

Advisory Board

P.J. Cameron *Queen Mary and Westfield College*
M.A.J. Chaplain *University of Dundee*
K. Erdmann *Oxford University*
L.C.G. Rogers *University of Bath*
E. Süli *Oxford University*
J.F. Toland *University of Bath*

Other books in this series

A First Course in Discrete Mathematics *I. Anderson*
Analytic Methods for Partial Differential Equations *G. Evans, J. Blackledge, P. Yardley*
Applied Geometry for Computer Graphics and CAD *D. Marsh*
Basic Linear Algebra *T.S. Blyth and E.F. Robertson*
Basic Stochastic Processes *Z. Brzezniak and T. Zastawniak*
Elementary Differential Geometry *A. Pressley*
Elementary Number Theory *G.A. Jones and J.M. Jones*
Elements of Abstract Analysis *M. Ó Searcoid*
Elements of Logic via Numbers and Sets *D.L. Johnson*
Further Linear Algebra *T.S. Blyth and E.F. Robertson*
Geometry *R. Fenn*
Groups, Rings and Fields *D.A.R. Wallace*
Hyperbolic Geometry *J.W. Anderson*
Information and Coding Theory *G.A. Jones and J.M. Jones*
Introduction to Laplace Transforms and Fourier Series *P.P.G. Dyke*
Introduction to Ring Theory *P.M. Cohn*
Introductory Mathematics: Algebra and Analysis *G. Smith*
Introductory Mathematics: Applications and Methods *G.S. Marshall*
Linear Functional Analysis *B.P. Rynne and M.A. Youngson*
Measure, Integral and Probability *M. Capiński and E. Kopp*
Multivariate Calculus and Geometry *S. Dineen*
Numerical Methods for Partial Differential Equations *G. Evans, J. Blackledge, P. Yardley*
Real Analysis *J.M. Howie*
Sets, Logic and Categories *P. Cameron*
Symmetries *D.L. Johnson*
Topics in Group Theory *G. Smith and O. Tabachnikova*
Topologies and Uniformities *I.M. James*
Vector Calculus *P.C. Matthews*

Andrew Baker

Matrix Groups

An Introduction to Lie Group Theory

With 16 Figures



Springer

Andrew Baker, BSc, PhD
Department of Mathematics, University of Glasgow, Glasgow G12 8QW, UK

Cover illustration elements reproduced by kind permission of
Aptech Systems, Inc., Publishers of the GAUSS Mathematical and Statistical System, 23804 S.E. Kent-Kangley Road, Maple Valley, WA 98038,
USA. Tel: (206) 432 - 7855 Fax (206) 432 - 7832 email: info@aptech.com URL: www.aptech.com
American Statistical Association: Chance Vol 8 No 1, 1995 article by KS and KW Heiner 'Tree Rings of the Northern Shawangunks' page 32 fig 2
Springer-Verlag: Mathematics in Education and Research Vol 4 Issue 3 1995 article by Roman E Maeder, Beatrice Amrhein and Oliver Gloer
'Illustrated Mathematics: Visualization of Mathematical Objects' page 9 fig 11, originally published as a CD ROM 'Illustrated Mathematics' by
TELOS: ISBN 0-387-14222-3, German edition by Birkhauser: ISBN 3-7643-5100-4.
Mathematics in Education and Research Vol 4 Issue 3 1995 article by Richard J Gaylord and Kazume Nishidate 'Traffic Engineering with Cellular
Automata' page 35 fig 2. Mathematics in Education and Research Vol 5 Issue 2 1996 article by Michael Trott 'The Implicitization of a Trefoil
Knot' page 14.
Mathematics in Education and Research Vol 5 Issue 2 1996 article by Lee de Cola 'Coins, Trees, Bars and Bells: Simulation of the Binomial Process'
page 19 fig 3. Mathematics in Education and Research Vol 5 Issue 2 1996 article by Richard Gaylord and Kazume Nishidate 'Contagious
Spreading' page 33 fig 1. Mathematics in Education and Research Vol 5 Issue 2 1996 article by Joe Buhler and Stan Wagon 'Secrets of the
Madelung Constant' page 50 fig 1.

British Library Cataloguing in Publication Data

Baker, Andrew

Matrix groups : an introduction to Lie group theory. -
(Springer undergraduate mathematics series)

1. Matrix groups 2. Lie groups

I. Title

512.5'5

ISBN 1852334703

Library of Congress Cataloging-in-Publication Data

Baker, Andrew, 1953-

Matrix groups : an introduction to Lie group theory / Andrew Baker.

p. cm. -- (Springer undergraduate mathematics series)

Includes bibliographical references and index.

ISBN 1-85233-470-3 (alk. paper)

1. Matrix groups. I. Title. II. Series.

QA174.2.B35 2001

512'.2—dc21

2001049261

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

Springer Undergraduate Mathematics Series ISSN 1615-2085

ISBN 1-85233-470-3 Springer-Verlag London Berlin Heidelberg

a member of BertelsmannSpringer Science+Business Media GmbH

<http://www.springer.co.uk>

© Springer-Verlag London Limited 2002

Printed in Great Britain

The use of registered names, trademarks etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant laws and regulations and therefore free for general use.

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Typesetting: Camera ready by the author

Printed and bound at the Athenaeum Press Ltd., Gateshead, Tyne & Wear

12/3830-543210 Printed on acid-free paper SPIN 10830978

This work provides a first taste of the theory of Lie groups accessible to advanced mathematics undergraduates and beginning graduate students, providing an appetiser for a more substantial further course. Although the formal prerequisites are kept as low level as possible, the subject matter is sophisticated and contains many of the key themes of the fully developed theory. We concentrate on *matrix groups*, *i.e.*, closed subgroups of real and complex general linear groups. One of the results proved is that every matrix group is in fact a Lie group, the proof following that in the expository paper of Howe [12]. Indeed, the latter, together with the book of Curtis [7], influenced our choice of goals for the present book and the course which it evolved from. As pointed out by Howe, Lie theoretic ideas lie at the heart of much of standard undergraduate linear algebra, and exposure to them can inform or motivate the study of the latter; we frequently describe such topics in enough detail to provide the necessary background for the benefit of readers unfamiliar with them.

Outline of the Chapters

Each chapter contains exercises designed to consolidate and deepen readers' understanding of the material covered. We also use these to explore related topics that may not be familiar to all readers but which should be in the toolkit of every well-educated mathematics graduate. Here is a brief synopsis of the chapters.

Chapter 1: The *general linear groups* $GL_n(\mathbf{k})$ for $\mathbf{k} = \mathbb{R}$ (the real numbers) and $\mathbf{k} = \mathbb{C}$ (the complex numbers) are introduced and studied both as groups and as topological spaces. *Matrix groups* are defined and a number of standard examples discussed, including *special linear groups* $SL_n(\mathbf{k})$, *orthogonal groups* $O(n)$ and *special orthogonal groups* $SO(n)$, *unitary groups* $U(n)$ and *special unitary groups* $SU(n)$, as well as more exotic examples such as *Lorentz groups*.

and *symplectic groups*. The relation of complex to real matrix groups is also studied. Along the way we discuss various algebraic, analytic and topological notions including *norms*, *metric spaces*, *compactness* and *continuous group actions*.

Chapter 2: The *exponential function* for matrices is introduced and *one-parameter subgroups* of matrix groups are studied. We show how these ideas can be used in the solution of certain types of differential equations.

Chapter 3: The idea of a *Lie algebra* is introduced and various algebraic properties are studied. *Tangent spaces* and Lie algebras of matrix groups are defined together with the *adjoint action*. The important special case of $SU(2)$ and its relationship to $SO(3)$ is studied in detail.

Chapters 4 and 5: *Finite dimensional algebras* over fields, especially \mathbb{R} or \mathbb{C} , are defined and their units viewed 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* $Sp(n)$ are also defined, completing the list of compact connected classical groups and their universal covers. *Automorphism groups* of algebras are also shown to provide further examples of matrix groups.

Chapter 6: The geometry and linear algebra of *Lorentz groups* which are of importance in Relativity are studied. The relationship of $SL_2(\mathbb{C})$ to the Lorentz group $Lor(3,1)$ is discussed, extending the work on $SU(2)$ and $SO(3)$ in Chapter 3.

Chapter 7: The general notion of a *Lie group* is introduced and we show that all matrix groups are *Lie subgroups* of general linear groups. Along the way we introduce the basic ideas of *differentiable manifolds* and *smooth maps*. We show that not every Lie group can be realised as a matrix group by considering the simplest *Heisenberg group*.

Chapters 8 and 9: *Homogeneous spaces* of Lie groups are defined and we show how to recognise them as orbits of *smooth actions*. We discuss *connectivity* of Lie groups and use homogeneous spaces to prove that many familiar Lie groups are path connected. We also describe some important families of homogeneous spaces such as *projective spaces* and *Grassmannians*, as well as examples related to special factorisations of matrices such as *polar form*.

Chapters 10, 11 and 12: The basic theory of *compact connected Lie groups* and their *maximal tori* is studied and the relationship to some well-known matrix diagonalisation results highlighted. We continue this theme by describing the classification theory of compact connected *simple Lie groups*, showing how the families we meet in earlier chapters provide all but a finite number of the isomorphism types predicted. *Root systems*, *Weyl groups* and *Dynkin diagrams* are defined and many examples described.

Some suggestions for using this book

For an advanced undergraduate course of about 30 lectures to students already equipped with basic real and complex analysis, metric spaces, linear algebra, group and ring theory, the material of Chapters 1, 3, 7 provide an introduction to matrix groups, while Chapters 4, 5, 6, 8, 9 supply extra material that might be quarried for further examples. A more ambitious course aimed at presenting the classical compact connected Lie groups might take in Chapters 4, 5 and perhaps lead on to some of the theory of compact connected Lie groups discussed in Chapters 10, 11, 12.

A reader (perhaps a graduate student) using the book on their own would find it useful to follow up some of the references [6, 8, 17, 18, 26, 29] to see more advanced approaches to the topics on differential geometry and topology covered in Chapters 7, 8, 9 and the classification theory of Chapters 10, 11, 12.

Each chapter has a set of Exercises of varying degrees of difficulty. Hints and solutions are provided for some of these, the more challenging questions being indicated by the symbols Δ or $\Delta\Delta$ with the latter intended for readers wishing to pursue the material in greater depth.

Prerequisites and assumptions

The material in Chapters 1, 3, 7 is intended to be accessible to a well-equipped advanced undergraduate, although many topics such as non-metric topological spaces, normed vector spaces and rings may be unfamiliar so we have given the relevant definitions. We do not assume much abstract algebra beyond standard notions of homomorphisms, subobjects, kernels and images and quotients; semi-direct products of groups are introduced, as are Lie algebras. A course on matrix groups is a good setting to learn algebra, and there are many significant algebraic topics in Chapters 4, 5, 11, 12. Good sources of background material are [5, 15, 16, 22, 28].

The more advanced parts of the theory which are described in Chapters 7, 8, 9, 10, 11, 12 should certainly challenge students and naturally point to more detailed studies of Differential Geometry and Lie Theory. Occasionally ideas from Algebraic Topology are touched upon (e.g., the fundamental group and Lefschetz Fixed Point Theorem) and an interested reader might find it helpful to consult an introductory book on the subject such as [9, 20, 26].

Typesetting

This book was produced using L^AT_EX and the American Mathematical Society's *amsmath* package. Diagrams were produced using *xypic*. The symbol Δ was produced by my colleague J. Nimmo, who also provided other help with T_EX and L^AT_EX.

Acknowledgements

I would like to thank the following: the Universität Bern for inviting me to visit and teach a course in the Spring of 2000; the students who spotted errors and obscurities in my notes and Z. Balogh who helped with the problem classes; the mathematicians of Glasgow University, especially I. Gordon, J. Nimmo, R. Odoni and J. Webb; the topologists and other mathematicians of Manchester University from whom I learnt a great deal over many years. Also thanks to Roger and Marliese Delaquis for providing a temporary home in the wonderful city of Bern. Finally, special thanks must go to Carole, Daniel and Laura for putting up with it all.

Contents

Part I. Basic Ideas and Examples

1.	Real and Complex Matrix Groups	3
1.1	Groups of Matrices	3
1.2	Groups of Matrices as Metric Spaces	5
1.3	Compactness	12
1.4	Matrix Groups	15
1.5	Some Important Examples	17
1.6	Complex Matrices as Real Matrices	29
1.7	Continuous Homomorphisms of Matrix Groups	31
1.8	Matrix Groups for Normed Vector Spaces	33
1.9	Continuous Group Actions	37
2.	Exponentials, Differential Equations and One-parameter Subgroups	45
2.1	The Matrix Exponential and Logarithm	45
2.2	Calculating Exponentials and Jordan Form	51
2.3	Differential Equations in Matrices	55
2.4	One-parameter Subgroups in Matrix Groups	56
2.5	One-parameter Subgroups and Differential Equations	59
3.	Tangent Spaces and Lie Algebras	67
3.1	Lie Algebras	67
3.2	Curves, Tangent Spaces and Lie Algebras	71
3.3	The Lie Algebras of Some Matrix Groups	76

3.4 Some Observations on the Exponential Function of a Matrix Group	84
3.5 $SO(3)$ and $SU(2)$	86
3.6 The Complexification of a Real Lie Algebra	92
4. Algebras, Quaternions and Quaternionic Symplectic Groups	99
4.1 Algebras	99
4.2 Real and Complex Normed Algebras	111
4.3 Linear Algebra over a Division Algebra	113
4.4 The Quaternions	116
4.5 Quaternionic Matrix Groups	120
4.6 Automorphism Groups of Algebras	122
5. Clifford Algebras and Spinor Groups	129
5.1 Real Clifford Algebras	130
5.2 Clifford Groups	139
5.3 Pinor and Spinor Groups	143
5.4 The Centres of Spinor Groups	151
5.5 Finite Subgroups of Spinor Groups	152
6. Lorentz Groups	157
6.1 Lorentz Groups	157
6.2 A Principal Axis Theorem for Lorentz Groups	165
6.3 $SL_2(\mathbb{C})$ and the Lorentz Group $Lor(3,1)$	171

Part II. Matrix Groups as Lie Groups

7. Lie Groups	181
7.1 Smooth Manifolds	181
7.2 Tangent Spaces and Derivatives	183
7.3 Lie Groups	187
7.4 Some Examples of Lie Groups	189
7.5 Some Useful Formulae in Matrix Groups	193
7.6 Matrix Groups are Lie Groups	199
7.7 Not All Lie Groups are Matrix Groups	203
8. Homogeneous Spaces	211
8.1 Homogeneous Spaces as Manifolds	211
8.2 Homogeneous Spaces as Orbits	215
8.3 Projective Spaces	217
8.4 Grassmannians	222

8.5 The Gram–Schmidt Process	224
8.6 Reduced Echelon Form	226
8.7 Real Inner Products	227
8.8 Symplectic Forms	229
9. Connectivity of Matrix Groups	235
9.1 Connectivity of Manifolds	235
9.2 Examples of Path Connected Matrix Groups	238
9.3 The Path Components of a Lie Group	241
9.4 Another Connectivity Result	244

Part III. Compact Connected Lie Groups and their Classification

10. Maximal Tori in Compact Connected Lie Groups	251
10.1 Tori	251
10.2 Maximal Tori in Compact Lie Groups	255
10.3 The Normaliser and Weyl Group of a Maximal Torus	259
10.4 The Centre of a Compact Connected Lie Group	262
11. Semi-simple Factorisation	267
11.1 An Invariant Inner Product	267
11.2 The Centre and its Lie Algebra	270
11.3 Lie Ideals and the Adjoint Action	272
11.4 Semi-simple Decompositions	276
11.5 The Structure of the Adjoint Representation	278
12. Roots Systems, Weyl Groups and Dynkin Diagrams	289
12.1 Inner Products and Duality	289
12.2 Roots systems and their Weyl groups	291
12.3 Some Examples of Root Systems	293
12.4 The Dynkin Diagram of a Root System	297
12.5 Irreducible Dynkin Diagrams	298
12.6 From Root Systems to Lie Algebras	299
Hints and Solutions to Selected Exercises	303
Bibliography	323
Index	325

Part I

Basic Ideas and Examples

Real and Complex Matrix Groups

Throughout, \mathbf{k} will denote a (commutative) field. Most of the time we will be interested in the cases of the fields $\mathbf{k} = \mathbb{R}$ (the real numbers) and $\mathbf{k} = \mathbb{C}$ (the complex numbers), however the general framework of this chapter is applicable to more general fields equipped with suitable norms in place of the absolute value. Indeed, as we will see in Chapter 4, much of it even applies to the case of a general *normed division algebra* or *skew field*, with the *quaternions* providing the most important non-commutative example.

1.1 Groups of Matrices

Let $M_{m,n}(\mathbf{k})$ be the set of $m \times n$ matrices whose entries are in \mathbf{k} . We will denote the (i,j) entry of an $m \times n$ matrix A by A_{ij} or a_{ij} and also write

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

We will use the special notations

$$M_n(\mathbf{k}) = M_{n,n}(\mathbf{k}), \quad \mathbf{k}^n = M_{n,1}(\mathbf{k}).$$

$M_{m,n}(\mathbf{k})$ is a \mathbf{k} -vector space with the operations of matrix addition and scalar multiplication. The zero vector is the $m \times n$ zero matrix $O_{m,n}$ which we

will often denote O when the size is clear from the context. The matrices E^{rs} with $r = 1, \dots, m$, $s = 1, \dots, n$ and

$$(E^{rs})_{ij} = \delta_{ir} \delta_{js} = \begin{cases} 1 & \text{if } i = r \text{ and } j = s, \\ 0 & \text{otherwise,} \end{cases}$$

form a basis of $M_{m,n}(\mathbf{k})$, hence its dimension as a \mathbf{k} -vector space is

$$\dim_{\mathbf{k}} M_{m,n}(\mathbf{k}) = mn. \quad (1.1)$$

When $n = 1$ we will denote the standard basis vectors of $\mathbf{k}^n = M_{n,1}(\mathbf{k})$ by

$$\mathbf{e}_r = E^{r1} \quad (r = 1, \dots, m).$$

As well as being a \mathbf{k} -vector space of dimension n^2 , $M_n(\mathbf{k})$ is also a ring with the usual addition and multiplication of square matrices, with zero $O_n = O_{n,n}$ and the $n \times n$ identity matrix I_n as its unity; $M_n(\mathbf{k})$ is not commutative except when $n = 1$. Later we will see that $M_n(\mathbf{k})$ is also an important example of a *finite dimensional \mathbf{k} -algebra* in the sense to be introduced in Chapter 4. The ring $M_n(\mathbf{k})$ acts on \mathbf{k}^n by left multiplication, giving \mathbf{k}^n the structure of a *left $M_n(\mathbf{k})$ -module*.

Proposition 1.1

The determinant function $\det: M_n(\mathbf{k}) \rightarrow \mathbf{k}$ has the following properties.

- i) For $A, B \in M_n(\mathbf{k})$, $\det(AB) = \det A \det B$.
- ii) $\det I_n = 1$.
- iii) $A \in M_n(\mathbf{k})$ is invertible if and only if $\det A \neq 0$.

We will use the notation

$$GL_n(\mathbf{k}) = \{A \in M_n(\mathbf{k}) : \det A \neq 0\}$$

for the set of invertible $n \times n$ matrices (also known as the set of units of the ring $M_n(\mathbf{k})$), and

$$SL_n(\mathbf{k}) = \{A \in M_n(\mathbf{k}) : \det A = 1\} \subseteq GL_n(\mathbf{k})$$

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

Theorem 1.2

The sets $GL_n(\mathbf{k})$, $SL_n(\mathbf{k})$ are groups under matrix multiplication. Furthermore, $SL_n(\mathbf{k}) \leq GL_n(\mathbf{k})$, i.e., $SL_n(\mathbf{k})$ is a subgroup of $GL_n(\mathbf{k})$.

