If N is strictly upper triangular, show that $\exp(N)$ is unipotent.

b) Determine $\exp(N)$ when N is an arbitrary upper triangular matrix.

The next two problems relate to Section 7.

1-13. Let (X_1, ρ_1), (X_2, ρ_2) be two metric spaces. Define a function

$$\rho\colon (X_1 \times X_2) \times (X_1 \times X_2) \longrightarrow \mathbb{R}^+; \quad \rho((x_1, x_2), (y_1, y_2)) = \sqrt{\rho_1(x_1, y_1)^2 + \rho_2(x_2, y_2)^2}.$$

a) Show that $(X_1 \times X_2, \rho)$ is a metric space.

b) Show that a sequence $\{(x_{1,r}, x_{2,r})\}_{r\geqslant 0}$ converges (i.e., has a limit) in $X_1 \times X_2$ if and only if the sequences $\{(x_{1,r})\}_{r\geqslant 0}$, $\{(x_{2,r})\}_{r\geqslant 0}$ converge in X_1 and X_2 respectively.

1-14. a) Using the previous question, define a metric on $\mathrm{M}_n(\Bbbk) \times \Bbbk^n$ and show that the product map

$$\varphi\colon \mathrm{M}_n(\Bbbk) \times \Bbbk^n \longrightarrow \Bbbk^n; \quad \varphi(A, \mathbf{x}) = A\mathbf{x},$$

is continuous.

b) Let $G \leqslant \mathrm{GL}_n(\Bbbk)$ be a matrix subgroup. By restricting the metric and product φ of (a) to the subset $G \times \Bbbk^n$, consider the resulting continuous group action of G on $\Bbbk^n$. Show that the stabilizer of $\mathbf{x} \in \Bbbk^n$,

$$\mathrm{Stab}_G(\mathbf{x}) = \{A \in G : A\mathbf{x} = \mathbf{x}\}$$

is a matrix subgroup of G. More generally, if $X \subseteq \Bbbk^n$ is a closed subset, show that

$$\mathrm{Stab}_G(X) = \{A \in G : AX = X\}$$

is a matrix subgroup of G, where $AX = \{A\mathbf{x} : \mathbf{x} \in X\}$.

c) For the standard basis vector $\mathbf{e}_n = [0, \cdots, 0, 1]^T$ and $X = \{t\mathbf{e}_n : t \in \mathbb{R}\}$, determine $\mathrm{Stab}_G(\mathbf{e}_n)$ and $\mathrm{Stab}_G(X)$ for each of the following matrix subgroups $G \leqslant \mathrm{GL}_n(\mathbb{R})$:

$$G = \mathrm{GL}_n(\mathbb{R}), \quad G = \mathrm{SL}_n(\mathbb{R}), \quad G = \mathrm{O}(n), \quad G = \mathrm{SO}(n).$$

Problems on Chapter 2

2-1. a) Solve the differential equation

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

by finding a solution of the form

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \alpha(t) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

with $\alpha\colon \mathbb{R} \longrightarrow \mathrm{GL}_2(\mathbb{R})$. Sketch the trajectory of this solution as a curve in the xy-plane. What happens for other initial values $x(0), y(0)$?

b) Repeat this with the equations

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix};$$

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

2-2. Let G be a matrix group and $U \in G$.

a) Show that each the functions

$$\begin{aligned} \mathrm{L}_U\colon G \longrightarrow G; &\quad \mathrm{L}_U(A) = UA, \\ \mathrm{R}_U\colon G \longrightarrow G; &\quad \mathrm{R}_U(A) = AU, \\ \mathrm{C}_U\colon G \longrightarrow G; &\quad \mathrm{C}_U(A) = UAU^{-1}, \end{aligned}$$

is a differentiable map and determine its derivative at I.

b) Using (a), show that there $\mathbb{R}$-linear isomorphisms

$$\lambda_U\colon \mathrm{T}_I\, G \longrightarrow \mathrm{T}_U\, G, \quad \rho_U\colon \mathrm{T}_I\, G \longrightarrow \mathrm{T}_U\, G, \quad \chi_U\colon \mathrm{T}_I\, G \longrightarrow \mathrm{T}_I\, G,$$

such that for all $U, V \in G$,

$$\lambda_{UV} = \lambda_U \circ \lambda_V, \quad \rho_{UV} = \rho_V \circ \rho_U, \quad \chi_{UV} = \chi_U \circ \chi_V.$$

2-3. For each of the following matrix groups G determine its Lie algebra $\mathfrak{g}$.