Because of these group structures, $GL_n(\mathbf{k})$ is called the $n \times n$ general linear group, while $SL_n(\mathbf{k})$ is called the $n \times n$ special linear or unimodular group. When $\mathbf{k} = \mathbb{R}$ or $\mathbf{k} = \mathbb{C}$ we will refer to $GL_n(\mathbb{R})$, $SL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$, $SL_n(\mathbb{C})$ as real or complex general linear groups. Of course, we can also consider subgroups of these groups, but before doing so we consider the topology of $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ as metric spaces.

1.2 Groups of Matrices as Metric Spaces

In this section we will always assume that $\mathbf{k} = \mathbb{R}$ or \mathbb{C} . Recall that $M_n(\mathbf{k})$ is a \mathbf{k} -vector space of dimension n^2 . We will define a norm $\|\cdot\|$ on $M_n(\mathbf{k})$. It is worth remarking that we choose this particular norm mainly for the convenience of its multiplicative properties; in fact, as explained in Section 1.8, any other vector space norm would give an equivalent metric topology on $M_n(\mathbf{k})$. Other useful norms on $M_n(\mathbf{k})$ are discussed in Strang [28].

We begin with the usual notion of length for a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbf{k}^n$,

namely

$$|\mathbf{x}| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}.$$

This is an example of a norm on the vector space \mathbf{k}^n as specified in Definition 1.51.

For $A \in M_n(\mathbf{k})$, consider the set

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

Then the subset

$$\mathcal{S}_A^1 = \{ |Ax| : \mathbf{x} \in \mathbf{k}^n, |\mathbf{x}| = 1 \} \subseteq \mathcal{S}_A$$

is actually equal to \mathcal{S}_A since if $\mathbf{x} \neq 0$ we have

$$\frac{|Ax|}{|\mathbf{x}|} = |Ax'|,$$

where $\mathbf{x}' = (1/|\mathbf{x}|)\mathbf{x}$ has length $|\mathbf{x}'| = 1$. The subset

$$\{\mathbf{x} \in \mathbf{k}^n : |\mathbf{x}| = 1\} \subseteq \mathbf{k}^n$$

is closed and bounded and so is compact in the sense of Section 1.3, hence by Corollary 1.23, the real-valued function

$$\{\mathbf{x} \in \mathbf{k}^n : |\mathbf{x}| = 1\} \longrightarrow \mathbb{R}; \quad \mathbf{x} \longmapsto |Ax|$$

is bounded and attains its supremum

$$\sup \mathcal{S}_A = \sup \mathcal{S}_A^1 = \max \mathcal{S}_A^1 = \max \mathcal{S}_A.$$

This means that the real number

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

is defined. This *norm* function $\|\cdot\|: M_n(\mathbb{k}) \rightarrow \mathbb{R}$ is called the *operator* or *sup* (= *supremum*) *norm* on $M_n(\mathbb{k})$. For the general notion of a *norm* on a \mathbb{k} -algebra see Definition 4.31.

For a real matrix $A \in M_n(\mathbb{R}) \subseteq M_n(\mathbb{R})$, at first sight there appear to be two distinct norms of this type, namely

$$\|A\|_{\mathbb{R}} = \{|Ax| : x \in \mathbb{R}^n, |x| = 1\}, \quad \|A\|_{\mathbb{C}} = \{|Ax| : x \in \mathbb{C}^n, |x| = 1\}.$$

Lemma 1.3

If $A \in M_n(\mathbb{R})$, then $\|A\|_{\mathbb{C}} = \|A\|_{\mathbb{R}}$.

Proof

It is obvious that $\|A\|_{\mathbb{R}} \leq \|A\|_{\mathbb{C}}$. Now for a vector $z \in \mathbb{C}^n$ with $|z| = 1$, write $z = x + iy$ with $x, y \in \mathbb{R}^n$. Then $|x|^2 + |y|^2 = 1$ and

$$\begin{aligned} |Az|^2 &\leq |Ax|^2 + |iAy|^2 \\ &\leq |x|^2 \|A\|_{\mathbb{R}}^2 + |y|^2 \|A\|_{\mathbb{R}}^2 \\ &= (|x|^2 + |y|^2) \|A\|_{\mathbb{R}}^2 \\ &= \|A\|_{\mathbb{R}}^2, \end{aligned}$$

giving $|Az| \leq \|A\|_{\mathbb{R}}$. Thus $\|A\|_{\mathbb{C}} \leq \|A\|_{\mathbb{R}}$ and hence $\|A\|_{\mathbb{R}} = \|A\|_{\mathbb{C}}$. □

Remark 1.4

There is a procedure for calculating $\|A\|$ which is important in numerical linear algebra. We describe this briefly; for further details see Strang [28].

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_{\max} . Then

$$\|A\| = \sqrt{\lambda_{\max}}.$$

In fact, for any unit eigenvector v of A^*A for the eigenvalue λ_{\max} , $\|A\| = |Av|$. When A is real, $A^*A = A^T A$ is real positive symmetric and there are unit length

eigenvectors $w \in \mathbb{R}^n \subseteq \mathbb{C}^n$ of A^*A for the eigenvalue λ for which $\|A\| = |Aw|$. In particular, this also shows that $\|A\|$ is independent of whether A is viewed as a real or complex matrix.

The main properties of $\|\cdot\|$ are summarised in the next result and imply that $\|\cdot\|$ is a \mathbf{k} -norm on $M_n(\mathbf{k})$. General \mathbf{k} -norms are discussed in Section 4.2.

Proposition 1.5

The function $\|\cdot\|$ has the following properties.

- i) If $t \in \mathbf{k}$, $A \in M_n(\mathbf{k})$, then $\|tA\| = |t| \|A\|$.
- ii) If $A, B \in M_n(\mathbf{k})$, then $\|AB\| \leq \|A\| \|B\|$.
- iii) If $A, B \in M_n(\mathbf{k})$, then $\|A + B\| \leq \|A\| + \|B\|$.
- iv) If $A \in M_n(\mathbf{k})$, then $\|A\| = 0$ if and only if $A = 0$.
- v) $\|I_n\| = 1$.

The norm $\|\cdot\|$ can be used to define a metric ρ on $M_n(\mathbf{k})$ by

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

together with the associated metric topology on $M_n(\mathbf{k})$. Then a sequence $\{A_r\}_{r \geq 0}$ of elements in $M_n(\mathbf{k})$ converges to a limit $A \in M_n(\mathbf{k})$ if $\|A_r - A\| \rightarrow 0$ as $r \rightarrow \infty$. We may also define continuous functions $M_n(\mathbf{k}) \rightarrow X$ into a topological space X .

For $A \in M_n(\mathbf{k})$ and $r > 0$, let

$$N_{M_n(\mathbf{k})}(A; r) = \{B \in M_n(\mathbf{k}) : \|B - A\| < r\},$$

which is the *open disc* of radius r in $M_n(\mathbf{k})$. Similarly, if $Y \subseteq M_n(\mathbf{k})$ and $A \in Y$, we set

$$N_Y(A; r) = \{B \in Y : \|B - A\| < r\} = N_{M_n(\mathbf{k})}(A; r) \cap Y.$$

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 $N_Y(A; \delta) \subseteq V$.

Definition 1.6

Let $Y \subseteq M_n(\mathbf{k})$ and (X, \mathcal{T}) be a topological space. Then a function $f: Y \rightarrow X$ 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

$$B \in N_Y(A; \delta) \implies f(B) \in U.$$

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

For a topological space (X, \mathcal{T}) , a subset $W \subseteq X$ is *closed* if $X - W \subseteq X$ is open. For a metric space this is equivalent to requiring that whenever a sequence in W has a limit in X , the limit is in W . 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 $X = \mathbf{k}$ and \mathcal{T} to be the natural metric space topology associated to the standard norm on \mathbf{k} and consider continuous functions $Y \rightarrow \mathbf{k}$.

Proposition 1.7

For $1 \leq r, s \leq n$, the *coordinate function*

$$\text{coord}_{rs}: M_n(\mathbf{k}) \rightarrow \mathbf{k}; \quad \text{coord}_{rs}(A) = A_{rs}$$

is continuous.

Proof

For the standard unit basis vectors e_i ($1 \leq i \leq n$) of \mathbf{k}^n we have

$$\begin{aligned} |A_{rs}| &\leq \sqrt{\sum_{i=1}^n |A_{is}|^2} \\ &= \left| \sum_{i=1}^n A_{is} e_i \right| \\ &= |Ae_s| \\ &\leq \|A\|, \end{aligned}$$

hence for $A, A' \in M_n(\mathbf{k})$,

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

If $A \in M_n(\mathbf{k})$ and $\epsilon > 0$, then $\|A' - A\| < \epsilon$ implies that $|A'_{rs} - A_{rs}| < \epsilon$. This shows that coord_{rs} is continuous at every $A \in M_n(\mathbf{k})$. \square

Corollary 1.8

If $f: \mathbf{k}^{n^2} \rightarrow \mathbf{k}$ is continuous, then the associated function

$$F: M_n(\mathbf{k}) \rightarrow \mathbf{k}; \quad F(A) = f((A_{ij})_{1 \leq i,j \leq n}),$$

is continuous.

Corollary 1.9

The determinant $\det: M_n(\mathbf{k}) \rightarrow \mathbf{k}$ and trace $\text{tr}: M_n(\mathbf{k}) \rightarrow \mathbf{k}$ are continuous functions.

Proof

The determinant is obtained by composing a continuous function $M_n(\mathbf{k}) \rightarrow \mathbf{k}^{n^2}$ identifying $M_n(\mathbf{k})$ with \mathbf{k}^{n^2} with a polynomial function $\mathbf{k}^{n^2} \rightarrow \mathbf{k}$. Similarly,

$$\text{tr } A = \sum_{k=1}^n A_{kk}$$

defines the trace as a polynomial function. \square

There is a kind of converse to these results.

Proposition 1.10

For $A \in M_n(\mathbf{k})$,

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

Proof

Let $\mathbf{x} = x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n$ with $|\mathbf{x}| = 1$. Since $|x_k| \leq 1$ for each k , we have

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

As this is true for all vectors \mathbf{x} with $|\mathbf{x}| = 1$,

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

by definition of $\|A\|$. \square

Definition 1.11

A sequence $\{A_r\}_{r \geq 0}$ in $M_n(\mathbf{k})$ is a *Cauchy sequence* if for every $\epsilon > 0$, there is a natural number N such that $\|A_r - A_s\| < \epsilon$ whenever $r, s > N$.

Theorem 1.12

For $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , every Cauchy sequence $\{A_r\}_{r \geq 0}$ in $M_n(\mathbf{k})$ has a unique limit $\lim_{r \rightarrow \infty} A_r$ in $M_n(\mathbf{k})$. Furthermore,

$$(\lim_{r \rightarrow \infty} A_r)_{ij} = \lim_{r \rightarrow \infty} (A_r)_{ij}. \quad (1.2)$$

Proof

It is standard that if such a limit exists it is unique so we need to show existence. By Proposition 1.7, the limit on the right-hand side of Equation (1.2) exists, so it is sufficient to show that the required limit is the matrix A for which

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

For the sequence $\{A_r - A\}_{r \geq 0}$, as $r \rightarrow \infty$ we have

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

so $A_r \rightarrow A$ by Proposition 1.10. □

Because of this result, the metric space $(M_n(\mathbf{k}), \| \cdot \|)$ is said to be *complete* with respect to the norm $\| \cdot \|$.

It can be shown that the metric topologies induced by $\| \cdot \|$ and the usual norm on \mathbf{k}^{n^2} agree in the sense that they have the same open sets. Actually this is true for any two norms on \mathbf{k}^{n^2} ; see Section 1.8 and [21, 22] for more on this. We summarise this as a useful criterion whose proof is left as an exercise.

Proposition 1.13

A function $F: M_m(\mathbf{k}) \rightarrow M_n(\mathbf{k})$ is continuous with respect to the norms $\| \cdot \|$ if and only if each of the component functions $F_{rs}: M_m(\mathbf{k}) \rightarrow \mathbf{k}$ is continuous. In particular, a function $f: M_m(\mathbf{k}) \rightarrow \mathbf{k}$ is continuous with respect to the norm $\| \cdot \|$ and the usual metric on \mathbf{k} if and only if it is continuous when viewed as a function $\mathbf{k}^{m^2} \rightarrow \mathbf{k}$.

Next we consider the topology of some subsets of $M_n(\mathbf{k})$, in particular some groups of matrices.

Proposition 1.14

Let $\mathbf{k} = \mathbb{R}$ or \mathbb{C} . Then

- i) $GL_n(\mathbf{k}) \subseteq M_n(\mathbf{k})$ is an open subset;
- ii) $SL_n(\mathbf{k}) \subseteq M_n(\mathbf{k})$ is a closed subset.

Proof

We know that the function $\det: M_n(\mathbf{k}) \rightarrow \mathbf{k}$ is continuous. Then

$$GL_n(\mathbf{k}) = M_n(\mathbf{k}) - \det^{-1}\{0\},$$

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

$$SL_n(\mathbf{k}) = \det^{-1}\{1\} \subseteq GL_n(\mathbf{k}),$$

which is closed in $M_n(\mathbf{k})$ and $GL_n(\mathbf{k})$ since the singleton set $\{1\}$ is closed in \mathbf{k} , so (ii) is true. \square

In the following we will make use of the *product topology* of two topological spaces X, Y ; this is the topology on $X \times Y$ in which every open set is a union of sets of the form

$$U \times V \quad (U \subseteq X, V \subseteq Y \text{ open}).$$

We refer to $X \times Y$ as the product space if it has the product topology. If the topologies on X and Y come from metrics, it is possible to define a metric whose associated topology agrees with the product topology; this is discussed in the Exercises.

The addition and multiplication maps

$$\text{add}: M_n(\mathbf{k}) \times M_n(\mathbf{k}) \rightarrow M_n(\mathbf{k}); \quad \text{add}(X, Y) = X + Y,$$

$$\text{mult}: M_n(\mathbf{k}) \times M_n(\mathbf{k}) \rightarrow M_n(\mathbf{k}); \quad \text{mult}(X, Y) = XY,$$

are also continuous, where we take the product topology on the domain $M_n(\mathbf{k}) \times M_n(\mathbf{k})$. Finally, the inverse map

$$\text{inv}: GL_n(\mathbf{k}) \rightarrow GL_n(\mathbf{k}); \quad \text{inv}(A) = A^{-1},$$

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

(polynomial in the entries A_{ij})

$$\det A$$

and as this is a continuous function of the entries of A it is a continuous function of A .

Definition 1.15

Let G be a topological space and view $G \times G$ as the product space. Suppose that G is also a group with multiplication map $\text{mult}: G \times G \rightarrow G$ and inverse map $\text{inv}: G \rightarrow G$. Then G is a *topological group* if mult and inv are continuous.

The simplest examples are obtained from arbitrary groups G given discrete topologies; in particular all finite groups can be viewed this way. Of course, the discussion above has already established the following.

Theorem 1.16

For $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , each of the groups $\text{GL}_n(\mathbf{k})$, $\text{SL}_n(\mathbf{k})$ is a topological group with the evident multiplication and inverse maps and the subspace topologies inherited from $M_n(\mathbf{k})$.

1.3 Compactness

In this section we discuss the idea of compactness for topological spaces and explain its significance for subsets of \mathbf{k}^n with the usual metric, where $\mathbf{k} = \mathbb{R}$ or \mathbb{C} . Many of the most useful results for continuous functions from a compact space into a metric space also apply more generally when the domain is *Hausdorff* in the sense of Definition 1.24.

Definition 1.17

A subset $X \subseteq \mathbf{k}^m$ is *compact* if and only if it is closed and bounded.

Example 1.18

Identifying subsets of $M_n(\mathbf{k})$ with subsets of \mathbf{k}^{n^2} , we can consider compact subsets of $M_n(\mathbf{k})$. In particular, a subgroup $G \leq \text{GL}_n(\mathbf{k})$ is compact if it is compact as a subset of $\text{GL}_n(\mathbf{k})$, or equivalently of $M_n(\mathbf{k})$.

Our next result is standard for metric spaces.

Proposition 1.19

$X \subseteq M_n(\mathbf{k})$ is compact if and only if the following two conditions are satisfied:

- there is a $b \in \mathbb{R}^+$ such that for all $A \in X$, $\|A\| \leq b$;
- every sequence $\{C_n\}_{n \geq 0}$ in X which is convergent in $M_n(\mathbf{k})$ has a limit in X , i.e., X is a closed subset of $M_n(\mathbf{k})$.

The following important characterisation of compact subsets of $M_n(\mathbf{k})$ leads to the general definition of *compact topological space*.

Theorem 1.20 (Heine–Borel Theorem)

$X \subseteq M_n(\mathbf{k})$ is compact if and only if every open cover $\{U_\alpha\}_{\alpha \in \Lambda}$ of X contains a finite subcover $\{U_{\alpha_1}, \dots, U_{\alpha_k}\}$.

Definition 1.21

A topological space X is *compact* if and only if every open cover $\{U_\alpha\}_{\alpha \in \Lambda}$ of X contains a finite subcover $\{U_{\alpha_1}, \dots, U_{\alpha_k}\}$.

Clearly our two notions of compactness coincide for a subset $X \subseteq \mathbf{k}^n$.

Proposition 1.22

Let X be a compact topological space and $f: X \rightarrow Y$ be a continuous function. Then the image $fX \subseteq Y$ is a compact subspace of Y .

Proof

Let $\{V_\alpha\}_{\alpha \in \Lambda}$ be an open cover of fX . Then by definition of the subspace topology, there is a collection of open subsets $\{V'_\alpha\}_{\alpha \in \Lambda}$ for which $fX \cap V'_\alpha = V_\alpha$. For each $\alpha \in \Lambda$,

$$f^{-1}V_\alpha = f^{-1}V'_\alpha,$$

so $\{f^{-1}V'_\alpha\}_{\alpha \in \Lambda}$ is an open covering of X . By compactness, there is a finite subcollection $\{f^{-1}V'_{\alpha_1}, \dots, f^{-1}V'_{\alpha_k}\}$ which also covers X , hence $\{V'_{\alpha_1}, \dots, V'_{\alpha_k}\}$ is a finite cover of fX . \square

Corollary 1.23

Let X be a compact topological space and $f: X \rightarrow \mathbb{R}$ be a continuous function. Then the image $fX \subseteq \mathbb{R}$ is a bounded subset and there are elements $x_+, x_- \in X$ for which

$$f(x_+) = \sup fX, \quad f(x_-) = \inf fX.$$

For later use we record some other useful results on continuous functions out of a compact space into a *Hausdorff* space. Omitted proofs can be found in books on point set topology. We start by introducing the notion of a *Hausdorff* space which provides a useful generalisation of the concept of a metric space particularly useful when dealing with quotient constructions such as the *homogeneous spaces* of Chapter 8.

Definition 1.24

A topological space X is *Hausdorff* if for every pair of points $u, v \in X$ with $u \neq v$, there are open subsets $U, V \subseteq X$ with $u \in U, v \in V$ and $U \cap V = \emptyset$.

Lemma 1.25

Every metric space is Hausdorff.

Proof

Let X be a metric space with metric ρ . If $u, v \in X$ are distinct, then $\rho(u, v) > 0$. If $r = \rho(u, v)/2$, the open discs $N_X(u; r)$ and $N_X(v; r)$ satisfy the conditions required for the open sets U and V . \square

Proposition 1.26

Let X be a compact topological space. If $f: X \rightarrow Y$ is a continuous function into a Hausdorff topological space Y , then $fX \subseteq Y$ is a closed subset. In particular, when $j: X \rightarrow Y$ is the inclusion function for a subspace $X \subseteq Y$, we obtain that X is a closed subset of Y .

Our next definition provides a notion of equivalence of topological spaces.

Definition 1.27

A continuous bijection $f: X \rightarrow Y$ between topological spaces X and Y is a *homeomorphism* if its inverse $f^{-1}: Y \rightarrow X$ is continuous.