(a) $$G = \{A \in \mathrm{GL}_2(\mathbb{R}) : AQA^T = Q\}, \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix};$$

(b) $$G = \{A \in \mathrm{GL}_2(\mathbb{R}) : AQA^T = Q\}, \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix};$$

(c) $$G = \{A \in \mathrm{GL}_3(\mathbb{R}) : AQA^T = Q\}, \quad Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix};$$

(d) $$G = \mathrm{Aff}_n(\Bbbk);$$

(e) $$G = \mathrm{Symp}_{2m}(\mathbb{R}), \quad (m = 1, 2, \ldots).$$

2-4. Consider the set of all $n \times n$ real special orthogonal matrices $\mathrm{SO}(n)$ and its subset

$$U = \{A \in \mathrm{SO}(n) : \det(I + A) \neq 0\} \subseteq \mathrm{SO}(n).$$

Define the function

$$\Phi\colon U \longrightarrow \mathrm{M}_n(\mathbb{R}); \quad \Phi(A) = (I - A)(I + A)^{-1}.$$

a) Show that $\operatorname{im}\Phi = \text{Sk-Sym}_n(\mathbb{R})$, the set of all $n \times n$ real skew symmetric matrices. Hence we might as well write $\Phi\colon U \longrightarrow \text{Sk-Sym}_n(\mathbb{R})$.
b) Find the inverse map $\Phi^{-1}\colon \text{Sk-Sym}_n(\mathbb{R}) \longrightarrow U$.
c) Use (b) to determine the dimension of $\mathrm{SO}(n)$.

[Φ **is the *real Cayley transform.***]

2-5. Consider the set of all $n \times n$ unitary matrices $\mathrm{U}(n)$ and its subset

$$V = \{A \in \mathrm{U}(n) : \det(I + A) \neq 0\} \subseteq \mathrm{U}(n).$$

Define the function

$$\Theta\colon V \longrightarrow \mathrm{M}_n(\mathbb{C}); \quad \Theta(A) = (I - A)(I + A)^{-1}.$$

a) Show that $\operatorname{im}\Theta = \text{Sk-Herm}_n(\mathbb{C})$, the set of all $n \times n$ skew hermitian matrices. Hence we might as well write $\Theta\colon V \longrightarrow \text{Sk-Herm}_n(\mathbb{C})$.
b) Find the inverse map $\Theta^{-1}\colon \text{Sk-Herm}_n(\mathbb{C}) \longrightarrow V$.
c) Use (b) to determine the dimension of $\mathrm{U}(n)$.
d) For the case $n = 2$ show that $\Theta(V \cap \mathrm{SU}(2)) \subseteq \text{Sk-Herm}^0_2(\mathbb{C})$ and $\Theta^{-1}\,\text{Sk-Herm}^0_2(\mathbb{C}) \subseteq \mathrm{SU}(2)$. Is this true for $n > 2$?

[Φ **is the *complex Cayley transform.***]

Problems on Chapter 3

3-1. Using the bases $\{1, i, j, k\}$ of $\mathbb{H}_{\mathbb{R}}$ over $\mathbb{R}$ and $\{1, j\}$ of $\mathbb{H}_{\mathbb{C}}$ over $\mathbb{C}$, determine the reduced determinants $\mathrm{Rdet}_{\mathbb{R}} \colon \mathrm{GL}_n(\mathbb{H}) \longrightarrow \mathbb{R}^{\times}$ and $\mathrm{Rdet}_{\mathbb{C}} \colon \mathrm{GL}_n(\mathbb{H}) \longrightarrow \mathbb{C}^{\times}$.

3-2. i) Verify that $\mathrm{M}_n(\mathbb{H})$ is complete with respect to the norm $\| \ \|$. Using this, explain how to define an exponential function $\exp\colon \mathrm{M}_n(\mathbb{H}) \longrightarrow \mathrm{GL}_n(\mathbb{H})$ with properties analogous to those for the exponential functions on $\mathrm{M}_n(\mathbb{R}), \mathrm{M}_n(\mathbb{C})$.
ii) When $n = 1$, determine $\exp(q)$ using the decomposition $q = r + su$ with $r, s \in \mathbb{R}$ and u a pure quaternion of unit length $|u| = 1$.