Proposition 1.28

Let X be a compact topological space and $f: X \rightarrow Y$ a continuous bijection into a Hausdorff topological space Y . Then f is a homeomorphism.

Our final result will be useful when working with compact matrix groups.

Proposition 1.29

Let (X, ρ) be a compact metric space and let $\{s_n\}_{n \geq 1}$ be a sequence in X . Then there is a convergent subsequence $\{s_{\sigma(n)}\}_{n \geq 1}$.

Proof

Suppose that this is false. Then for every $x \in X$ which is not of the form $x = s_n$ for some n , there is an $r_x > 0$ for which

$$\rho(x, s_n) \geq r_x \quad (n \geq 1).$$

Also, for each $m \geq 1$, there is an $r_{s_m} > 0$ and an $n_m \geq m$ for which

$$\rho(s_m, s_n) \geq r_x \quad (n \geq n_m).$$

The open discs $N_X(x; r_x)$ form an open cover of X , hence by compactness there are elements x_1, \dots, x_k for which

$$X = N_X(x_1; r_{x_1}) \cup \dots \cup N_X(x_k; r_{x_k}).$$

But for sufficiently large n , s_n cannot be in any of the discs $N_X(x_j; r_{x_j})$, so this provides a contradiction. \square

1.4 Matrix Groups

Definition 1.30

A subgroup $G \leq \mathrm{GL}_n(\mathbf{k})$ which is also a closed subspace is a *matrix group over \mathbf{k}* or a *\mathbf{k} -matrix group*. In order to make the value of n explicit, we will sometimes say that G is a *matrix subgroup* of $\mathrm{GL}_n(\mathbf{k})$.

Before considering some examples we record some useful general properties of matrix groups.

Proposition 1.31

Let $G \leq \mathrm{GL}_n(\mathbf{k})$ be a matrix subgroup. Then a closed subgroup $H \leq G$ is a matrix subgroup H of $\mathrm{GL}_n(\mathbf{k})$.

Proof

Every sequence $\{A_n\}_{n \geq 0}$ in H with a limit in $GL_n(\mathbf{k})$ actually has its limit in G since $A_n \in H \subseteq G$ for every n and G is closed in $GL_n(\mathbf{k})$. Since H is closed in G , this means that $\{A_n\}_{n \geq 0}$ has a limit in H . So H is closed in $GL_n(\mathbf{k})$ which shows that it is a matrix subgroup. \square

This result suggests another definition.

Definition 1.32

A closed subgroup $H \leq G$ of a matrix group G is called a *matrix subgroup* of G .

Proposition 1.33

Let G be a matrix group and $H \leq K, K \leq G$ be matrix subgroups. Then H is a matrix subgroup of G .

Proof

This is a straightforward generalisation of Proposition 1.31. \square

Example 1.34

$SL_n(\mathbf{k}) \leq GL_n(\mathbf{k})$ is a matrix group over \mathbf{k} .

Proof

$SL_n(\mathbf{k})$ is closed in $M_n(\mathbf{k})$ by Proposition 1.14 and $SL_n(\mathbf{k}) \leq GL_n(\mathbf{k})$. \square

Example 1.35

We may consider $GL_n(\mathbf{k})$ as a subgroup of $GL_{n+1}(\mathbf{k})$ by identifying the $n \times n$ matrix $A = [a_{ij}]$ with

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

and it is easily verified that $GL_n(\mathbf{k})$ is closed in $GL_{n+1}(\mathbf{k})$, hence $GL_n(\mathbf{k})$ is a matrix subgroup of $GL_{n+1}(\mathbf{k})$. We can restrict this to an embedding of $SL_n(\mathbf{k})$ which then appears as a closed subgroup of $SL_{n+1}(\mathbf{k}) \leq GL_{n+1}(\mathbf{k})$. So $SL_n(\mathbf{k})$ is a matrix subgroup of $SL_{n+1}(\mathbf{k})$. More generally, with the aid of this embedding, any matrix subgroup of $GL_n(\mathbf{k})$ can also be viewed as a matrix subgroup of $GL_{n+1}(\mathbf{k})$.

Given a matrix subgroup $G \leq GL_n(\mathbf{k})$, it is often useful to restrict the determinant on $GL_n(\mathbf{k})$ to a function

$$\det_G : G \longrightarrow \mathbf{k}^\times; \quad \det_G A = \det A.$$

We usually write this function as \det when no ambiguity is likely to result. Of course, \det_G is always a continuous group homomorphism.

When $\mathbf{k} = \mathbb{R}$, 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 $GL_1(\mathbb{R}) = \mathbb{R}^\times$ which is both closed and open as a subset, while \mathbb{R}^- is an open subset; thus \mathbb{R}^+ and \mathbb{R}^- are *clopen* subsets, i.e., both closed and open. For $G \leq GL_n(\mathbb{R})$,

$$\det_G^{-1} \mathbb{R}^+ = G \cap \det^{-1} GL_n(\mathbb{R}),$$

and also

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

Hence G is a disjoint union of the clopen subsets

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

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 $GL_n(\mathbb{R})$. 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.

1.5 Some Important Examples

We now discuss some important examples of real and complex matrix groups.

Groups of Upper Triangular Matrices

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

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

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 subgroup* or *Borel subgroup* of $\mathrm{GL}_n(\mathbf{k})$ is

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

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

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

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

For the case

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

the function

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

is a continuous group homomorphism which is an isomorphism with continuous inverse. This allows us to view \mathbf{k} as a matrix group.

Affine Groups

The n -dimensional affine group over \mathbf{k} is

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

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

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

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

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

and this is a closed subgroup. There is also the closed subgroup

$$\left\{ \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} : A \in \mathrm{GL}_n(\mathbf{k}) \right\} \leqslant \mathrm{Aff}_n(\mathbf{k})$$

which we will identify with $\mathrm{GL}_n(\mathbf{k})$. The following is a standard notion in group theory.

Definition 1.36

Let G be a group with $H \leqslant G$ and $N \triangleleft G$. Then G is the *semi-direct product* of H and N if $G = HN$ and $H \cap N = \{1\}$; this is often denoted by $G = H \ltimes N$ or $G = N \rtimes H$.

When $G = H \ltimes N$, there is a group isomorphism $q: G/N \longrightarrow H$ as well as the inclusion homomorphism $j: H \longrightarrow G$ and these satisfy $q \circ j = \mathrm{Id}_H$. Notice that H acts on N by conjugation since N is normal in G . The simplest kind of semi-direct product is the *direct product* $H \times N$, where the conjugation action of H on N is trivial.

Proposition 1.37

$\mathrm{Trans}_n(\mathbf{k})$ is a normal subgroup of $\mathrm{Aff}_n(\mathbf{k})$ and $\mathrm{Aff}_n(\mathbf{k})$ can be expressed as the semi-direct product of $\mathrm{Trans}_n(\mathbf{k})$ and $\mathrm{GL}_n(\mathbf{k})$,

$$\mathrm{Aff}_n(\mathbf{k}) = \mathrm{GL}_n(\mathbf{k}) \ltimes \mathrm{Trans}_n(\mathbf{k}) = \{AT : A \in \mathrm{GL}_n(\mathbf{k}), T \in \mathrm{Trans}_n(\mathbf{k})\},$$

with $\mathrm{Trans}_n(\mathbf{k}) \cap \mathrm{GL}_n(\mathbf{k}) = \{I_{n+1}\}$.

Proof

To see that $\text{Trans}_n(\mathbf{k}) \triangleleft \text{Aff}_n(\mathbf{k})$, notice that if

$$\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} \in \text{GL}_n(\mathbf{k}), \quad \begin{bmatrix} I & t \\ 0 & 1 \end{bmatrix} \in \text{Trans}_n(\mathbf{k}),$$

then

$$\begin{aligned} \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} A & At \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} I & At \\ 0 & 1 \end{bmatrix}, \end{aligned}$$

which is in $\text{Trans}_n(\mathbf{k})$. The equality

$$\text{Trans}_n(\mathbf{k}) \cap \text{GL}_n(\mathbf{k}) = \{I_{n+1}\}$$

follows from the fact that non-trivial translations do not fix 0 while all elements of $\text{GL}_n(\mathbf{k})$ do. \square

Orthogonal and Isometry Groups

For $n \geq 1$, an $n \times n$ real matrix A for which $A^T A = I_n$ is called an *orthogonal matrix*; here A^T is the *transpose* of $A = [a_{ij}]$, whose entries are given by

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

Such an orthogonal matrix has an inverse, namely A^T , and the product of two orthogonal matrices A, B is orthogonal since

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

Notice also that $I_n \in O(n)$. So the subset

$$O(n) = \{A \in \text{GL}_n(\mathbb{R}) : A^T A = I_n\} \subseteq M_n(\mathbb{R})$$

is a subgroup of $\text{GL}_n(\mathbb{R})$ and is called the $n \times n$ (*real*) *orthogonal group*. The single matrix equation $A^T A = I_n$ is equivalent to the n^2 equations

$$\sum_{k=1}^n a_{ki} a_{kj} = \delta_{ij} \tag{1.3}$$

for the n^2 real numbers a_{ij} , 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 $O(n)$ is a closed subset of $M_n(\mathbb{R})$ and so is a matrix subgroup of $GL_n(\mathbb{R})$.

Consider the determinant function restricted to $O(n)$, $\det: O(n) \rightarrow \mathbb{R}^\times$. For $A \in O(n)$,

$$(\det A)^2 = \det A^T \det A = \det(A^T A) = \det I_n = 1,$$

which implies that $\det A = \pm 1$. Thus we have

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

where

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

Notice that

$$O(n)^+ \cap O(n)^- = \emptyset,$$

so $O(n)$ is the *disjoint union* of the subsets $O(n)^+$ and $O(n)^-$. The important subgroup

$$SO(n) = O(n)^+ \leq O(n)$$

is the $n \times n$ *special orthogonal group*.

One of the main reasons for the study of the orthogonal groups $O(n)$ and $SO(n)$ is their relationship with *isometries*, where an *isometry* of \mathbb{R}^n is a distance-preserving bijection $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, i.e.,

$$|f(x) - f(y)| = |x - y| \quad (x, y \in \mathbb{R}^n).$$

If such an isometry fixes the origin 0 then it is actually a linear transformation, often referred to as a *linear isometry*, and so with respect to the standard basis it corresponds to a matrix $A \in GL_n(\mathbb{R})$. Here is a more precise statement.

Proposition 1.38

If $A \in GL_n(\mathbb{R})$, then the following conditions are equivalent.

- A is a linear isometry.
- $Ax \cdot Ay = x \cdot y$ for all vectors $x, y \in \mathbb{R}^n$.
- $A^T A = I_n$, i.e., A is orthogonal.

Proof

If A is a linear isometry then for every $\mathbf{v} \in \mathbb{R}^n$, $|A\mathbf{v}| = |\mathbf{v}|$. Now for every pair of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\begin{aligned}|A(\mathbf{x} - \mathbf{y})|^2 &= (A\mathbf{x} - A\mathbf{y}) \cdot (A\mathbf{x} - A\mathbf{y}) \\&= |A\mathbf{x}|^2 + |A\mathbf{y}|^2 - 2A\mathbf{x} \cdot A\mathbf{y} \\&= |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2A\mathbf{x} \cdot A\mathbf{y},\end{aligned}$$

and similarly,

$$\begin{aligned}|A(\mathbf{x} - \mathbf{y})|^2 &= |\mathbf{x} - \mathbf{y}|^2 \\&= (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) \\&= |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2\mathbf{x} \cdot \mathbf{y}.\end{aligned}$$

Hence,

$$A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}.$$

For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$ and $(A\mathbf{u}) \cdot (A\mathbf{v}) = \mathbf{u}^T A^T A \mathbf{v}$. For $i, j = 1, \dots, n$,

$$\mathbf{e}_i^T A^T A \mathbf{e}_j = (i, j) \text{ entry of } A^T A$$

and $\mathbf{e}_i^T \mathbf{e}_j = \delta_{ij}$. Thus $A^T A = I_n$.

Finally, if $A^T A = I_n$ then for each $\mathbf{w} \in \mathbb{R}^n$,

$$\begin{aligned}|Aw|^2 &= (Aw) \cdot (Aw) \\&= \mathbf{w}^T A^T A \mathbf{w} \\&= \mathbf{w}^T \mathbf{w} \\&= |\mathbf{w}|^2,\end{aligned}$$

showing that A is a linear isometry. \square

Elements of $\mathrm{SO}(n)$ are often called *direct isometries* or *rotations*, while elements of $\mathrm{O}(n)^-$ are sometimes called *indirect isometries*. We can also define the full *isometry group* of \mathbb{R}^n ,

$$\mathrm{Isom}_n(\mathbb{R}) = \{f: \mathbb{R}^n \rightarrow \mathbb{R}^n : f \text{ is an isometry}\},$$

which clearly contains the subgroup of translations. In fact,

$$\mathrm{Isom}_n(\mathbb{R}) \leq \mathrm{Aff}_n(\mathbb{R})$$

and is actually a closed subgroup, hence is a matrix subgroup. There is also a semi-direct product decomposition

$$\mathrm{Isom}_n(\mathbb{R}) = \left\{ \begin{bmatrix} A & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} : A \in \mathrm{O}(n), \mathbf{t} \in \mathbb{R}^n \right\}.$$

Proposition 1.39

$\text{Trans}_n(\mathbb{R})$ is a normal subgroup of $\text{Isom}_n(\mathbb{R})$ and $\text{Isom}_n(\mathbb{R})$ can be expressed as the semi-direct product of $\text{Trans}_n(\mathbb{R})$ and $O(n)$,

$$\text{Isom}_n(\mathbb{R}) = O(n) \ltimes \text{Trans}_n(\mathbb{R}) = \{AT : A \in O(n), T \in \text{Trans}_n(\mathbb{R})\},$$

with $\text{Trans}_n(\mathbb{R}) \cap O(n) = \{I_{n+1}\}$.

Proof

This is proved in a similar fashion to Proposition 1.37. \square

For later use, we record some important ideas about elements of $O(n)$. First we recall that a subspace $H \subseteq \mathbb{R}^n$ of dimension $\dim H = (n - 1)$ is called a *hyperplane* in \mathbb{R}^n . Associated with such a hyperplane is a linear transformation $\theta_H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ called *reflection in the hyperplane* H . To define θ_H , observe that every element $x \in \mathbb{R}^n$ can be uniquely expressed as $x = x_H + x'_H$ with $x_H \in H$ and $y \cdot x'_H = 0$ for every $y \in H$. Then

$$\theta_H : \mathbb{R}^n \rightarrow \mathbb{R}^n; \quad \theta_H(x) = x_H - x'_H. \quad (1.4)$$

Lemma 1.40

For a hyperplane $H \subseteq \mathbb{R}^n$, the hyperplane reflection in H is an indirect isometry of \mathbb{R}^n , $\theta_H \in O(n)$.

Proof

This is proved by observing that on choosing an orthonormal basis for H and adjoining a unit vector orthogonal to H , the matrix of θ_H with respect to this basis has the block form

$$\begin{bmatrix} I_{n-1} & O_{(n-1) \times 1} \\ O_{1 \times (n-1)} & -1 \end{bmatrix} = \text{diag}(1, \dots, 1, -1),$$

which clearly has determinant -1 . \square

We will refer to an element of $O(n)$ as a hyperplane reflection if it represents a hyperplane reflection with respect to the standard basis of \mathbb{R}^n , hence it has the form

$$P \text{diag}(1, \dots, 1, -1) P^T$$

for some $P \in O(n)$.

Proposition 1.41

Every element $A \in O(n)$ is a product of hyperplane reflections. The number of these is even if $A \in SO(n)$ and odd if $A \in O(n)^-$.

Proof

We sketch a direct proof, noting that the result also follows from the Principal Axis Theorem 10.13.

We proceed by induction on n . When $n = 1$, $A = \pm 1$ and as $1 = (-1)^2$, the result is true. Now assume that it holds for all $B \in O(k)$ with $k < n$. Suppose that $A \in O(n)$.

If ± 1 is an eigenvalue of A with corresponding unit eigenvector \mathbf{u} , then taking

$$H = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} \cdot \mathbf{u} = 0\},$$

we find that for $\mathbf{v} \in H$,

$$(A\mathbf{v}) \cdot \mathbf{u} = \mathbf{v} \cdot A^T \mathbf{u} = \pm \mathbf{v} \cdot \mathbf{u} = 0,$$

where we have used the fact that

$$A^T \mathbf{u} = A^{-1} \mathbf{u} = \pm \mathbf{u}.$$

Using an orthonormal basis for H extended by \mathbf{u} we see that

$$A = P \begin{bmatrix} A_{n-1} & O_{(n-1) \times 1} \\ O_{1 \times (n-1)} & \pm 1 \end{bmatrix} P^T$$

for some $P \in O(n)$ and $A_{n-1} \in O(n-1)$. By induction, A_{n-1} is a product of hyperplane reflections and so also is A .

We must still consider the case where all the eigenvalues of A have the form $e^{\pm \alpha i}$ with $0 < \alpha < \pi$. Given a unit eigenvector $\mathbf{v} \in \mathbb{C}^n$ for the eigenvalue $e^{\alpha i}$, $\bar{\mathbf{v}}$ is a unit eigenvector for the eigenvalue $e^{-\alpha i}$. With respect to the usual hermitian inner product on \mathbb{C}^n , \mathbf{v} and $\bar{\mathbf{v}}$ are orthogonal. Then

$$\mathbf{v}' = \frac{1}{\sqrt{2}} (\mathbf{v} + \bar{\mathbf{v}}), \quad \mathbf{v}'' = \frac{i}{\sqrt{2}} (\mathbf{v} - \bar{\mathbf{v}})$$

are real, orthogonal unit vectors which also satisfy the equations

$$A\mathbf{v}' = \cos \alpha \mathbf{v}' + \sin \alpha \mathbf{v}'', \quad A\mathbf{v}'' = -\sin \alpha \mathbf{v}' + \cos \alpha \mathbf{v}''.$$

Consider the subspace

$$H' = \{\mathbf{w} \in \mathbb{R}^n : \mathbf{w} \cdot \mathbf{v}' = \mathbf{w} \cdot \mathbf{v}'' = 0\} \subseteq \mathbb{R}^n.$$

An orthonormal basis of H' can be extended to one for \mathbb{R}^n by adding v', v'' . Then with respect to this basis, the matrix of the linear transformation A has the form.

$$\begin{bmatrix} A_{n-2} & O_{(n-2) \times 2} \\ O_{2 \times (n-2)} & R_2(\alpha) \end{bmatrix},$$

where

$$A_{n-2} \in O(n-2), \quad R_2(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \in SO(2).$$

So for some $Q \in O(n)$,

$$A = Q \begin{bmatrix} A_{n-2} & O_{(n-2) \times 2} \\ O_{2 \times (n-2)} & R_2(\alpha) \end{bmatrix} Q^T.$$

By the inductive assumption, A_{n-2} is a product of hyperplane reflections in \mathbb{R}^{n-2} , so it only remains to show that the matrix $R_2(\alpha)$ is a product of hyperplane reflections in \mathbb{R}^2 . By direct calculation we obtain

$$R_2(\alpha) = \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

where the first factor represents reflection in the line

$$\cos(\alpha/2)y = \sin(\alpha/2)x$$

and the second represents reflection in the x -axis, $y = 0$.

The other statements follow from the fact that a hyperplane reflection is an indirect isometry. \square

Generalised Orthogonal Groups

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

$$O_Q = \{A \in GL_n(\mathbb{R}) : A^T Q A = Q\}.$$

It is easy to see that this is a closed subgroup of $GL_n(\mathbb{R})$, i.e., a matrix group. Moreover, if $\det Q \neq 0$, then for $A \in O_Q$ we have $\det A = \pm 1$. We can also define

$$O_Q^+ = \det^{-1} \mathbb{R}^+, \quad O_Q^- = \det^{-1} \mathbb{R}^-$$

and can write O_Q as a disjoint union of clopen subsets $O_Q = O_Q^+ \cup O_Q^-$ 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 $O_Q^+ \cap \mathrm{SL}_2(\mathbb{R})$ which preserves each of the two connected components of the hyperboloid

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

We will study this example in greater detail in Chapter 6.