3-3. For each of the following matrix groups G, determine the Lie algebra $\mathfrak{g}$ and dimension $\dim G$:
i) $G = \mathrm{GL}_n(\mathbb{H})$;
ii) $G = \mathrm{Sp}_n(\mathbb{H})$;
iii) $G = \ker \mathrm{Rdet}_{\Bbbk} \colon \mathrm{GL}_n(\mathbb{H}) \longrightarrow \Bbbk^{\times}$ where $\Bbbk = \mathbb{R}, \mathbb{C}$;
iv) $G = \ker \mathrm{Rdet}_{\Bbbk} \colon \mathrm{Sp}(n) \longrightarrow \Bbbk^{\times}$ where $\Bbbk = \mathbb{R}, \mathbb{C}$.

3-4. The group of unit quaternions

$$\mathrm{Sp}(1) = \{q \in \mathbb{H} : |q| = 1\}$$

has an $\mathbb{R}$-linear action on $\mathbb{H}$ given by

$$q \cdot x = qxq^{-1} = qx\overline{q} \quad (x \in \mathbb{H}).$$

i) By identifying $\mathbb{H}$ with $\mathbb{R}^4$ using the basis $\{i, j, k, 1\}$, show that this defines a Lie homomorphism $\mathrm{Sp}(1) \longrightarrow \mathrm{SO}(4)$.
ii) Show that this action restricts to an action of $\mathrm{Sp}(1)$ on the space of pure quaternions and by identifying this with $\mathbb{R}^3$ using the basis $\{i, j, k\}$, show that this defines a surjective Lie homomorphism $\alpha\colon \mathrm{Sp}(1) \longrightarrow \mathrm{SO}(3)$. Show that the kernel of α is $\{1, -1\}$.

3-5. Using the surjective homomorphism $\alpha\colon \mathrm{Sp}(1) \longrightarrow \mathrm{SO}(3)$ of the previous question, for a subgroup $G \leqslant \mathrm{SO}(3)$ set

$$\widetilde{G} = \alpha^{-1}G = \{g \in \mathrm{Sp}(1) : \alpha(g) \in G\} \leqslant \mathrm{Sp}(1).$$

From now on assume that G is *finite.*
i) What is the order of $\widetilde{G}$?
ii) Show that the order of the centre of $\widetilde{G}$, $\mathrm{C}(\widetilde{G})$, is even.
iii) If G contains an element of order 2, show that the group homomorphism $\alpha\colon \widetilde{G} \longrightarrow G$ is not *split* in the sense that there is no group homomorphism $\beta\colon \widetilde{G} \longrightarrow G$ for which $\alpha \circ \beta = \mathrm{Id}_G$.

$$\begin{tikzcd} G \arrow[r, "\beta"] \arrow[dr, "\mathrm{Id}_G"'] & \widetilde{G} \arrow[d, "\alpha"] \\ & G \end{tikzcd}$$

iv) Show that $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$ is a subgroup of $\mathrm{Sp}(1)$ and find a geometric interpretation as a group of symmetries for $\alpha Q_8 \leqslant \mathrm{SO}(3)$. Generalize this by considering for each $n \geqslant 2$,

$$Q_{2n} = \left\{ e^{2\pi i r/n} : r = 0, \ldots, n-1 \right\} \cup \left\{ e^{2\pi i r/n} j : r = 0, \ldots, n-1 \right\}.$$

v) Show that the set T_{24} consisting of the 24 elements

$$\pm 1, \quad \pm i, \quad \pm j, \quad \pm k, \quad \frac{1}{2}(\pm 1 \pm i \pm j \pm k)$$

is a subgroup of $\mathrm{Sp}(1)$ and find a geometric interpretation for the group $\alpha T_{24} \leqslant \mathrm{SO}(3)$.
vi) [Challenge question: not for the fainthearted!] Let **Icos** be a regular icosahedron in $\mathbb{R}^3$ centred at the origin. The group of direct symmetries of **Icos** is known to be isomorphic to the alternating group, $\mathrm{Symm}^+(\mathbf{Icos}) \cong A_5$. Find $\alpha^{-1}\,\mathrm{Symm}^+(\mathbf{Icos}) \leqslant \mathrm{Sp}(1)$.