Symplectic Groups

Symplectic geometry has become an actively studied mathematical topic and is the geometry associated with Hamiltonian Mechanics and therefore with Quantum Mechanics; it is also important as an area of differential geometry and in the study of 4-dimensional manifolds. Symplectic groups are the natural symmetry groups for such geometries. We will discuss the related notion of a *symplectic form* in Section 8.8.

Similar considerations to those in the last section apply to an $n \times n$ real *skew symmetric* matrix S , i.e., one for which $S^T = -S$. For such a matrix,

$$\det S^T = \det(-S) = (-1)^n \det S,$$

giving

$$\det S = (-1)^n \det S. \quad (1.5)$$

The most interesting case occurs if $\det S \neq 0$ when n must be even and we then write $n = 2m$. The standard example of this is built up using the 2×2 block

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

If $m \geq 1$ we have the non-degenerate skew symmetric matrix

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

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*.

There is an alternative version of the symplectic group defined using the skew symmetric matrix

$$J'_{2m} = \begin{bmatrix} O_m & I_n \\ -I_n & O_m \end{bmatrix}$$

in place of J_{2m} . For the precise relationship see the discussion preceding Corollary 8.25. Then we define

$$\text{Symp}'_{2m}(\mathbb{R}) = \{A \in \text{GL}_{2m}(\mathbb{R}) : A^T J'_{2m} A = J'_{2m}\} \leq \text{GL}_{2m}(\mathbb{R}).$$

There is a simple relationship between $\text{Symp}'_{2m}(\mathbb{R})$ and $\text{Symp}_{2m}(\mathbb{R})$, see the discussion at the end of Section 8.8 for further explanation.

Proposition 1.42

Let $P \in \text{GL}_{2m}(\mathbb{R})$ be any matrix for which $J'_{2m} = P^T J_{2m} P$. Then

$$\text{Symp}'_{2m}(\mathbb{R}) = P^{-1} \text{Symp}_{2m}(\mathbb{R}) P = \{P^{-1} AP : A \in \text{Symp}_{2m}(\mathbb{R})\}.$$

Remark 1.43

It is straightforward to see that

$$\text{Symp}_2(\mathbb{R}) = \text{Symp}'_2(\mathbb{R}) = \text{SL}_2(\mathbb{R}),$$

but in general

$$\text{Symp}_{2m}(\mathbb{R}) \neq \text{SL}_{2m}(\mathbb{R}), \quad \text{Symp}'_{2m}(\mathbb{R}) \neq \text{SL}_{2m}(\mathbb{R}).$$

It is also easy to show that $\det A = \pm 1$ if $A \in \text{Symp}_{2m}(\mathbb{R})$ or $A \in \text{Symp}'_{2m}(\mathbb{R})$. In fact the methods of Chapters 8 and 9 can be used to prove that $\det A = 1$, so

$$\text{Symp}_{2m}(\mathbb{R}) \leq \text{SL}_{2m}(\mathbb{R}), \quad \text{Symp}'_{2m}(\mathbb{R}) \leq \text{SL}_{2m}(\mathbb{R}).$$

Unitary Groups

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

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

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

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

Again the unitary condition amounts to n^2 equations for the n^2 complex numbers a_{ij} (compare Equation (1.3)),

$$\sum_{k=1}^n \bar{a}_{ki} a_{kj} = \delta_{ij}. \quad (1.6)$$

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

$$\begin{cases} \sum_{k=1}^n \bar{a}_{ki} a_{kj} = \delta_{ij} & (1 \leq i, j \leq n), \\ \det A = 1. \end{cases} \quad (1.7)$$

Of course, $\det A$ is a polynomial in the entries 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

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

where $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$. Note that \cdot is not \mathbb{C} -linear but for $u, v \in \mathbb{C}$ it satisfies

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

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 $\mathbf{x} = 0$. Then a matrix $A \in \mathrm{M}_n(\mathbb{C})$ is unitary if and only if

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

1.6 Complex Matrices as Real Matrices

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 using one of the following constructions.

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

$$\rho: \mathbb{C} \longrightarrow M_2(\mathbb{R}); \quad \rho(x + yi) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}.$$

This can also be expressed as

$$\rho(x + yi) = xI_2 - yJ$$

where $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ was introduced in Section 1.5. Then ρ is an injective ring homomorphism, so we can view \mathbb{C} as a subring of $M_2(\mathbb{R})$ by identifying \mathbb{C} with its image under ρ ,

$$\rho\mathbb{C} = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R}) : d = a, c = -b \right\}.$$

Notice that complex conjugation corresponds to transposition since

$$\rho(\bar{z}) = \rho(z)^T. \tag{1.8}$$

We will describe two different ways to extend this to a function $M_n(\mathbb{C}) \longrightarrow M_{2n}(\mathbb{R})$. For $Z = [z_{rs}] \in M_n(\mathbb{C})$ with $z_{rs} = x_{rs} + y_{rs}i$, we can write

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

where $X = [x_{rs}]$ and $Y = [y_{rs}]$ are $n \times n$ real matrices.

First we define the function

$$\rho_n: M_n(\mathbb{C}) \longrightarrow M_{2n}(\mathbb{R})$$

for which $\rho_n(Z)$ is the $n \times n$ matrix of 2×2 real blocks with $\rho(z_{rs})$ as the one in the (r, s) place,

$$\rho_n(Z) = \begin{bmatrix} \rho(z_{11}) & \rho(z_{12}) & \cdots & \rho(z_{1n}) \\ \rho(z_{21}) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \rho(z_{n1}) & \cdots & \cdots & \rho(z_{nn}) \end{bmatrix}$$

Then ρ_n is an injective ring homomorphism which is also continuous. In the notation of Section 1.5, this gives

$$\rho_n(iI_n) = -J_{2n}.$$

More generally, for $n \times n$ real matrices X, Y ,

$$\rho_n(X + iY) = \rho_n(X) - J_{2n}\rho_n(Y)J_{2n},$$

where $\rho_n(X)$ and $\rho_n(Y)$ are $2n \times 2n$ real matrices consisting of 2×2 scalar blocks $x_{rs}I_2$ and $y_{rs}I_2$.

Our second function is

$$\rho'_n: M_n(\mathbb{C}) \longrightarrow M_{2n}(\mathbb{R}); \quad \rho'_n(X + iY) = \begin{bmatrix} X & -Y \\ Y & X \end{bmatrix} \quad (X, Y \in M_n(\mathbb{R})),$$

which is an injective ring homomorphism. ρ'_n is easily seen to be continuous. Using the matrix J'_{2m} of Section 1.5, we have the following identities for $X, Y \in M_n(\mathbb{R})$ and $Z \in M_n(\mathbb{C})$:

$$(J'_{2n})^2 = -I_{2n},$$

$$(J'_{2n})^T = -J'_{2n},$$

$$\rho'_n(X + iY) = \begin{bmatrix} X & O_n \\ O_n & X \end{bmatrix} - \begin{bmatrix} Y & O_n \\ O_n & Y \end{bmatrix} J'_{2n},$$

$$\rho'_n(Z) = \rho'_n(Z)^T.$$

The images of ρ_n and ρ'_n , $\rho_n GL_n(\mathbb{C}) \leqslant GL_{2n}(\mathbb{R})$ and $\rho'_n GL_n(\mathbb{C}) \leqslant GL_{2n}(\mathbb{R})$, are clearly closed subgroups of $GL_{2n}(\mathbb{R})$, so any matrix subgroup $G \leqslant GL_n(\mathbb{C})$ can be viewed as a matrix subgroup of $GL_{2n}(\mathbb{R})$ by identifying it with either of its images $\rho_n G$ or $\rho'_n G$. It is sometimes useful to characterise $\rho_n GL_n(\mathbb{C})$ and $\rho'_n GL_n(\mathbb{C})$ as in the following result.

Proposition 1.44

We have

$$\rho_n M_n(\mathbb{C}) = \{A \in M_{2n}(\mathbb{R}): AJ_{2n} = J_{2n}A\}, \tag{1.9a}$$

$$\rho_n GL_n(\mathbb{C}) = \{A \in GL_{2n}(\mathbb{R}): AJ_{2n} = J_{2n}A\}, \tag{1.9b}$$

and

$$\rho'_n M_n(\mathbb{C}) = \{A \in M_{2n}(\mathbb{R}): AJ'_{2n} = J'_{2n}A\}, \tag{1.9c}$$

$$\rho'_n GL_n(\mathbb{C}) = \{A \in GL_{2n}(\mathbb{R}): AJ'_{2n} = J'_{2n}A\}. \tag{1.9d}$$

Proof

We give the proof of the second pair of equations, the first pair being similar.
Writing

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

for some $n \times n$ matrices A_{rs} , we see that $AJ'_{2n} = J'_{2n}A$ if and only if

$$\begin{bmatrix} A_{12} & -A_{11} \\ A_{22} & -A_{21} \end{bmatrix} = \begin{bmatrix} -A_{21} & -A_{22} \\ A_{11} & A_{22} \end{bmatrix},$$

from which we obtain

$$A = \begin{bmatrix} A_{11} & -A_{21} \\ A_{21} & A_{11} \end{bmatrix} \in \rho'_n M_n(\mathbb{C}),$$

giving the result. \square

1.7 Continuous Homomorphisms of Matrix Groups

In studying groups, the notion of a homomorphism of groups plays a central rôle. For matrix groups we need to be careful about topological properties as well as the algebraic ones.

Definition 1.45

Let G, H be two matrix groups. A group homomorphism $\varphi: G \rightarrow H$ is a *continuous homomorphism* of matrix groups if it is continuous and its image $\varphi G \leq H$ is a closed subspace of H .

Example 1.46

The function

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

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

To see why the closure condition on the image in the above definition is desirable, consider the following example.

Example 1.47

Let

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

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

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

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

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

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

Whenever we have a homomorphism of matrix groups $\varphi: G \longrightarrow H$ 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 1.48

Let $\varphi: G \longrightarrow H$ 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 $\bar{\varphi}: G/\ker \varphi \longrightarrow \varphi G$.

Proof

Since φ is continuous, whenever it makes sense in G we have

$$\lim_{n \rightarrow \infty} \varphi(A_n) = \varphi\left(\lim_{n \rightarrow \infty} A_n\right),$$

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 . By Proposition 1.31, $\ker \varphi \leqslant G$ is a matrix group. \square

Remark 1.49

$G/\ker \varphi$ has a natural *quotient topology* discussed in Section 8.1; this is not obviously a metric topology. Then $\bar{\varphi}$ is always a homoeomorphism.

Remark 1.50

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 and 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.

1.8 Matrix Groups for Normed Vector Spaces

Our approach to matrix groups has been in terms of \mathbb{R}^n and \mathbb{C}^n , naturally relying on coordinates and the standard notion of distance on these vector spaces. It is often desirable to generalise this to arbitrary finite dimensional real or complex vector spaces and also to allow a more general notion of distance defined in terms of a *norm*. From now on we assume that $\mathbf{k} = \mathbb{R}$ or \mathbb{C} .

Definition 1.51

Let V be a finite dimensional \mathbf{k} -vector space with a function $\nu: V \rightarrow \mathbb{R}^+$. Then ν is called a \mathbf{k} -*norm* on V if it satisfies the conditions

- i) $\nu(tv) = |t|\nu(v)$ for $t \in \mathbf{k}$, $v \in V$;
- ii) $\nu(v_1 + v_2) \leq \nu(v_1) + \nu(v_2)$ for $v_1, v_2 \in V$;
- iii) for $v \in V$, $\nu(v) = 0$ if and only if $v = 0$.

We denote such a normed vector space by writing (V, ν) .

There are many examples of such norms on \mathbf{k}^n apart from the standard one.

Example 1.52

Consider the function

$$\|\cdot\|_1: \mathbb{k}^n \longrightarrow \mathbb{R}^+; \quad \|\mathbf{x}\|_1 = \max\{|x_i| : i = 1, 2, \dots, n\}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Then $\|\cdot\|_1$ is a \mathbb{k} -norm on \mathbb{k}^n .

Given a finite dimensional normed \mathbb{k} -vector space V , there is a metric ρ_ν on V defined by

$$\rho_\nu(x, y) = \nu(x - y).$$

We can use the associated topology to define continuous functions between V and any other topological space, including other metric spaces. We will regard a normed vector space as a metric space in this way.

The next result is standard and depends on the fact that \mathbb{R} and \mathbb{C} are complete metric spaces.

Theorem 1.53

Let (U, μ) and (V, ν) be two finite dimensional normed \mathbb{k} -vector spaces of the same dimension. Then any linear isomorphism $\varphi: U \longrightarrow V$ is continuous with continuous inverse, i.e., a homeomorphism.

Corollary 1.54

If μ and ν are two \mathbb{k} -norms on a finite dimensional \mathbb{k} -vector space, then they give rise to the same topology.

Proof

The identity function $\text{Id}_V: V \longrightarrow V$ is a linear isomorphism, which must be continuous. This implies that every open set in the topology of ν is open in the topology of μ . Similarly, every open set in the topology of μ is open in the topology of ν . \square

Given a finite dimensional normed \mathbb{k} -vector space (V, ν) and a linear transformation $\alpha: V \longrightarrow V$, we define the *operator norm* of α with respect to ν by

$$\|\alpha\|_\nu = \sup\{\nu(\alpha(x)) : x \in V, \nu(x) = 1\}.$$

If we denote by $\text{End}_k(V)$ the set of all linear transformations $V \rightarrow V$, this defines a k -norm on $\text{End}_k(V)$ in the sense of Proposition 1.5. This makes $\text{End}_k(V)$ into a finite dimensional normed k -vector space with associated topology.

Proposition 1.55

If μ and ν are two norms on finite dimensional k -vector space V , then the topologies associated to $(\text{End}_k(V), \|\cdot\|_\mu)$ and $(\text{End}_k(V), \|\cdot\|_\nu)$ are the same.

Proof

Apply Theorem 1.53. \square

Inside of $\text{End}_k(V)$ we have the group of linear isomorphisms, i.e., invertible linear transformations, $\text{GL}_k(V) \subseteq \text{End}_k(V)$, the *general linear group* of the k -vector space V . If ν is a norm on V , $\text{GL}_k(V)$ inherits the metric and is an open subset of $\text{End}_k(V)$. The following is straightforward to verify.

Proposition 1.56

Let (U, μ) and (V, ν) be two finite dimensional normed k -vector spaces of the same dimension. If $\varphi: U \rightarrow V$ is a linear isomorphism, then it induces a continuous group isomorphism with continuous inverse,

$$\varphi_*: \text{GL}_k(U) \rightarrow \text{GL}_k(V); \quad \varphi_*(\alpha) = \varphi \circ \alpha \circ \varphi^{-1}.$$

Recall that for $V = k^n$ we can identify $\text{End}_k(V)$ with $M_n(k)$ and $\text{GL}_k(V)$ with $\text{GL}_n(k)$. More generally, given a finite basis for V , there is an associated k -linear isomorphism $V \cong k^n$ for some n , a ring isomorphism $\text{End}_k(V) \cong M_n(k)$ and a group isomorphism $\text{GL}_k(V) \cong \text{GL}_n(k)$.

Theorem 1.57

Let (V, ν) be a finite dimensional normed k -vector space (V, ν) of dimension $\dim_k V = n$ and let $\varphi: V \rightarrow k^n$ be a linear isomorphism. Then $\varphi_*: \text{GL}_k(V) \rightarrow \text{GL}_n(k)$ is a continuous group isomorphism with continuous inverse. Hence $\text{GL}_k(V)$ is a matrix group.

This result shows that although we could have set up a theory of matrix groups based on finite dimensional normed vector spaces (V, ν) , defining matrix groups to be closed subgroups of $\text{GL}_k(V)$, we would have obtained an essen-

tially equivalent theory to the one we have described. The extra flexibility is sometimes useful and many accounts are based on it.

We end this section with some useful observations about a finite dimensional normed \mathbb{k} -vector space (V, ν) and a \mathbb{k} -vector subspace $W \subseteq V$.

Proposition 1.58

Let $v \in V$. Then there is a vector $w_0 \in W$ for which

$$\nu(w_0 - v) = \inf\{\nu(w - v) : w \in W\}.$$

Proof

Since $0 \in W$, if we set

$$\lambda_W = \inf\{\nu(w - v) : w \in W\},$$

then $0 \leq \lambda_W \leq \nu(v)$. Now consider the set

$$S = \{w \in W : \nu(w) \leq 2\lambda_W\}.$$

This is a closed and bounded subset of the normed vector space (W, ν) so is compact. By Corollary 1.23, the continuous function

$$f: S \rightarrow \mathbb{R}; \quad f(w) = \nu(w - v)$$

has values bounded below by 0 and attains its infimum. \square

Proposition 1.59

The subset $W \subseteq V$ is closed.

Proof

If $v \notin W$, then $\lambda_W > 0$ since otherwise

$$0 = \lambda_W = \nu(w_0 - v),$$

which can only happen if $w_0 = v$. So the open disc $N_V(v; \lambda_W)$ has trivial intersection with W ,

$$N_V(v; \lambda_W) \cap W = \emptyset.$$

So the complement of W in V is open, hence W is closed. \square

1.9 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 $\mu: G \times X \rightarrow X$ 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 *stabiliser* of x is

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

while the *orbit* of x is

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

Theorem 1.60

Let G act on X .

- For $x \in X$, $\text{Stab}_G(x) \leq G$, i.e., $\text{Stab}_G(x)$ is a subgroup of G .
- For $x, y \in X$, $y \in \text{Orb}_G(x)$ if and only if $\text{Orb}_G(y) = \text{Orb}_G(x)$.

For $x \in X$, there is a bijection

$$\varphi: G / \text{Stab}_G(x) \rightarrow \text{Orb}_G(x); \quad \varphi(g) = gx.$$

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

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

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

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

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

Definition 1.61

Let G be a topological group and X be a topological space. Then a group action $\mu: G \times X \rightarrow X$ is a *continuous group action* if the function μ is continuous.

In this definition, $G \times X$ has the product topology which is obtained from a suitable metric when G and X are metric spaces.

If X is Hausdorff (in particular if it is a metric space) then any one-element subset $\{x\} \subseteq X$ is closed and the stabiliser of x , $\text{Stab}_G(x) \leq G$, is a closed subgroup. This provides a useful source of closed subgroups.

For us, the most important type of action for matrix groups arises when a matrix group G has a continuous homomorphism $\varphi: G \rightarrow \text{GL}_k(V)$ where (V, ν) is a k -norm on the finite dimensional k -vector space V . Then the associated action

$$\mu_\varphi: G \times V \rightarrow V; \quad \mu_\varphi(g, v) = \varphi(g)(v)$$

is continuous. It is worth remarking that by Proposition 1.59, any vector subspace $W \subseteq V$ is closed, so the stabiliser

$$\text{Stab}_G(W) = \{g \in G : \varphi(g)W = W\} \leq G$$

is a closed subgroup, as is

$$\bigcap_{w \in W} \text{Stab}_G(w) \leq G.$$

See the exercises at the end of this chapter for more on this.

Definition 1.62

If $\varphi: G \rightarrow \text{GL}_k(V)$ is a continuous group homomorphism, then the associated action μ_φ is called a (*continuous*) *linear action* or *representation* of G on V .

By choosing a basis for V and applying the ideas discussed at the end of Section 1.8 and Theorem 1.57, we may as well assume that $V = k^n$. Then a continuous action is essentially the same thing as a continuous group homomorphism $G \rightarrow \text{GL}_n(k)$.

Here is an example that illustrates how ubiquitous this idea is. Recall that for indeterminates x_1, \dots, x_k , a polynomial $f(x_1, \dots, x_k) \in k[x_1, \dots, x_k]$ is *homogeneous* of degree n if $f(x_1, \dots, x_k)$ is a linear combination of monomials $x_1^{r_1} \cdots x_k^{r_k}$ with degree

$$r_1 + \cdots + r_k = n.$$

We write $k[x_1, \dots, x_k]_n$ for the subset of all elements of $k[x_1, \dots, x_k]$ homogeneous of degree n . It is easy to see that with addition and scalar multiplication $k[x_1, \dots, x_k]$ is a k -vector space and $k[x_1, \dots, x_k]_n$ is a finite dimensional subspace with a basis consisting of all monomials of degree n .

Example 1.63

For indeterminates x, y , let $V_n = \mathbf{k}[x, y]_n$ with $\dim_{\mathbf{k}} V_n = (n + 1)$. There is a \mathbf{k} -linear action $\mathrm{GL}_2(\mathbf{k}) \rightarrow \mathrm{GL}_{\mathbf{k}}(V_n)$ given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} f(x, y) = f(ax + cy, bx + dy).$$

On a monomial this yields

$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix} x^r y^{n-r} &= (ax + cy)^r (bx + dy)^{n-r} \\ &= \sum_{i=0}^r \sum_{j=0}^{n-r} \binom{r}{i} \binom{n-r}{j} (ax)^i (cy)^{r-i} (bx)^j (dy)^{n-r-j} \\ &= \sum_{k=0}^n \left(\sum_{i=0}^k \binom{r}{i} \binom{n-r}{k-i} a^i c^{r-i} b^{k-i} d^{n-r-k+i} \right) x^k y^{n-k}, \end{aligned}$$

giving for example,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} x^2 y = a^2 b x^3 + (a^2 d + 2abc)x^2 y + (2acd + c^2 b)xy^2 + c^2 d y^3.$$

Taking the basis of monomials of degree n in the order

$$x^n, x^{n-1}y, \dots, xy^{n-1}, y^n,$$

this can be interpreted as a homomorphism $\varphi_n: \mathrm{GL}_2(\mathbf{k}) \rightarrow \mathrm{GL}_{n+1}(\mathbf{k})$.

Of course, any matrix subgroup $G \leq \mathrm{GL}_2(\mathbf{k})$ will also act on V_n by a continuous linear action. When $\mathbf{k} = \mathbb{C}$, the representations of $\mathrm{SU}(2)$ on the V_n are particularly important, see Sternberg [27] for an illuminating discussion of these representations and their applications.

EXERCISES

- 1.1. Determine the norm $\|A\|$ for each of the following matrices A and real numbers $t, u, v \in \mathbb{R}$.

$$\begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix}, \quad \begin{bmatrix} u & 1 \\ 0 & u \end{bmatrix}, \quad \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}, \quad \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}.$$

What can be said when $u, v \in \mathbb{C}$?

1.2. Suppose that $A \in M_n(\mathbb{C})$.

a) If $B \in U(n)$, show that $\|BAB^{-1}\| = \|A\|$.

b) For a general element $C \in GL_n(\mathbb{C})$, what can be said about $\|CAC^{-1}\|$?

1.3. [This problem expands on Remark 1.4 and requires knowledge of the diagonalisation of hermitian matrices.] Let $A \in M_n(\mathbb{C})$.

a) Show that

$$\begin{aligned}\|A\|^2 &= \sup\{\mathbf{x}^* A^* A \mathbf{x} : \mathbf{x} \in \mathbb{C}^n, |\mathbf{x}| = 1\} \\ &= \max\{\mathbf{x}^* A^* A \mathbf{x} : \mathbf{x} \in \mathbb{C}^n, |\mathbf{x}| = 1\}.\end{aligned}$$

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 , $\|A\| = |Av|$.

1.4. If $\{A_r\}_{r \geq 0}$ is a sequence of matrices $A_r \in M_n(\mathbb{k})$, prove the following version of the *ratio test*.

- a) If $\lim_{r \rightarrow \infty} \frac{\|A_{r+1}\|}{\|A_r\|} < 1$, the series $\sum_{r=0}^{\infty} A_r$ converges in $M_n(\mathbb{k})$.
- b) If $\lim_{r \rightarrow \infty} \frac{\|A_{r+1}\|}{\|A_r\|} > 1$, the series $\sum_{r=0}^{\infty} A_r$ diverges in $M_n(\mathbb{k})$.
- c) Develop other convergence tests for $\sum_{r=0}^{\infty} A_r$.

1.5. Suppose that $A \in M_n(\mathbb{k})$ and $\|A\| < 1$.

a) Show that the series

$$\sum_{r=0}^{\infty} A^r = I + A + A^2 + A^3 + \dots$$

converges in $M_n(\mathbb{k})$.

- b) Show that $(I - A)$ is invertible and find a formula for $(I - A)^{-1}$.
- c) If A is 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 $O(n)$ is compact.

b) Show that the set of all $n \times n$ unitary matrices $U(n)$ is compact.

c) Show that $GL_n(\mathbb{k})$ and $SL_n(\mathbb{k})$ are not compact if $n \geq 2$.

d) Investigate which of the other matrix groups of Section 1.5 are compact.

1.7. Recall that in a topological space X , the *closure* \overline{U} of a subset $U \subseteq X$ is the smallest closed subset $V \subseteq X$ for which $U \subseteq V$. Then $U \subseteq X$ is closed in X if and only if $\overline{U} = U$. It is also useful to note that if $u \in U$, then every open set W containing u intersects U .

- If G is a matrix group and $H \leq G$ is a subgroup, show that the closure $\overline{H} \subseteq G$ of H in G is also a subgroup.
- Generalise (a) to an arbitrary topological group G .
- Let

$$\Gamma = \{A \in GL_n(\mathbb{R}) : \det A \in \mathbb{Q}\} \leq GL_n(\mathbb{R}).$$

Show that Γ is not a closed subgroup of $GL_n(\mathbb{R})$ and find its closure $\overline{\Gamma}$ in $GL_n(\mathbb{R})$.

- Show that a compact subgroup $G \leq GL_n(\mathbb{R})$ is a matrix group.
- Let G be a compact matrix group and suppose that the matrix subgroup $H \leq G$ is discrete as a subspace of G , i.e., each singleton subset $\{h\} \subseteq H$ is open in H . Show that H is finite.
- Using Example 1.35, verify each of the following for $n \geq 1$.
 - $O(n)$ is a matrix subgroup of $O(n+1)$;
 - $SO(n)$ is a matrix subgroup of $SO(n+1)$;
 - $U(n)$ is a matrix subgroup of $U(n+1)$;
 - $SU(n)$ is a matrix subgroup of $SU(n+1)$.
- a) If $A \in \text{Symp}_{2m}(\mathbb{R})$, prove that $\det A = \pm 1$.
b) Prove that $\text{Symp}_2(\mathbb{R}) = SL_2(\mathbb{R})$.
- Recall Section 1.6 and in particular Proposition 1.44. Verify the following sets of equalities:

$$\begin{aligned}\rho_n U(n) &= O(2n) \cap \rho_n GL_n(\mathbb{C}) \\ &= O(2n) \cap \text{Symp}_{2n}(\mathbb{R}) \\ &= \rho_n GL_n(\mathbb{C}) \cap \text{Symp}_{2n}(\mathbb{R}), \\ \rho'_n U(n) &= O(2n) \cap \rho'_n GL_n(\mathbb{C}) \\ &= O(2n) \cap \text{Symp}'_{2n}(\mathbb{R}) \\ &= \rho'_n GL_n(\mathbb{C}) \cap \text{Symp}'_{2n}(\mathbb{R}),\end{aligned}$$

- Let $(X_1, \rho_1), (X_2, \rho_2)$ be two metric spaces. Consider the function

$$\begin{aligned}\rho: (X_1 \times X_2) \times (X_1 \times X_2) &\longrightarrow \mathbb{R}^+; \\ \rho((x_1, x_2), (y_1, y_2)) &= \sqrt{\rho_1(x_1, y_1)^2 + \rho_2(x_2, y_2)^2}.\end{aligned}$$

a) Show that $(X_1 \times X_2, \rho)$ is a metric space whose open sets are the same as those of the product topology.

b) Show that a sequence $\{(x_{1,r}, x_{2,r})\}_{r \geq 0}$ converges (i.e., has a limit) in $X_1 \times X_2$ if and only if the sequences $\{(x_{1,r})\}_{r \geq 0}, \{(x_{2,r})\}_{r \geq 0}$ converge in X_1 and X_2 respectively.

1.14. Let $\mu: G \times X \rightarrow X$ be a continuous group action as in Definition 1.61 and assume that X is Hausdorff (for example a metric space).

a) If $x \in X$, show that the stabiliser

$$\text{Stab}_G(x) = \{g \in G : gx = x\}$$

is a closed subgroup of G .

b) If $W \subseteq X$ is a closed subset, show that

$$\text{Stab}_G(W) = \{g \in G : gW = W\} \leqslant G, \quad \bigcap_{w \in W} \text{Stab}_G(w) \leqslant G,$$

are closed subgroups, where $gW = \{gw : w \in W\}$.

1.15. Let $\mathbf{k} = \mathbb{R}$ or \mathbb{C} .

a) Making use of a suitable metric ρ on the product space $M_n(\mathbf{k}) \times \mathbf{k}^n$, show that the product map

$$\varphi: M_n(\mathbf{k}) \times \mathbf{k}^n \rightarrow \mathbf{k}^n; \quad \varphi(A, x) = Ax,$$

is continuous.

b) Let $G \leqslant \text{GL}_n(\mathbf{k})$ be a matrix subgroup. By restricting the metric ρ and product map φ of (a) to the subset $G \times \mathbf{k}^n$, consider the resulting continuous group action of G on \mathbf{k}^n . Show that the stabiliser of $x \in \mathbf{k}^n$,

$$\text{Stab}_G(x) = \{A \in G : Ax = x\},$$

is a matrix subgroup of G . More generally, if $X \subseteq \mathbf{k}^n$ is a closed subset, show that

$$\text{Stab}_G(X) = \{A \in G : AX = X\}$$

is a matrix subgroup of G , where $AX = \{Ax : x \in X\}$.

c) For the standard basis vector $e_n \in \mathbf{k}^n$ and

$$X = \{te_n : t \in \mathbb{R}\} \subseteq \mathbf{k}^n,$$

determine $\text{Stab}_G(e_n)$ and $\text{Stab}_G(X)$ for each of the following matrix subgroups $G \leqslant \text{GL}_n(\mathbb{R})$: $\text{GL}_n(\mathbb{R}), \text{SL}_n(\mathbb{R}), \text{O}(n), \text{SO}(n)$.

1.16. Let $n \geq 1$. Consider the \mathbb{C} -linear action of $SU(2)$ on the $(n+1)$ -dimensional complex vector space of $V_n = \mathbb{C}[x, y]_n$ introduced in Example 1.63. Let $\varphi_n: SU(2) \rightarrow GL_{n+1}(\mathbb{C})$ be the continuous homomorphism obtained by working with matrices relative to the basis of monomials in x and y ordered by increasing degree in x .

- Determine φ_n explicitly for small values of n .
- Determine the stabiliser of $x^n \in V_n$.
- When $n = 2$, determine the stabiliser of $xy \in V_2$.
- Show that

$$\ker \varphi_n = \begin{cases} \{I\} & \text{if } n \text{ is odd,} \\ \{I, -I\} & \text{if } n \text{ is even.} \end{cases}$$

Exponentials, Differential Equations and One-parameter Subgroups

The matrix versions of the familiar real and complex exponential and logarithm functions are fundamental for the study of many aspects of matrix group theory, particularly the *one-parameter subgroups*. Indeed, the *matrix exponential function* provides the link between the *Lie algebra* of a matrix group and the group itself. In the case of a compact connected matrix group, the exponential is even surjective, allowing a parametrisation of such a group by a region in \mathbb{R}^n for some n ; see Chapter 10 for details. Just as in the theory of ordinary differential equations, matrix exponential functions also play a central rôle in the theory of certain types of differential equations for matrix-valued functions and these are important in many applications of Lie theory.

2.1 The Matrix Exponential and Logarithm

Throughout this section, we will assume that $\mathbf{k} = \mathbb{R}$ or \mathbb{C} . The power series

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

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

Let $A \in M_n(\mathbf{k})$. The matrix-valued series

$$\text{Exp}(A) = \sum_{n \geq 0} \frac{1}{n!} A^n = I + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \dots,$$

$$\text{Log}(A) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} A^n = A - \frac{1}{2} A^2 + \frac{1}{3} A^3 - \frac{1}{4} A^4 + \dots,$$

converge whenever $\|A\| < r.o.c.$. So $\text{Exp}(A)$ makes sense for every $A \in M_n(\mathbf{k})$ while $\text{Log}(A)$ only exists if $\|A\| < 1$.

Proposition 2.1

Let $A \in M_n(\mathbf{k})$.

- i) For $u, v \in \mathbf{C}$, $\text{Exp}((u+v)A) = \text{Exp}(uA) \text{Exp}(vA)$.
- ii) $\text{Exp}(A) \in \text{GL}_n(\mathbf{k})$ and $\text{Exp}(A)^{-1} = \text{Exp}(-A)$.

Proof

(i) By expanding the first series we obtain

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

By a sequence of obvious manipulations that are justified since these series are all absolutely convergent,

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

(ii) From part (i),

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

so $\text{Exp}(A)$ is invertible with inverse $\text{Exp}(-A)$. \square

Using these series we can define the matrix version of the *exponential function*:

$$\exp: M_n(\mathbf{k}) \longrightarrow GL_n(\mathbf{k}); \quad \exp(A) = \text{Exp}(A).$$

Proposition 2.2

If $A, B \in M_n(\mathbf{k})$ commute then

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

Proof

Again we expand the series and perform a sequence of manipulations, all of which are justified because these series are absolutely convergent. The result is

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

Of course the identity

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

depends crucially on the fact that A and B commute. \square

We define the *logarithm function* by

$$\log: N_{M_n(k)}(I; 1) \longrightarrow M_n(k); \quad \log(A) = \text{Log}(A - I).$$

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

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

Proposition 2.3

The functions \exp and \log satisfy

- i) if $\|A - I\| < 1$, then $\exp(\log(A)) = A$;
- ii) if $\|\exp(B) - I\| < 1$, then $\log(\exp(B)) = B$.

Proof

These results follow from the formal identities between power series

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

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

which are proved by comparing coefficients. \square

The functions \exp and \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

$$N_{M_n(k)}(O; \delta_1) \subseteq \exp^{-1} N_{GL_n(k)}(I; 1).$$

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

$$\exp N_{M_n(k)}(O; r) \subseteq N_{M_n(k)}(I; e^r - 1).$$

Hence we have the following results.

Proposition 2.4

The exponential function \exp is injective when restricted to the open subset $N_{M_n(k)}(O; \ln 2) \subseteq M_n(k)$, 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 $A \in M_n(\mathbf{k})$. When $B \in M_n(\mathbf{k})$ commutes with A ,

$$\begin{aligned}\frac{d}{dt}_{|t=0} \exp(A + tB) &= \lim_{h \rightarrow 0} \frac{1}{h} (\exp(A + hB) - \exp(A)) \\ &= \exp(A)B = B\exp(A).\end{aligned}\quad (2.1)$$

The general situation is more complicated.

For a variable X consider the series

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

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

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

to elements of $M_n(\mathbf{C})$. In particular we can consider

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

where

$$\text{ad } A: M_n(\mathbf{C}) \longrightarrow M_n(\mathbf{C}); \quad \text{ad } A(C) = AC - CA,$$

is viewed as a \mathbf{C} -linear operator which acts on $M_n(\mathbf{C})$ by the *adjoint action* of A . Then

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

Proposition 2.5

For $A, B \in M_n(\mathbf{C})$ we have

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

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

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

Proof

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

$$F(D)|_{s=0} f(s) = \int_0^1 f(s) \, ds. \quad (2)$$

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

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

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

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

Evaluating at $s = 0$ gives Equation (2.2).

The matrix-valued function

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

satisfies

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

since for $m, n \geq 1$,

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

So

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

giving

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

which is the desired formula.

2.2 Calculating Exponentials and Jordan Form

Since exponentials of matrices occur throughout this subject matter, it is important to be able to calculate them. It is an easy exercise to show that for $A \in M_n(\mathbb{C})$ and $B \in GL_n(\mathbb{C})$,

$$\exp(BAB^{-1}) = B \exp(A)B^{-1}.$$

When A is *diagonalisable*, i.e., $A = C \text{diag}(\lambda_1, \dots, \lambda_n)C^{-1}$ for some $C \in GL_n(\mathbb{C})$, we have

$$\begin{aligned}\exp(A) &= C \exp(\text{diag}(\lambda_1, \dots, \lambda_n))C^{-1} \\ &= C \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})C^{-1},\end{aligned}\tag{2.3}$$

since

$$\exp(\text{diag}(\lambda_1, \dots, \lambda_n)) = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}).$$

This means that the problem of calculating the exponential of a diagonalisable matrix is solved once an explicit diagonalisation is found. Many important types of matrices are indeed diagonalisable, including (skew) symmetric, (skew) hermitian, orthogonal, unitary and normal matrices. However, there are also many non-diagonalisable matrices. The general situation is best discussed in terms of the *Jordan form*, a good reference for which is Strang [28], although many books on linear algebra contain accounts of this material. Similar results hold over any algebraically closed field but we work over the field of complex numbers \mathbb{C} .

We start by recalling that for a matrix $A \in M_n(\mathbb{C})$, the *characteristic polynomial* of A is the monic polynomial $\text{char}_A(X) \in \mathbb{C}[X]$ of degree n given by

$$\text{char}_A(X) = \det(XI_n - A).$$

We will sometimes write

$$\text{char}_A(X) = X^n + c_{n-1}(A)X^{n-1} + \cdots + c_1(A)X + c_0(A),$$

where $c_k(A) \in \mathbb{C}$. We recall the following important result.

Theorem 2.6 (Cayley–Hamilton Theorem)

A complex matrix $A \in M_n(\mathbb{C})$ satisfies its own characteristic polynomial, i.e.,

$$\text{char}_A(A) = A^n + c_{n-1}(A)A^{n-1} + \cdots + c_1(A)A + c_0(A)I_n = O_n.$$

We also recall the *minimal polynomial* of A , $\min_A(X) \in \mathbb{C}[X]$. This is the non-zero monic polynomial of minimal degree for which

$$\min_A(A) = O_n.$$

This always exists and has the following property.

Proposition 2.7

If $f(X) \in \mathbb{C}[X]$ satisfies $f(A) = O_n$ then $\min_A(X) \mid f(X)$, i.e., there is a $g(X) \in \mathbb{C}[X]$ such that $\min_A(X)g(X) = f(X)$.

Corollary 2.8

$\min_A(X) \mid \text{char}_A(X)$, hence $\deg \min_A(X) \leq n$.

The complex roots of $\text{char}_A(X)$ are the eigenvalues of A , so every root of $\min_A(X)$ is an eigenvalue of A . In fact the converse is also true. So if the *distinct* eigenvalues of A are $\lambda_1, \dots, \lambda_d$ and

$$\text{char}_A(X) = (X - \lambda_1)^{m_1} \cdots (X - \lambda_d)^{m_d}$$

with $m_j \geq 1$, then

$$\min_A(X) = (X - \lambda_1)^{m'_1} \cdots (X - \lambda_d)^{m'_d}$$

where $m_j \geq m'_j \geq 1$.