This requires a good way to view the icosahedron relative to the x, y, z-axes. Nice graphics and information on the icosahedron can be found at

`http://mathworld.wolfram.com/Icosahedron.html`

The resulting subgroup of $\mathrm{Sp}(1)$ *is called the* binary icosahedral group *since it double covers the symmetry* $\mathrm{Symm}^+(\mathbf{Icos})$*; it also provides a non-split double covering $\widetilde{A}_5 \longrightarrow A_5$ of the simple group A_5.*

3-6. In the Clifford algebra Cl_n, let $u, v \in \mathbb{R}^n \subseteq \mathrm{Cl}_n$.
i) If $|u| = 1$, by expressing v as a sum $v_1 + v_2$ with $v_1 = tu$ and $u \cdot v_2 = 0$, find a general formula for uvu.
ii) If $\{u_1, \ldots, u_n\}$ is any orthonormal basis for $\mathbb{R}^n$, show that

$$u_j u_i = \begin{cases} -1 & \text{if } j = i, \\ -u_i u_j & \text{if } j \neq i. \end{cases}$$

Deduce that every element $A \in \mathrm{O}(n)$ induces an automorphism $A_* \colon \mathrm{Cl}_n \longrightarrow \mathrm{Cl}_n$ for which $A_* x = Ax$ if $x \in \mathbb{R}^n$.

3-7. In the following, use the Universal Property of Theorem 3.23.
i) Show that the natural embedding

$$i_n \colon \mathbb{R}^n \longrightarrow \mathbb{R}^{n+1}; \qquad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \longmapsto \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 0 \end{bmatrix},$$

induces an injective $\mathbb{R}$-algebra homomorphism $i'_n \colon \mathrm{Cl}_n \longrightarrow \mathrm{Cl}_{n+1}$ for which $i'_n(x) = i_n(x)$ whenever $x \in \mathbb{R}^n$. Determine the image $\operatorname{im} i'_n \subseteq \mathrm{Cl}_{n+1}$.
ii) Show that the $\mathbb{R}$-linear transformation

$$k_n \colon \mathbb{R}^n \longrightarrow \mathrm{Cl}_{n+1}; \quad k_n(x) = x e_{n+1},$$

induces an injective $\mathbb{R}$-algebra homomorphism $k'_n \colon \mathrm{Cl}_n \longrightarrow \mathrm{Cl}_{n+1}$ for which $k'_n(x) = k_n(x)$ whenever $x \in \mathbb{R}^n$. Show that $\operatorname{im} k'_n = \mathrm{Cl}^+_{n+1}$.

Problems on Chapter 4

4-1. Show that the subset

$$M = \{(A, b) \in \mathrm{M}_n(\mathbb{R}) \times \mathbb{R} : b \det A = 1\} \subseteq \mathrm{M}_n(\mathbb{R}) \times \mathbb{R}$$

is a closed submanifold and determine $\mathrm{T}_{(A,b)} M$ for $(A, b) \in M$.

Show that M has the structure of a Lie group with multiplication μ given by

$$\mu((A_1, b_1), (A_2, b_2)) = (A_1 A_2, b_1 b_2).$$

To which standard matrix group is M isomorphic?

Repeat this with $\mathbb{R}$ replaced by $\mathbb{C}$.

4-2. Write out the details of the calculation in Example 4.21 for the cases $n = 2, 3$.

4-3. Modify the details of Example 4.21 to show that $\mathrm{U}(n) \leqslant \mathrm{GL}_n(\mathbb{C})$ is a Lie subgroup. It might be helpful to do the cases $n = 1, 2, 3$ first.

Use the determinant function

$$\det\colon \mathrm{U}(n) \longrightarrow \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$$

together with the Identity Check Trick 4.20, to show that $\mathrm{SU}(n) \leqslant \mathrm{U}(n)$ is a Lie subgroup.