For $\lambda \in \mathbb{C}$ and $r \geq 1$, we have the *Jordan block matrix*

$$J(\lambda, r) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & 0 & \lambda & \end{bmatrix} \in M_r(\mathbb{C}).$$

The characteristic polynomial of $J(\lambda, r)$ is

$$\text{char}_{J(\lambda, r)}(X) = \det(XI_r - J(\lambda, r)) = (X - \lambda)^r \quad (2.4a)$$

and by the Cayley-Hamilton Theorem 2.6,

$$(J(\lambda, r) - \lambda I_r)^r = O_r.$$

Notice that

$$(J(\lambda, r) - \lambda I_r)^{r-1} e_r = e_1,$$

which implies that

$$(J(\lambda, r) - \lambda I_r)^{r-1} \neq O_r.$$

Hence we also have

$$\min_{J(\lambda, r)}(X) = (X - \lambda)^r. \quad (2.4b)$$

The main result on Jordan form is the following.

Theorem 2.9

Let $A \in M_n(\mathbb{C})$. Then there is an invertible matrix $P \in GL_n(\mathbb{C})$ for which $A = P J(A) P^{-1}$ and $J(A) \in M_n(\mathbb{C})$ is the matrix having block form

$$J(A) = \begin{bmatrix} J(\lambda_1, r_1) & O & O & \cdots & O \\ O & J(\lambda_2, r_2) & O & \cdots & O \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ O & \cdots & \cdots & O & J(\lambda_m, r_m) \end{bmatrix}.$$

This form is unique except for the order in which the Jordan blocks $J(\lambda_i, r_i)$ occur.

Corollary 2.10

A is diagonalisable if and only if $A = P J(A) P^{-1}$ where $J(A)$ has $r_1 = \cdots = r_m = 1$.

In this *Jordan form* for A , the λ_i are eigenvalues of A and in fact the characteristic polynomial of A is

$$\text{char}_A(X) = (X - \lambda_1)^{r_1} \cdots (X - \lambda_m)^{r_m}. \quad (2.5)$$

However the minimal polynomial is more complicated to specify. First notice that in general the λ_i are not all distinct. For example, the matrix

$$A = \begin{bmatrix} 7 & 0 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} J(7, 1) & O & 0 \\ O & J(7, 2) & O \\ 0 & O & J(5, 1) \end{bmatrix}$$

has two Jordan blocks for the eigenvalue 7. In this case, the characteristic polynomial is

$$\text{char}_A(X) = (X - 7)^1 (X - 7)^2 (X - 5) = (X - 7)^3 (X - 5),$$

while the minimal polynomial is

$$\min_A(X) = (X - 7)^2(X - 5).$$

Let us calculate the exponential of this matrix. We have

$$\exp(A) = \begin{bmatrix} \exp(J(7, 1)) & 0 & 0 \\ 0 & \exp(J(7, 2)) & 0 \\ 0 & 0 & \exp(J(5, 1)) \end{bmatrix},$$

so it suffices to determine $\exp(J(7, 2))$. Since for $k \geq 1$

$$J(7, 2)^k = \begin{bmatrix} 7^k & k7^{k-1} \\ 0 & 7^k \end{bmatrix},$$

we find

$$\begin{aligned} \exp(J(7, 2)) &= \sum_{k=0}^{\infty} \frac{1}{k!} J(7, 2)^k \\ &= I_2 + \sum_{k=1}^{\infty} \frac{1}{k!} \begin{bmatrix} 7^k & k7^{k-1} \\ 0 & 7^k \end{bmatrix} \\ &= \text{diag}(e^7, e^7) + \begin{bmatrix} 0 & \sum_{k=1}^{\infty} 7^{k-1}/(k-1)! \\ 0 & 0 \end{bmatrix} \\ &= e^7 I_2 + e^7 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= e^7 \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

So we have

$$\exp(A) = \begin{bmatrix} e^7 & 0 & 0 & 0 \\ 0 & e^7 & e^7 & 0 \\ 0 & 0 & e^7 & 0 \\ 0 & 0 & 0 & e^5 \end{bmatrix}.$$

In the general case, for each eigenvalue λ_j of A , the factor $(X - \lambda_j)^{m'_j}$ of $\min_A(X)$ is determined by taking

$$m'_j = \min\{r_i : \lambda_i = \lambda_j\}. \quad (2.6)$$

2.3 Differential Equations in Matrices

Definition 2.11

A *differentiable curve* in $M_n(\mathbf{k})$ is a function

$$\alpha: (a, b) \longrightarrow M_n(\mathbf{k})$$

for which the derivative $\alpha'(t)$ exists for each $t \in (a, b)$. Here $\alpha'(t)$ is defined as an element of $M_n(\mathbf{k})$ by

$$\alpha'(t) = \lim_{s \rightarrow t} \frac{1}{(s - t)} (\alpha(s) - \alpha(t)),$$

provided this limit exists.

Let $A \in M_n(\mathbb{R})$. 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 notation

$$\alpha'(t) = \frac{d}{dt} \alpha(t)$$

for the derivative of α .

Consider the first order differential equation

$$\alpha'(t) = \alpha(t)A, \tag{2.7}$$

in which $\alpha: (a, b) \longrightarrow M_n(\mathbb{R})$ is assumed to be a differentiable curve.

If $n = 1$ then taking $A = a$ 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, namely the function of t given by the power series

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

This is indicative of the general case $n \geq 1$.

Theorem 2.12

For $A, C \in M_n(\mathbb{R})$ with A non-zero, and $a < 0 < b$, the differential equation of (2.7) has a unique solution $\alpha: (a, b) \longrightarrow M_n(\mathbb{R})$ for which $\alpha(0) = C$. Furthermore, if C is invertible then so is $\alpha(t)$ for $t \in (a, b)$, hence $\alpha: (a, b) \longrightarrow GL_n(\mathbb{R})$.

Proof

First we will solve the equation subject to the boundary condition $\alpha(0) = I$. By Section 2.1, for each $t \in (a, b)$, the series

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

converges, so the function

$$\alpha: (a, b) \longrightarrow M_n(\mathbb{R}); \quad \alpha(t) = \exp(tA),$$

is defined and differentiable with

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

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

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

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 then so is $C \exp(tA)$ for all t . \square

2.4 One-parameter Subgroups in Matrix Groups

Let $G \leqslant \mathrm{GL}_n(\mathbb{k})$ be a matrix group. Since $G \subseteq M_n(\mathbb{k})$, the next definition is an obvious modification of Definition 2.11, in particular the derivative is defined by

$$\gamma'(t) = \lim_{s \rightarrow t} \frac{1}{(s-t)} (\gamma(s) - \gamma(t)) \in M_n(\mathbb{k}),$$

provided this limit exists.

Definition 2.13

A *differentiable curve* in G is a function $\gamma: (a, b) \rightarrow G$ for which the derivative $\gamma'(t)$ exists at each $t \in (a, b)$.

Since $G \subseteq M_n(\mathbf{k})$ such a curve is also a curve in $M_n(\mathbf{k})$. We will usually assume that $a < 0 < b$ when considering such curves.

Now suppose that $\varepsilon > 0$ or $\varepsilon = \infty$.

Definition 2.14

A *one-parameter semigroup in G* is a continuous function $\gamma: (-\varepsilon, \varepsilon) \rightarrow G$ which is differentiable at 0 and also 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 $\gamma: \mathbb{R} \rightarrow G$ 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 2.15

Let $\gamma: (-\varepsilon, \varepsilon) \rightarrow G$ be a one-parameter semigroup in G . Then for every $t \in (-\varepsilon, \varepsilon)$, γ is differentiable at t and

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

Proof

For small $h \in \mathbb{R}$,

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

Hence

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

and similarly $\gamma'(t) = \gamma(t)\gamma'(0)$. □

Of course, part of the importance of this result is the implication that γ is a differentiable curve in G even though we only assumed it was continuous.

Proposition 2.16

Let $\gamma: (-\varepsilon, \varepsilon) \rightarrow G$ be a one-parameter semigroup in G . Then there is a unique extension to a one-parameter group $\tilde{\gamma}: \mathbb{R} \rightarrow G$ in G , i.e., a function $\tilde{\gamma}$ for which $\tilde{\gamma}(t) = \gamma(t)$ for all $t \in (-\varepsilon, \varepsilon)$.

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 , $\gamma(t/n), \gamma(t/n)^n \in G$. Since $mn \geq m$ and $mn \geq n$, we also have $t/mn \in (-\varepsilon, \varepsilon)$ and therefore

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

Thus $\gamma(t/n)^n = \gamma(t/m)^m$, which shows that we obtain a well-defined element of G for every real number t . This defines a function

$$\tilde{\gamma}: \mathbb{R} \rightarrow G; \quad \tilde{\gamma}(t) = \gamma(t/n)^n \quad \text{for large } n.$$

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

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

Theorem 2.17

Let $\gamma: \mathbb{R} \rightarrow G$ be a one-parameter group in G . Then it has the form

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

for some $A \in M_n(\mathbf{k})$.

Proof

Let $A = \gamma'(0)$. By Proposition 2.15, γ satisfies the differential equation

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

By Theorem 2.12, this equation has the unique solution $\gamma(t) = \exp(tA)$. \square

Remark 2.18

We cannot yet reverse this process and decide for which $A \in M_n(\mathbf{k})$ the one-parameter group

$$\gamma: \mathbb{R} \longrightarrow GL_n(\mathbf{k}); \quad \gamma(t) = \exp(tA)$$

actually takes values in G . The answer involves the *Lie algebra* of G which will be defined in Chapter 3. Notice that we also have the curious phenomenon 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 sometimes merely continuity) often lead to much stronger conclusions.

2.5 One-parameter Subgroups and Differential Equations

In this section we show how ideas about one-parameter subgroups can be applied to solving differential equations. A good source for applications of linear algebra is the book of Strang [28].

Consider the following differential equation for a function $v: \mathbb{R} \longrightarrow \mathbf{C}^n$ which is assumed to be differentiable:

$$v'(t) = Av(t), \quad v(0) = v_0. \tag{2.8}$$

Here $A \in M_n(\mathbf{C})$. In examples we will often find that solutions take values in \mathbb{R}^n , however it is more convenient to work over \mathbf{C} since eigenvalues and eigenvectors are often involved in calculations.

If $v_0 \neq 0$, then we can look for a solution of the form

$$v(t) = \alpha(t)v_0, \tag{2.9}$$

where $\alpha: \mathbb{R} \longrightarrow GL_n(\mathbf{k})$ is a one-parameter subgroup as discussed in Section 2.4 and satisfying

$$\alpha'(t) = \alpha(t)A.$$

By Theorem 2.17, for $t \in \mathbb{R}$ we have

$$\alpha(t) = \exp(tA),$$

hence a solution of Equation (2.8) is

$$\mathbf{v}(t) = \exp(tA)\mathbf{v}_0. \quad (2.10)$$

In fact, this is the unique solution.

Example 2.19

For a skew symmetric matrix $S \in M_n(\mathbb{R})$ and non-zero $\mathbf{v}_0 \in \mathbb{R}^n$, the differential equation

$$\mathbf{v}'(t) = S\mathbf{v}(t), \quad \mathbf{v}(0) = \mathbf{v}_0.$$

has a solution for which

$$|\mathbf{v}(t)| = |\mathbf{v}_0| \quad (t \in \mathbb{R}).$$

Proof

By Equation (3.13), for all $t \in \mathbb{R}$ we have $\exp(tS) \in SO(n)$, hence

$$|\mathbf{v}(t)| = |\exp(tS)\mathbf{v}(0)| = |\mathbf{v}(0)|.$$

Notice that since $|\mathbf{v}(t)|^2 = |\mathbf{v}(0)|^2$ is a constant,

$$\begin{aligned} 2\mathbf{v}(t) \cdot \mathbf{v}'(t) &= \mathbf{v}(t) \cdot \mathbf{v}'(t) + \mathbf{v}'(t) \cdot \mathbf{v}(t) \\ &= \frac{d}{dt}(\mathbf{v}(t) \cdot \mathbf{v}(t)) \\ &= \frac{d}{dt}|\mathbf{v}(t)|^2 \\ &= 0. \end{aligned}$$

This shows that $\mathbf{v}'(t)$ is tangent to the sphere centred at the origin and of radius $|\mathbf{v}(0)|$ at the point $\mathbf{v}(t)$. \square

Here is an explicit example of this type.

Example 2.20

When $n = 3$ and

$$S = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & -2 & 0 \end{bmatrix},$$

the differential equation

$$\mathbf{v}'(t) = S\mathbf{v}(t), \quad \mathbf{v}(0) = \mathbf{v}_0,$$

has the solution

$$\mathbf{v}(t) = \frac{1}{5} \begin{bmatrix} 4 + \cos \sqrt{5}t & \sqrt{5} \sin \sqrt{5}t & 2 - 2 \cos \sqrt{5}t \\ -\sqrt{5} \sin \sqrt{5}t & \cos \sqrt{5}t & 2\sqrt{5} \sin \sqrt{5}t \\ 2 - 2 \cos \sqrt{5}t & -2\sqrt{5} \sin \sqrt{5}t & 1 + 4 \cos \sqrt{5}t \end{bmatrix} \mathbf{v}_0.$$

Proof

One approach to computing $\exp(tS)$ involves determining the eigenvalues of S and then diagonalising it over \mathbb{C} ; the details are left as an exercise for the reader. We obtain the special orthogonal matrix

$$\exp(tS) = \frac{1}{5} \begin{bmatrix} 4 + \cos \sqrt{5}t & \sqrt{5} \sin \sqrt{5}t & 2 - 2 \cos \sqrt{5}t \\ -\sqrt{5} \sin \sqrt{5}t & \cos \sqrt{5}t & 2\sqrt{5} \sin \sqrt{5}t \\ 2 - 2 \cos \sqrt{5}t & -2\sqrt{5} \sin \sqrt{5}t & 1 + 4 \cos \sqrt{5}t \end{bmatrix},$$

giving for the required solution

$$\mathbf{v}(t) = \frac{1}{5} \begin{bmatrix} 2 - 2 \cos \sqrt{5}t \\ 2\sqrt{5} \sin \sqrt{5}t \\ 1 + 4 \cos \sqrt{5}t \end{bmatrix}.$$

□

When the matrix A in Equation (2.8) is diagonalisable, this approach works well, provided an explicit diagonalisation can actually be found. In particular, symmetric, skew symmetric, hermitian and skew hermitian matrices are always diagonalisable over \mathbb{C} , as are matrices without multiple eigenvalues. For more general matrices with multiple eigenvalues it may be necessary to use Jordan forms as discussed in Section 2.2. This is illustrated in the next example.

Example 2.21

If we take

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \mathbf{v}_0 \neq 0,$$

then the differential equation

$$\mathbf{v}'(t) = A\mathbf{v}(t), \quad \mathbf{v}(0) = \mathbf{v}_0$$

has solution

$$\begin{bmatrix} e^{2t} & te^{2t} & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^t \end{bmatrix} \mathbf{v}_0.$$

Proof

This follows from Equation (3.6). Here we take

$$D = \text{diag}(2t, 2t, 2t), \quad N = \begin{bmatrix} 0 & t & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and find that $N^2 = O$. This gives

$$\begin{aligned} \exp(tA) &= e^{2t}I_3 + \sum_{k \geq 0} \frac{1}{(k+1)!} (k+1)N \text{diag}(2t, 2t, 2t)^k \\ &= e^{2t}I_3 + N \sum_{k \geq 0} \frac{1}{k!} \text{diag}(2t, 2t, 2t)^k \\ &= \begin{bmatrix} e^{2t} & te^{2t} & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^t \end{bmatrix}. \end{aligned}$$

This gives the claimed solution. \square

Now suppose we have an equation of the form Equation (2.8). Write

$$\mathbf{v}(t) = \begin{bmatrix} v_1(t) \\ \vdots \\ v_n(t) \end{bmatrix},$$

where $v_k : \mathbb{R} \rightarrow \mathbb{C}$ ($k = 1, \dots, n$). Since

$$\mathbf{v}^{(r)}(t) = \frac{d^r}{dt^r} \mathbf{v}(t) = A^r \mathbf{v}(t),$$

using the Cayley–Hamilton Theorem 2.6 we obtain

$$\mathbf{v}^{(n)}(t) + c_{n-1}(A)\mathbf{v}^{(n-1)}(t) + \cdots + c_1(A)\mathbf{v}^{(1)}(t) + c_0(A)\mathbf{v}(t) = \mathbf{0}, \quad (2.11a)$$

hence each of the coordinate functions v_k satisfies the ordinary differential equation

$$v_k^{(n)}(t) + c_{n-1}(A)v_k^{(n-1)}(t) + \cdots + c_1(A)v_k^{(1)}(t) + c_0(A)v_k(t) = 0. \quad (2.11b)$$

Of course, there are also boundary conditions coming from the initial condition $\mathbf{v}(0) = \mathbf{v}_0$.

Example 2.22

In Example 2.21, if $v_0 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, then the coordinate functions of the solution are

$$v_1(t) = ae^{2t} + bte^{2t}, \quad v_2(t) = be^{2t}, \quad v_3(t) = ce^t,$$

all of which satisfy the differential equations

$$v_k^{(3)}(t) - 6v_k^{(2)}(t) + 12v_k^{(1)}(t) - 8v_k(t) = 0 \quad (k = 1, 2, 3).$$

Of course the solution of a ordinary differential equation of form (2.11b) subject to suitable boundary conditions should be familiar, the novel point being the solution of an equivalent coupled system of first order differential equations for the v_k as a single vector differential equation solved using a one-parameter subgroup in $\text{GL}_n(\mathbb{C})$.

The geometry underlying this can sometimes be indicated in a diagram, at least in the case where everything is happening in \mathbb{R}^2 . The *phase portrait* shows at each point with coordinates (x, y) an arrow corresponding to the vector $A \begin{bmatrix} x \\ y \end{bmatrix}$, while the solution through a point can also be plotted as a *flow line*.

Example 2.23

The differential equation

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 4x(t) + 2y(t) \\ 4y(t) \end{bmatrix}$$

has solutions of the form

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} e^{4t} & 2te^{4t} \\ 0 & e^{4t} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} e^{4t}x_0 + 2te^{4t}y_0 \\ e^{4t}y_0 \end{bmatrix},$$

for arbitrary initial vectors $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$.

This follows from the fact that

$$\exp \left(\begin{bmatrix} 4t & 2t \\ 0 & 4t \end{bmatrix} \right) = e^{4t}I_2 + e^{4t}I_2 \left(2t \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) = \begin{bmatrix} e^{4t} & 2te^{4t} \\ 0 & e^{4t} \end{bmatrix}.$$

It is well known that the general solution of an ordinary differential equation of the form given in (2.11) is built up as a sum of polynomial functions multiplied by exponential functions.

EXERCISES

2.1. For $t \in \mathbb{R}$, determine each of the matrices

$$\exp\left(\begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix}\right), \quad \exp\left(\begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix}\right), \quad \exp\left(\begin{bmatrix} t & 0 \\ -2 & t \end{bmatrix}\right).$$

2.2. Let $\mathbf{k} = \mathbb{R}$ or \mathbb{C} and $A \in M_n(\mathbf{k})$.

a) Show that for $B \in GL_n(\mathbf{k})$,

$$\exp(BAB^{-1}) = B \exp(A)B^{-1}.$$

b) If D is diagonalisable with $D = C \operatorname{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})C^{-1}$ for some $C \in GL_n(\mathbf{k})$, show that

$$\exp(D) = C \operatorname{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})C^{-1}.$$

c) Use (b) to find the matrices

$$\exp\left(\begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix}\right), \quad \exp\left(\begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix}\right).$$

2.3. For $\mathbf{k} = \mathbb{R}, \mathbb{C}$ and $n \geq 1$, let $N \in M_n(\mathbf{k})$.

a) If N is strictly upper triangular, show that $\exp(N)$ is unipotent.
 b) Determine $\exp(N)$ when N is an upper triangular matrix with a fixed number t in every entry on its main diagonal.

2.4. a) If $S \in M_n(\mathbb{R})$ is skew symmetric, show that $\exp(S)$ is orthogonal, i.e., $\exp(S)^T = \exp(S)^{-1}$.

b) If $S \in M_n(\mathbb{C})$ is skew hermitian, show that $\exp(S)$ is unitary, i.e., $\exp(S)^* = \exp(S)^{-1}$.