4-4. Let G be a matrix group. Use Theorem 4.23 to show that each of the following subgroups of G is a Lie subgroup. In each case, try to find a proof that works when G is an arbitrary Lie group.

a) For $g \in G$, the centralizer of g, $\mathrm{C}_G(g) = \{x \in G : xgx^{-1} = g\}$.

b) The centre of G, $\mathrm{C}(G) = \bigcap_{g \in G} \mathrm{C}_G(g)$.

c) For $H \leqslant G$ a closed subgroup, the normalizer of H, $\mathrm{N}_G(g) = \{x \in G : xHx^{-1} = H\}$.

d) The kernel of φ, $\ker \varphi$, where $\varphi\colon G \longrightarrow H$ is a continuous homomorphism into a matrix group H.

4-5. Let G be a matrix group and M a smooth manifold. Suppose that $\mu\colon G \times M \longrightarrow M$ be a continuous group action as defined in Chapter 1 Section 7 and investigated in the Problems on Chapter 1. Also suppose that μ is smooth.

a) Show that for each $x \in M$, $\mathrm{Stab}_G(x) \leqslant G$ is a Lie subgroup.

b) If $X \subseteq M$ is a closed subset, show that $\mathrm{Stab}_G(X) = \{g \in G : gX = X\} \leqslant G$ is a Lie subgroup.

4-6. For a Lie group G and a closed subgroup $H \leqslant G$, show that the cosets gH, Hg and conjugate gHg^{-1} are submanifolds of G. In each case, identify the the tangent space at a point in terms of a suitable tangent space to H.

4-7. Let G and H be Lie groups and $\varphi\colon G \longrightarrow H$ a Lie homomorphism. Show that $\ker \varphi \leqslant G$ is a Lie subgroup and identify the tangent space $\mathrm{T}_g \ker \varphi$ at $g \in \ker \varphi$.

4-8. Determine the Lie bracket $[\ ,\ ]$ for the Lie algebra $\mathfrak{heis}_4$ of the Heisenberg group Heis_4.

Problems on Chapters 5 and 6

5-1. Let $\mathrm{SL}_n(\mathbb{R})$ act smoothly on $\mathbb{R}^n$ by matrix multiplication.

a) Find $\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$ and $\mathrm{Orb}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$.

b) Identify the homogeneous space $\mathrm{SL}_n(\mathbb{R})/\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$ and show that it is path connected if $n \geqslant 2$. Use this to give another proof that $\mathrm{GL}_n(\mathbb{R})$ has two path components.

5-2. Let $\mathrm{SL}_n(\mathbb{C})$ act smoothly on $\mathbb{C}^n$ by matrix multiplication.

a) Find $\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{C})}(\mathbf{e}_n)$ and $\mathrm{Orb}_{\mathrm{SL}_n(\mathbb{C})}(\mathbf{e}_n)$.

b) Identify the homogeneous space $\mathrm{SL}_n(\mathbb{C})/\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{C})}(\mathbf{e}_n)$ and show that it is path connected. Use this to prove that $\mathrm{SL}_n(\mathbb{C})$ is path connected.

c) By making use of the determinant $\det\colon \mathrm{GL}_n(\mathbb{C}) \longrightarrow \mathbb{C}^\times$, deduce that $\mathrm{GL}_n(\mathbb{C})$ is path connected.

5-3. Let $A \in \mathrm{GL}_n(\mathbb{R})$.

a) Show that the symmetric matrix $S = AA^T$ is positive definite, i.e., its eigenvalues are all positive real numbers. Deduce that S has a positive definite real symmetric square root, i.e., there is a positive definite real symmetric matrix S_1 satisfying $S_1^2 = S$.

b) Show that $S_1^{-1}A$ is orthogonal.

c) If $PR = QS$ where P, Q are positive definite real symmetric and $R, S \in \mathrm{O}(n)$, show that $P^2 = Q^2$.

d) Let S_2 be a positive definite real symmetric matrix which satisfies $S_2^2 = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Show that $S_2 = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$.

e) Show that A can be uniquely expressed as $A = PR$ where P is positive definite real symmetric and $R \in \mathrm{O}(n)$. If $\det A > 0$, show that $R \in \mathrm{SO}(n)$.

f) Show that the homogeneous space $\mathrm{GL}_n^+(\mathbb{R})/\mathrm{SO}(n)$ is path connected. Using Example 6.15, deduce that $\mathrm{GL}_n^+(\mathbb{R})$ is path connected.

g) Let $B \in \mathrm{GL}_n(\mathbb{C})$. By suitably modifying the details of the real case, show that B can be uniquely expressed as $B = QT$ with Q positive definite Hermitian and $T \in \mathrm{U}(n)$. Using Example 6.16, deduce that $\mathrm{GL}_n(\mathbb{C})$ is path connected.