2.5. 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 $\alpha: \mathbb{R} \rightarrow GL_2(\mathbb{R})$. Sketch the trajectory of this solution as a curve in the xy -plane. Investigate what happens for other initial

values $x(0), y(0)$.

b) Repeat this with the equations

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix};$$

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

2.6. If $A \in M_n(\mathbb{R})$ is skew symmetric, show that for a solution $\mathbf{x}: \mathbb{R} \rightarrow \mathbb{R}^n$ of the differential equation

$$\mathbf{x}'(t) = A\mathbf{x}(t),$$

$|\mathbf{x}(t)|$ is constant. In particular, if $|\mathbf{x}(0)| = 1$ then for all $t \in \mathbb{R}$,

$$\mathbf{x}'(t) \in T_{\mathbf{x}(t)} S^{n-1},$$

the tangent space to the unit sphere at $\mathbf{x}(t)$.

2.7. Let $\mathbf{k} = \mathbb{R}$ or \mathbb{C} .

a) Let $J(\lambda, r) \in M_r(\mathbf{k})$ be a Jordan block matrix. Show that there is a sequence of diagonalisable matrices $\{A_n\}_{n \geq 1}$ with $A_n \in M_r(\mathbf{k})$ and $A_n \rightarrow J(\lambda, r)$ as $n \rightarrow \infty$. If $\lambda \neq 0$, show that we can assume that $A_n \in GL_r(\mathbf{k})$ for all n .

b) Deduce that every matrix $A \in M_r(\mathbf{k})$ is a limit of diagonalisable matrices in $M_r(\mathbf{k})$ and if A is invertible, show that it is a limit of invertible diagonalisable matrices.

Tangent Spaces and Lie Algebras

In this chapter we show how to ‘linearise’ a matrix group G by considering its tangent space at the identity, which has the algebraic structure of a *Lie algebra*; the definition and basic properties of Lie algebra are introduced in Section 3.1. Amazingly, the Lie algebra of G captures enough of the properties of G to act as a more manageable substitute for many purposes, at least when G is *simply connected*. The geometric aspects of this will be studied in Chapter 7 when we investigate G as a *Lie group*.

3.1 Lie Algebras

In this section we collect together some basic concepts and results of the theory of Lie algebras which will be required at various times. We make no attempt at completeness, merely giving brief indications of how the algebraic theory develops.

Definition 3.1

A \mathbf{k} -*Lie algebra* or *Lie algebra over \mathbf{k}* consists of a vector space \mathfrak{a} over a field \mathbf{k} , together with a \mathbf{k} -bilinear map $[,] : \mathfrak{a} \times \mathfrak{a} \longrightarrow \mathfrak{a}$ called the *Lie bracket*, such

that for $x, y, z \in \mathfrak{a}$,

$$[x, y] = -[y, x], \quad (\text{Skew symmetry})$$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0. \quad (\text{Jacobi identity})$$

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

$$[r_1 x_1 + r_2 x_2, y] = r_1 [x_1, y] + r_2 [x_2, y],$$

$$[x, s_1 y_1 + s_2 y_2] = s_1 [x, y_1] + s_2 [x, y_2].$$

Example 3.2

Let $\mathbf{k} = \mathbb{R}$ and $\mathfrak{a} = \mathbb{R}^3$ and set

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

the *vector product* or *cross product* of \mathbf{x} and \mathbf{y} . For the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ we have the formulæ

$$\begin{cases} [\mathbf{e}_1, \mathbf{e}_2] = -[\mathbf{e}_2, \mathbf{e}_1] = \mathbf{e}_3, \\ [\mathbf{e}_2, \mathbf{e}_3] = -[\mathbf{e}_3, \mathbf{e}_2] = \mathbf{e}_1, \\ [\mathbf{e}_3, \mathbf{e}_1] = -[\mathbf{e}_1, \mathbf{e}_3] = \mathbf{e}_2. \end{cases} \quad (3.1)$$

Then \mathbb{R}^3 equipped with this bracket operation is an \mathbb{R} -Lie algebra. 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 $A, B \in M_n(\mathbf{k})$, their *commutator* or *Lie bracket* is the matrix

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

This defines a \mathbf{k} -bilinear function

$$[,]: M_n(\mathbf{k}) \times M_n(\mathbf{k}) \longrightarrow M_n(\mathbf{k}); \quad (A, B) \mapsto [A, B]$$

which is easily seen to satisfy the conditions of Definition 3.1.

Example 3.3

The \mathbf{k} -vector space $M_n(\mathbf{k})$ with the commutator bracket $[,]$ is a \mathbf{k} -Lie algebra.

Remark 3.4

Recall that A, B commute if $AB = BA$. Then $[A, B] = O_n$ if and only if A, B commute. So the Lie algebra structure on $M_n(\mathbf{k})$ gives a measure of how pairs of matrices fail to commute.

For later use, in the remainder of this section we develop some of the basic algebraic notions of Lie algebra theory. This material is analogous to the basic theory of groups or rings in that it introduces Lie subalgebras, Lie ideals, homomorphisms, etc.

Definition 3.5

If \mathfrak{a} is a \mathbf{k} -Lie algebra with bracket $[,]$, then a \mathbf{k} -subspace $\mathfrak{b} \subseteq \mathfrak{a}$ is a \mathbf{k} -Lie subalgebra of \mathfrak{a} if it is closed under taking commutators of pairs of elements in \mathfrak{b} , i.e., if whenever $x, y \in \mathfrak{b}$ then $[x, y] \in \mathfrak{b}$; we write $\mathfrak{b} \leqslant \mathfrak{a}$.

Of course, \mathfrak{a} is a \mathbf{k} -Lie subalgebra of itself, and $\{0\} \subseteq \mathfrak{a}$ is also a Lie subalgebra.

A Lie algebra in which all brackets are trivial is called *abelian*.

Suppose that $\mathfrak{n} \subseteq M_n(\mathbf{k})$ is a \mathbf{k} -vector subspace. Recalling Definition 3.5, we see that \mathfrak{n} is a \mathbf{k} -Lie subalgebra of $M_n(\mathbf{k})$ if it is closed under taking commutators of pairs of elements in \mathfrak{n} , i.e., whenever $A, B \in \mathfrak{n}$, then $[A, B] \in \mathfrak{n}$.

Definition 3.6

If \mathfrak{a} is a \mathbf{k} -Lie algebra with bracket $[,]$, then a \mathbf{k} -vector subspace $\mathfrak{n} \subseteq \mathfrak{a}$ is a Lie ideal of \mathfrak{a} if $[z, x] \in \mathfrak{n}$ for all $z \in \mathfrak{n}$ and $x \in \mathfrak{a}$; we then write $\mathfrak{n} \triangleleft \mathfrak{a}$. A Lie ideal $\mathfrak{n} \triangleleft \mathfrak{a}$ is proper if $\mathfrak{n} \neq \mathfrak{a}$ and it is non-trivial if $\mathfrak{n} \neq \{0\}$; the ideal $\{0\} \triangleleft \mathfrak{a}$ is the trivial ideal.

Definition 3.7

Let \mathfrak{a} be a \mathbf{k} -Lie algebra with bracket $[,]$.

- The centre of \mathfrak{a} is

$$z(\mathfrak{a}) = \{z \in \mathfrak{a} : \forall x \in \mathfrak{a}, [z, x] = 0\}.$$

- The commutator or derived subalgebra $\mathfrak{a}' \leqslant \mathfrak{a}$ is the vector subspace spanned by all the brackets $[x, y]$ ($x, y \in \mathfrak{a}$).

Proposition 3.8

$z(\mathfrak{a})$ and \mathfrak{a}' are Lie ideals of \mathfrak{a} , i.e., $z(\mathfrak{a}) \triangleleft \mathfrak{a}$ and $\mathfrak{a}' \triangleleft \mathfrak{a}$.

Proof

It is easy to see that $z(\mathfrak{a})$ is a vector subspace. For $z \in z(\mathfrak{a})$ and $x \in \mathfrak{a}$, $[x, z] = 0$ so $z(\mathfrak{a}) \triangleleft \mathfrak{a}$.

For \mathfrak{a}' , if $x, y, z \in \mathfrak{a}$ then $[x, [y, z]] \in \mathfrak{a}'$, so $\mathfrak{a}' \triangleleft \mathfrak{a}$. \square

Here is a useful result about the derived subalgebra which is left as an exercise.

Proposition 3.9

Let $\Phi: \mathfrak{a} \rightarrow \mathfrak{b}$ be a surjective homomorphism of \mathbb{k} -Lie algebras. Then \mathfrak{b} is abelian if and only if $\mathfrak{a}' \subseteq \ker \Phi$.

Definition 3.10

A \mathbb{k} -Lie algebra \mathfrak{a} is *simple* if every proper ideal $\mathfrak{b} \triangleleft \mathfrak{a}$ is trivial, i.e., $\mathfrak{b} = \{0\}$.

Definition 3.11

A \mathbb{k} -Lie algebra \mathfrak{a} has *trivial centre* or is *centreless* if $z(\mathfrak{g}) = \{0\}$.

Definition 3.12

Let $\mathfrak{g}, \mathfrak{h}$ be \mathbb{k} -Lie algebras. A \mathbb{k} -linear transformation $\Phi: \mathfrak{g} \rightarrow \mathfrak{h}$ is a *homomorphism of Lie algebras* if for all $x, y \in \mathfrak{g}$,

$$\Phi([x, y]) = [\Phi(x), \Phi(y)].$$

Such a homomorphism is an *isomorphism of Lie algebras* if it is also a \mathbb{k} -linear isomorphism.

For an ideal $\mathfrak{n} \triangleleft \mathfrak{a}$, the quotient vector space $\mathfrak{a}/\mathfrak{n}$ inherits a bracket operation defined by

$$[x + \mathfrak{n}, y + \mathfrak{n}] = [x, y] + \mathfrak{n}$$

for cosets $x + \mathfrak{n}, y + \mathfrak{n} \in \mathfrak{a}/\mathfrak{n}$. This is easily seen to make $\mathfrak{a}/\mathfrak{n}$ into a Lie algebra so that the quotient linear transformation $\mathfrak{a} \rightarrow \mathfrak{a}/\mathfrak{n}$ is a homomorphism of Lie algebras. $\mathfrak{a}/\mathfrak{n}$ is referred to as the *quotient Lie algebra* of \mathfrak{a} with respect to the ideal \mathfrak{n} . The following result is easily verified.

Proposition 3.13

Let $\Phi: \mathfrak{a} \rightarrow \mathfrak{b}$ be a homomorphism of \mathbf{k} -Lie algebras. Then $\ker \Phi \triangleleft \mathfrak{a}$ and there is an isomorphism of Lie algebras $\text{im } \Phi \cong \mathfrak{a}/\ker \Phi$.

The notion of *semi-direct product* of Lie algebras in the next definition is analogous to the group theoretic concept of Definition 1.36.

Definition 3.14

Let \mathfrak{a} be \mathbf{k} -Lie algebra, $\mathfrak{b} \leqslant \mathfrak{a}$ and $\mathfrak{n} \triangleleft \mathfrak{a}$. Then \mathfrak{a} is known as the *semi-direct product* of \mathfrak{b} and \mathfrak{n} if $\mathfrak{b} \cap \mathfrak{n} = \{0\}$ and $\mathfrak{a} = \mathfrak{b} + \mathfrak{n} = \{0\}$ as \mathbf{k} -vector spaces; we then write $\mathfrak{a} = \mathfrak{b} \ltimes \mathfrak{n}$.

As in group theory, there is a notion of *direct product* of two Lie algebras \mathfrak{a} and \mathfrak{b} . This is the vector space

$$\mathfrak{a} \times \mathfrak{b} = \{(x, y) : x \in \mathfrak{a}, y \in \mathfrak{b}\}$$

endowed with the bracket

$$[(x_1, y_1), (x_2, y_2)] = ([x_1, x_2], [y_1, y_2]).$$

This is a special case of the semi-direct product.

Finally, the following is a useful notation. For subspaces $U, V \leqslant \mathfrak{g}$ of a \mathbf{k} -Lie algebra \mathfrak{g} , set

$$[U, V] = \{\sum_j c_j [x_j, y_j] : c_j \in \mathbf{k}, x_j \in U, y_j \in V\} \subseteq \mathfrak{g}.$$

3.2 Curves, Tangent Spaces and Lie Algebras

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

Definition 3.15

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

$$T_U G = \{\gamma'(0) \in M_n(\mathbf{k}) : \gamma \text{ a differentiable curve in } G \text{ with } \gamma(0) = U\}.$$

Proposition 3.16

$T_U G$ is a real vector subspace of $M_n(\mathbf{k})$.

Proof

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

$$\gamma: \text{dom } \alpha \cap \text{dom } \beta \rightarrow G; \quad \gamma(t) = \alpha(t)U^{-1}\beta(t),$$

is also a differentiable curve in G with $\gamma(0) = U$. The Product Rule 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 $T_U G$ 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 $T_U G$ is closed under real scalar multiplication. \square

Definition 3.17

The *dimension* of the real matrix group G is $\dim G = \dim_{\mathbb{R}} T_I G$. If G is complex then its *complex dimension* is $\dim_{\mathbb{C}} G = \dim_{\mathbb{C}} T_I G$.

We will use the notation $\mathfrak{g} = T_I G$ for this real vector subspace of $M_n(\mathbb{k})$. In fact, \mathfrak{g} has the algebraic structure of a real Lie algebra as we now show.

Theorem 3.18

- i) If $G \leqslant \text{GL}_n(\mathbb{k})$ is a matrix subgroup, then \mathfrak{g} is an \mathbb{R} -Lie subalgebra of $M_n(\mathbb{k})$.
- ii) If $G \leqslant \text{GL}_m(\mathbb{C})$ is a matrix subgroup and \mathfrak{g} is a \mathbb{C} -subspace of $M_m(\mathbb{C})$, then \mathfrak{g} is a \mathbb{C} -Lie subalgebra.

Proof

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

Consider the function

$$F: \text{dom } \alpha \times \text{dom } \beta \rightarrow G; \quad F(s, t) = \alpha(s)\beta(t)\alpha(s)^{-1}.$$

This is clearly continuous and differentiable with respect to each of the variables s, t . For each $s \in \text{dom } \alpha$, the function $F(s, \cdot): \text{dom } \beta \rightarrow G$ is a differentiable curve in G with $F(s, 0) = I_n$. On differentiating we find

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

and so

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

Since \mathfrak{g} is a closed subspace of $M_n(\mathbb{k})$, whenever this limit exists we also have

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

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

$$\frac{d}{dt} (\alpha(t)^{-1}) = -\alpha(t)^{-1} \alpha'(t) \alpha(t)^{-1}. \quad (3.2)$$

We have

$$\begin{aligned} \lim_{s \rightarrow 0} \frac{1}{s} (\alpha(s)\beta'(0)\alpha(s)^{-1} - \beta'(0)) &= \frac{d}{ds} \Big|_{s=0} \alpha(s)\beta'(0)\alpha(s)^{-1} \\ &= \alpha'(0)\beta'(0)\alpha(0) - \alpha(0)\beta'(0)\alpha(0)^{-1}\alpha'(0)\alpha(0)^{-1} \\ &\quad [\text{by Equation (3.2)}] \\ &= \alpha'(0)\beta'(0)\alpha(0) - \alpha(0)\beta'(0)\alpha'(0) \\ &= \alpha'(0)\beta'(0) - \beta'(0)\alpha'(0) \\ &= [\alpha'(0), \beta'(0)]. \end{aligned}$$

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

Part (ii) is left as an exercise. □

So for each matrix group G there is a Lie algebra $\mathfrak{g} = T_I G$. A suitable type of homomorphism $G \rightarrow H$ between matrix groups gives rise to a linear transformation $\mathfrak{g} \rightarrow \mathfrak{h}$ which is a homomorphism of Lie algebras in the sense of Definition 3.12.

Definition 3.19

Let $G \leqslant \mathrm{GL}_n(\mathbb{k})$, $H \leqslant \mathrm{GL}_m(\mathbb{k})$ be matrix groups and $\varphi: G \rightarrow H$ be a continuous map. Then φ is said to be a *differentiable map* if it satisfies the following two conditions:

- for every differentiable curve $\gamma: (a, b) \rightarrow G$, the curve $\varphi \circ \gamma: (a, b) \rightarrow H$ is differentiable and has derivative

$$(\varphi \circ \gamma)'(t) = \frac{d}{dt} \varphi(\gamma(t));$$

- if two differentiable curves $\alpha, \beta: (a, b) \rightarrow G$ satisfy

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

then

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

A differentiable map which is also a group homomorphism is called a *differentiable homomorphism*. A continuous homomorphism of matrix groups that is also a differentiable map is called a *Lie homomorphism*.

Later we will see that the technical restriction in this definition is unnecessary. For now we note that if $\varphi: G \rightarrow H$ is the restriction of a differentiable map $\Phi: \mathrm{GL}_n(\mathbf{k}) \rightarrow \mathrm{GL}_m(\mathbf{k})$ then φ is also a differentiable map.

Proposition 3.20

Let G, H, K be matrix groups with $\varphi: G \rightarrow H$ and $\theta: H \rightarrow K$ differentiable homomorphisms. Then the following are true.

- i) For each $A \in G$ there is an \mathbb{R} -linear transformation $d\varphi: T_A G \rightarrow T_{\varphi(A)} H$ given by

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

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

- ii) We have

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

- iii) For the identity map $\mathrm{Id}_G: G \rightarrow G$ and $A \in G$,

$$d\mathrm{Id}_G = \mathrm{Id}_{T_A G}.$$

Proof

- (i) The definition of $d\varphi_A$ makes sense since by the definition of differentiability, given $X \in T_A G$, for any curve γ with

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

the derivative $(\varphi \circ \gamma)'(0)$ depends only on X and not on γ . Linearity is established using similar ideas to the proof of Proposition 3.16.

The identities of (ii) and (iii) are straightforward to verify. \square

If $\varphi: G \rightarrow H$ is a differentiable homomorphism then since $\varphi(I) = I$, $d\varphi_I: T_I G \rightarrow T_I H$ is a linear transformation called the *derivative* of φ which is usually denoted $d\varphi: \mathfrak{g} \rightarrow \mathfrak{h}$.

Theorem 3.21

Let G, H be matrix groups and $\varphi: G \rightarrow H$ be a differentiable homomorphism. Then the derivative $d\varphi: \mathfrak{g} \rightarrow \mathfrak{h}$ is a homomorphism of Lie algebras.

Proof

Following ideas and notation in the proof of Theorem 3.18, 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 that

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

□

As a special case, for each $A \in G$ the conjugation map

$$\chi_A: G \rightarrow G; \quad \chi_A(U) = AUA^{-1}$$

has as its derivative at I the *adjoint action* of A on \mathfrak{g} , namely the linear transformation

$$\text{Ad}_A = d\chi_A: \mathfrak{g} \rightarrow \mathfrak{g}; \quad \text{Ad}_A(X) = AXA^{-1}.$$

This satisfies

$$\text{Ad}_{AB} = \text{Ad}_A \circ \text{Ad}_B \quad (A, B \in G),$$

and on choosing a basis for \mathfrak{g} defines a continuous (and in fact differentiable) group homomorphism

$$\text{Ad}: G \rightarrow \text{GL}_{\dim G}(\mathbb{k}); \quad A \mapsto \text{Ad}_A.$$

Given a differentiable curve $\alpha: (-\varepsilon, \varepsilon) \rightarrow G$ with $\alpha(0)$, there is another curve

$$\tilde{\alpha}: (-\varepsilon, \varepsilon) \rightarrow \text{GL}_{\dim G}(\mathbb{k}); \quad \tilde{\alpha}(t) = \text{Ad}_{\alpha(t)},$$

whose derivative at 0 is given by

$$\tilde{\alpha}'(0)(X) = [\alpha'(0), X].$$

This defines the adjoint action of $U \in \mathfrak{g}$ on \mathfrak{g} ,

$$\text{ad}_U: \mathfrak{g} \rightarrow \mathfrak{g}; \quad \text{ad}_U(X) = [U, X].$$

It is easy to verify that

$$\text{ad}_{[U, V]}(X) = \text{ad}_U \circ \text{ad}_V(X) - \text{ad}_V \circ \text{ad}_U(X) \quad (U, V, X \in \mathfrak{g}).$$

So once a basis of \mathfrak{g} is chosen, the adjoint action gives rise to a homomorphism of Lie algebras

$$\text{ad}: \mathfrak{g} \rightarrow \text{M}_n(\mathbb{k}); \quad U \mapsto \text{ad}_U.$$

Ad and ad are commonly referred to as the *adjoint representations* of G and \mathfrak{g} .