[*These factorizations are known as* polar decompositions *of A and B.*]

5-4. For $\Bbbk = \mathbb{R}, \mathbb{C}$ and $n \geqslant 1$, the affine group $\mathrm{Aff}_n(\Bbbk)$ acts on $\Bbbk^n$ as explained in Chapter 1.

a) Find $\mathrm{Stab}_{\mathrm{Aff}_n(\Bbbk)}(\mathbf{0})$ and $\mathrm{Orb}_{\mathrm{Aff}_n(\Bbbk)}(\mathbf{0})$.

b) Show that the affine group $\mathrm{Aff}_n(\mathbb{R})$ has two path components, while $\mathrm{Aff}_n(\mathbb{C})$ is path connected.

Problems on Chapter 7

7-1. Show that there are exactly two Lie isomorphisms $\mathbb{T} \longrightarrow \mathbb{T}$, but infinitely many Lie isomorphisms $\mathbb{T}^r \longrightarrow \mathbb{T}^r$ when $r \geqslant 2$.

7-2. a) Show that the be the subgroup D consisting of all the diagonal matrices $\mathrm{diag}(\alpha, \beta)$ is a maximal torus of $\mathrm{U}(2)$. Determine $\mathrm{N}_{\mathrm{U}(2)}(D)$, $\pi_0 \, \mathrm{N}_{\mathrm{U}(2)}(D)$ and its action by conjugation on D.
b) Show that the subgroup of diagonal matrices in $\mathrm{SU}(3)$ is a maximal torus and determine its normalizer and group of path components and describe the conjugation action of the latter on this torus.
c) Let $T_2 = \{\mathrm{diag}(u, v) \in \mathrm{Sp}(2) : u, v \in \mathbb{C}\} \leqslant \mathrm{Sp}(2)$. Show that T_2 is a maximal torus of $\mathrm{Sp}(2)$. Using Example 6.21, determine $\mathrm{N}_{\mathrm{Sp}(2)}(T_2)$ and $\pi_0 \, \mathrm{N}_{\mathrm{Sp}(2)}(T_2)$ and describe its conjugation action on T_2.

7-3. Show that the group

$$A = \{(\cos\theta_1 + \sin\theta_1 e_1 e_2)(\cos\theta_2 + \sin\theta_2 e_3 e_4) \cdots (\cos\theta_n + \sin\theta_n e_{2n-1} e_{2n}) : \theta_1, \ldots, \theta_n \in [0, 2\pi)\}$$

is a maximal torus in each of the spinor groups $\mathrm{Spin}(2n)$, $\mathrm{Spin}(2n+1)$.

For small values of n, determine the normalizers and Weyl groups of A in $\mathrm{Spin}(2n)$ and $\mathrm{Spin}(2n+1)$. Find the conjugation action of each Weyl group on A.

Under the double covering maps ρ of Chapter 3, how are these maximal tori related to the maximal tori of $\mathrm{SO}(2n)$ and $\mathrm{SO}(2n+1)$ given by Proposition 7.2?

7-4. a) For $n \geqslant 1$, show that the group

$$\left\{ \begin{bmatrix} \varepsilon_1 & 0 & \cdots & 0 \\ 0 & \varepsilon_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \varepsilon_n \end{bmatrix} : \varepsilon_1, \ldots, \varepsilon_n = \pm 1 \right\}$$

is a maximal abelian subgroup of $\mathrm{O}(n)$.
b) Let $T_{2n} \leqslant \mathrm{SO}(2n)$, $T_{2n+1} \leqslant \mathrm{SO}(2n+1)$ be the maximal tori given by Proposition 7.2 and

$$T'_{2n} = T_{2n}$$

$$T'_{2n+1} = T_{2n+1} \cup \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & -1 \end{bmatrix} T_{2n+1}.$$

Show that $T'_{2n} \leqslant \mathrm{O}(2n)$ and $T'_{2n+1} \leqslant \mathrm{O}(2n+1)$ are maximal abelian subgroups.
c) Explain why these results are compatible with those of Chapter 7.