3.3 The Lie Algebras of Some Matrix Groups

In this section we determine the Lie algebras of some important families of real and complex matrix groups. As usual $\mathbf{k} = \mathbb{R}$ or \mathbb{C} .

General and Special Linear Groups

We will start with the real matrix group $\mathrm{GL}_n(\mathbb{R}) \subseteq M_n(\mathbb{R})$. For $A \in M_n(\mathbb{R})$ and $\epsilon > 0$ there is a differentiable curve

$$\alpha: (-\epsilon, \epsilon) \longrightarrow M_n(\mathbb{R}); \quad \alpha(t) = I + tA.$$

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

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

then $\mathrm{im} \alpha \subseteq \mathrm{GL}_n(\mathbb{R})$, so we will view α as a function $\alpha: (-\epsilon, \epsilon) \longrightarrow \mathrm{GL}_n(\mathbb{R})$. Calculating the derivative we find that $\alpha'(t) = A$, hence $\alpha'(0) = A$. This shows that $A \in T_I \mathrm{GL}_n(\mathbb{R})$. Since $A \in M_n(\mathbb{R})$ was arbitrary, we have

$$\begin{cases} \mathfrak{gl}_n(\mathbb{R}) = T_I \mathrm{GL}_n(\mathbb{R}) = M_n(\mathbb{R}), \\ \dim \mathrm{GL}_n(\mathbb{R}) = n^2. \end{cases} \quad (3.3)$$

Similarly,

$$\begin{cases} \mathfrak{gl}_n(\mathbb{C}) = T_I \mathrm{GL}_n(\mathbb{C}) = M_n(\mathbb{C}), \\ \dim_{\mathbb{C}} \mathrm{GL}_n(\mathbb{C}) = n^2, \\ \dim \mathrm{GL}_n(\mathbb{C}) = 2n^2. \end{cases} \quad (3.4)$$

For $\mathrm{SL}_n(\mathbb{R}) \leq \mathrm{GL}_n(\mathbb{R})$, suppose that $\alpha: (a, b) \longrightarrow \mathrm{SL}_n(\mathbb{R})$ 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

$$\frac{d(\det \alpha(t))}{dt} \Big|_{t=0} = 0.$$

Lemma 3.22

We have

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

Proof

Recall that for $A \in M_n(\mathbb{k})$,

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

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

$$\partial(\gamma_1 \gamma_2) = (\partial\gamma_1)\gamma_2(0) + \gamma_1(0)\partial\gamma_2. \quad (3.5)$$

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

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

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 yields

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

giving the result. □

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

$$\mathfrak{sl}_n(\mathbb{R}) = T_I \text{SL}_n(\mathbb{R}) \subseteq \ker \text{tr} \subseteq M_n(\mathbb{R}).$$

If $A \in \ker \text{tr} \subseteq M_n(\mathbb{R})$, the function

$$\alpha: (-\varepsilon, \varepsilon) \rightarrow M_n(\mathbb{R}); \quad \alpha(t) = \exp(tA) = \sum_{k \geq 0} \frac{t^k}{k!} A^k,$$

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

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

We will make use of the following lemma.

Lemma 3.23

For $A \in M_n(\mathbb{C})$ we have

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

We will give two proofs here, while another appears in Section 7.5.

Proof

Approach using differential equations: Consider the curve

$$\gamma: \mathbb{R} \rightarrow \text{GL}_1(\mathbb{C}) = \mathbb{C}^\times; \quad \gamma(t) = \det \exp(tA).$$

By applying Lemma 3.22 to the curve $t \mapsto \det \exp(tA)$ we obtain

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

So α satisfies the same differential equation and initial condition as the curve $t \mapsto e^{t \text{tr } A}$. Hence

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

by the uniqueness part of Theorem 2.12. □

Proof

Approach using Jordan form: If $S \in \text{GL}_n(\mathbb{C})$,

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

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 the theory of *Jordan forms*, described in Section 2.2, there is a suitable choice of such an S for which

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

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

We have

$$\begin{aligned} \exp(B) &= \sum_{k \geq 0} \frac{1}{k!} (D + N)^k \\ &= \sum_{k \geq 0} \frac{1}{k!} D^k + \sum_{k \geq 0} \frac{1}{(k+1)!} \left(\sum_{j=0}^k \binom{k+1}{j} D^j N^{k+1-j} \right) \\ &= \exp(D) + \sum_{k \geq 0} \frac{1}{(k+1)!} N \left(\sum_{j=0}^k \binom{k+1}{j} D^j N^{k-j} \right). \end{aligned} \quad (3.6)$$

For each $k \geq 0$, the matrix

$$N \sum_{j=0}^k \binom{k+1}{j} D^j N^{k-j}$$

is strictly upper triangular, hence

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

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

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

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

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

□

Using this lemma and the function α , we obtain

$$\begin{cases} \mathfrak{sl}_n(\mathbb{R}) = T_I \mathrm{SL}_n(\mathbb{R}) = \ker \mathrm{tr} \subseteq M_n(\mathbb{R}), \\ \dim \mathrm{SL}_n(\mathbb{R}) = n^2 - 1. \end{cases} \quad (3.7)$$

Working over \mathbb{C} we also have

$$\begin{cases} \mathfrak{sl}_n(\mathbb{C}) = T_I \mathrm{SL}_n(\mathbb{C}) = \ker \mathrm{tr} \subseteq M_n(\mathbb{C}), \\ \dim_{\mathbb{C}} \mathrm{SL}_n(\mathbb{C}) = n^2 - 1, \\ \dim \mathrm{SL}_n(\mathbb{C}) = 2n^2 - 2. \end{cases} \quad (3.8)$$

Affine Groups

Recall the affine group $\mathrm{Aff}_n(\mathbf{k}) \leq \mathrm{GL}_n(\mathbf{k})$ which by Proposition 1.37 is the semi-direct product $\mathrm{GL}_n(\mathbf{k}) \ltimes \mathrm{Trans}_n(\mathbf{k})$. Using similar methods to those we used for $\mathrm{GL}_n(\mathbf{k})$ we find that

$$\mathfrak{aff}_n(\mathbf{k}) = \left\{ \begin{bmatrix} A & t \\ 0 & 0 \end{bmatrix} : A \in \mathfrak{gl}_n(\mathbf{k}), t \in \mathbf{k}^n \right\} = \left\{ \begin{bmatrix} A & t \\ 0 & 0 \end{bmatrix} : A \in M_n(\mathbf{k}), t \in \mathbf{k}^n \right\}.$$

Notice that $\mathfrak{aff}_n(\mathbf{k})$ is a \mathbf{k} -vector space and indeed it is a \mathbf{k} -Lie algebra. The Lie algebra structure here is interesting since although the translation subgroup $\mathrm{Trans}_n(\mathbf{k})$ is abelian and hence its Lie algebra $\mathrm{trans}_n(\mathbf{k})$ has trivial bracket, we find that

$$\begin{aligned} \begin{bmatrix} A_1 & t_1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_2 & t_2 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A_2 & t_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_1 & t_1 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} A_1 A_2 & A_1 t_2 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A_2 A_1 & A_2 t_1 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} [A_1, A_2] & A_1 t_2 - A_2 t_1 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

This shows in particular that $\mathrm{trans}_n(\mathbf{k})$ is a \mathbf{k} -Lie ideal in $\mathfrak{aff}_n(\mathbf{k})$ since it satisfies

$$[a, t] \in \mathrm{trans}_n(\mathbf{k}) \quad (a \in \mathfrak{aff}_n(\mathbf{k}), t \in \mathrm{trans}_n(\mathbf{k})).$$

Recalling Definition 3.14, we see that $\mathfrak{aff}_n(\mathbf{k})$ is the semi-direct product

$$\mathfrak{aff}_n(\mathbf{k}) = \mathfrak{gl}_n(\mathbf{k}) \ltimes \mathrm{trans}_n(\mathbf{k}).$$

To summarise, for $\mathbf{k} = \mathbb{R}$ or \mathbb{C} we have

$$\begin{cases} \mathfrak{aff}_n(\mathbf{k}) = T_I \mathrm{Aff}_n(\mathbf{k}) = \mathfrak{gl}_n(\mathbf{k}) \ltimes \mathrm{trans}_n(\mathbf{k}), \\ \dim_{\mathbf{k}} \mathrm{Aff}_n(\mathbf{k}) = n^2 + n, \\ \dim \mathrm{Aff}_n(\mathbf{k}) = (n^2 + n) \dim_{\mathbb{R}} \mathbf{k}. \end{cases} \quad (3.9)$$

Upper Triangular and Unipotent Groups

For $n \geq 1$, recall the upper triangular and unipotent subgroups $\text{UT}_n(\mathbf{k})$ and $\text{SUT}_n(\mathbf{k})$ of $\text{GL}_n(\mathbf{k})$. Let

$$\alpha: (-\varepsilon, \varepsilon) \longrightarrow \text{UT}_n(\mathbb{R})$$

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

$$\alpha: (-\varepsilon, \varepsilon) \longrightarrow \text{UT}_n(\mathbf{k}); \quad \alpha(t) = I + tA,$$

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

$$\begin{cases} \text{ut}_n(\mathbf{k}) = \text{set of upper triangular matrices in } M_n(\mathbf{k}), \\ \dim \text{ut}_n(\mathbf{k}) = \binom{n+1}{2} \dim_{\mathbb{R}} \mathbf{k}. \end{cases} \quad (3.10)$$

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

$$\begin{cases} \text{sut}_n(\mathbf{k}) = \text{set of strictly upper triangular matrices in } M_n(\mathbf{k}), \\ \dim \text{sut}_n(\mathbf{k}) = \binom{n}{2} \dim_{\mathbb{R}} \mathbf{k}. \end{cases} \quad (3.11)$$

Orthogonal and special orthogonal groups

Let $O(n)$ be the $n \times n$ orthogonal group, i.e.,

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

Given a curve $\alpha: (a, b) \longrightarrow O(n)$ satisfying $\alpha(0) = I$ we have

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

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

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

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

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

$$\alpha: (-\epsilon, \epsilon) \longrightarrow \text{GL}_n(\mathbb{R}); \quad \alpha(t) = \exp(tA).$$

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 $\alpha: (-\epsilon, \epsilon) \longrightarrow O(n)$. Since $\alpha'(0) = A$, this shows that

$$\text{Sk-Sym}_n(\mathbb{R}) \subseteq \mathfrak{o}(n) = T_I O(n)$$

and so

$$\mathfrak{o}(n) = T_I O(n) = \text{Sk-Sym}_n(\mathbb{R}).$$

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

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

hence $\text{tr } A = 0$. By Lemma 3.23, $\det \exp(tA) = 1$, so $\alpha: (-\epsilon, \epsilon) \longrightarrow \text{SO}(n)$ where $\text{SO}(n)$ is the $n \times n$ special orthogonal group. It is also easy to show that

$$\dim \text{Sk-Sym}_n(\mathbb{R}) = \binom{n}{2}.$$

So we have actually shown

$$\begin{cases} \mathfrak{so}(n) = T_I \text{SO}(n) = \mathfrak{o}(n) = T_I O(n) = \text{Sk-Sym}_n(\mathbb{R}), \\ \dim \text{SO}(n) = \binom{n}{2}. \end{cases} \quad (3.12)$$

We have also shown that if $A \in \text{Sk-Sym}_n(\mathbb{R})$, then

$$\exp(tA) \in \text{SO}(n) \quad (t \in \mathbb{R}). \quad (3.13)$$

Unitary and special unitary groups

Now consider the $n \times n$ unitary group

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

For a curve α in $U(n)$ 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

$$u(n) = T_I U(n) \subseteq \text{Sk-Herm}_n(\mathbb{C}),$$

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

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

$$\eta: (-\varepsilon, \varepsilon) \longrightarrow GL_n(\mathbb{C}); \quad \eta(t) = \exp(tH)$$

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 $\eta: (-\varepsilon, \varepsilon) \longrightarrow U(n)$. Since $\eta'(0) = H$, this shows that

$$\text{Sk-Herm}_n(\mathbb{C}) \subseteq u(n) = T_I U(n).$$

Hence since

$$\text{Sk-Herm}_n(\mathbb{C}) = n + 2 \binom{n}{2} = n^2,$$

we have

$$\begin{cases} u(n) = T_I U(n) = \text{Sk-Herm}_n(\mathbb{C}), \\ \dim U(n) = n^2. \end{cases} \quad (3.14)$$

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

$$su(n) = T_I SU(n) \subseteq \text{Sk-Herm}_n(\mathbb{C}).$$

But also if $\eta: (a, b) \longrightarrow SU(n)$ is a curve with $\eta(0) = I$, then as in the analysis for $SL_n(\mathbb{R})$, we have $\text{tr } \eta'(0) = 0$. Writing

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

we see that

$$\dim \text{Sk-Herm}_n^0(\mathbb{C}) = \dim \text{Sk-Herm}_n(\mathbb{C}) - 1 = n^2 - 1,$$

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

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

takes values in $\text{SU}(n)$ by Lemma 3.23 and satisfies $\eta'(0) = H$. Hence

$$\begin{cases} \mathfrak{su}(n) = T_I \text{SU}(n) \subseteq \text{Sk-Herm}_n^0(\mathbb{C}), \\ \dim \text{SU}(n) = n^2 - 1. \end{cases} \quad (3.15)$$

3.4 Some Observations on the Exponential Function of a Matrix Group

Later we will see that for a matrix group $G \leq \text{GL}_n(\mathbb{R})$, the following statements are always true and are often helpful in determining Lie algebras of matrix groups using the approaches seen in this section.

- The function

$$\exp_G: \mathfrak{g} \longrightarrow \text{GL}_n(\mathbb{R}); \quad \exp_G(X) = \exp(X),$$

has image contained in G , $\exp_G \mathfrak{g} \subseteq G$; so we will usually write $\exp_G: \mathfrak{g} \longrightarrow G$ for the exponential of G and sometimes even just \exp .

- If G is compact and connected then $\exp_G \mathfrak{g} = G$.
- There is an open disc $N_{\mathfrak{g}}(O; r) \subseteq \mathfrak{g}$ on which \exp is injective and gives a homeomorphism

$$\exp: N_{\mathfrak{g}}(O; r) \longrightarrow \exp N_{\mathfrak{g}}(O; r),$$

where $\exp N_{\mathfrak{g}}(O; r) \subseteq G$ is in fact an open subset.

Here is an example which shows that the exponential map need not be surjective for a connected matrix group.

Example 3.24

Consider the exponential $\exp_{\text{SL}_2(\mathbb{R})}: \mathfrak{sl}_2(\mathbb{R}) \longrightarrow \text{SL}_2(\mathbb{R})$. Then for $\delta \geq 0$, the matrix

$$\begin{bmatrix} -(2 + \delta) & 0 \\ 0 & \frac{-1}{(2 + \delta)} \end{bmatrix} \in \text{SL}_2(\mathbb{R})$$

is not in the image of $\exp_{\text{SL}_2(\mathbb{R})}$.

Proof

Let

$$A \in \mathfrak{sl}_2(\mathbb{R}) = \{A \in M_2(\mathbb{R}) : \text{tr } A = 0\}.$$

By the Cayley–Hamilton Theorem 2.6, we have

$$A^2 + (\det A)I_2 = 0.$$

If $\det A = 0$ then the only eigenvalue of A is 0 and for any $t \in \mathbb{R}$ we have

$$\exp(tA) = I_2 + tA,$$

from which we obtain

$$\text{tr } \exp(tA) = 2.$$

If $\det A \neq 0$, then A has two distinct non-zero eigenvalues, namely

$$\begin{cases} \pm\sqrt{\det A}i & \text{if } \det A > 0, \\ \pm\sqrt{-\det A} & \text{if } \det A < 0. \end{cases}$$

In either case there would be an invertible matrix $P \in GL_2(\mathbb{C})$ for which

$$P^{-1}AP = \begin{cases} \text{diag}(\sqrt{\det A}i, -\sqrt{\det A}i) & \text{if } \det A > 0, \\ \text{diag}(\sqrt{-\det A}, -\sqrt{-\det A}) & \text{if } \det A < 0. \end{cases}$$

So for $t \in \mathbb{R}$ we have

$$P^{-1}\exp(tA)P = \begin{cases} \text{diag}\left(e^{t\sqrt{\det A}i}, e^{-t\sqrt{\det A}i}\right) & \text{if } \det A > 0, \\ \text{diag}\left(e^{t\sqrt{-\det A}}, e^{-t\sqrt{-\det A}}\right) & \text{if } \det A < 0, \end{cases}$$

which yields

$$\exp(tA) = \begin{cases} \cos(t\sqrt{\det A})I_2 + \frac{\sin(t\sqrt{\det A})}{\sqrt{\det A}}A & \text{if } \det A > 0, \\ \cosh(t\sqrt{-\det A})I_2 + \frac{\sinh(t\sqrt{-\det A})}{\sqrt{-\det A}}A & \text{if } \det A < 0. \end{cases}$$

This in turn gives

$$\text{tr } \exp(tA) = \begin{cases} 2\cos(t\sqrt{\det A}) & \text{if } \det A > 0, \\ 2\cosh(t\sqrt{-\det A}) & \text{if } \det A < 0. \end{cases}$$

Hence, whenever $\text{tr } A = 0$ and $t \in \mathbb{R}$, $\text{tr } \exp(tA) \geq -2$. Since

$$\text{tr} \begin{bmatrix} -(2+\delta) & 0 \\ 0 & \frac{-1}{(2+\delta)} \end{bmatrix} < -2,$$

the matrix $\begin{bmatrix} -(2+\delta) & 0 \\ 0 & -1/(2+\delta) \end{bmatrix}$ cannot be of the form $\exp(A)$ for any real traceless matrix A . \square

3.5 SO(3) and SU(2)

In this section we will discuss the groups SO(3) and SU(2) and their Lie algebras in detail. It is more usual to do this using the identification of SU(2) with the unit *quaternions* which we define in Chapter 4, but here we will develop the ideas without that interpretation. The Lie algebras $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$ are both 3-dimensional real vector spaces, for example having the following bases:

$$\mathfrak{so}(3): \quad P = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$\mathfrak{su}(2): \quad H = \frac{1}{2} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad E = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad F = \frac{1}{2} \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

The non-trivial Lie brackets amongst these are

$$[P, Q] = R, \quad [Q, R] = P, \quad [R, P] = Q, \tag{3.16a}$$

$$[H, E] = F, \quad [E, F] = H, \quad [F, H] = E. \tag{3.16b}$$

This implies that the \mathbb{R} -linear isomorphism

$$\varphi: \mathfrak{su}(2) \longrightarrow \mathfrak{so}(3); \quad \varphi(xH + yE + zF) = xP + yQ + zR \quad (x, y, z \in \mathbb{R}), \tag{3.17}$$

satisfies

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

and so 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 describing this, notice also that for the Lie algebra of Example 3.2, the \mathbb{R} -linear transformation

$$\theta_0: \mathbb{R}^3 \longrightarrow \mathfrak{so}(3); \quad \theta_0(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3) = xP + yQ + zR,$$

is an isomorphism of \mathbb{R} -Lie algebras by the equations of (3.1).

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

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

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

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

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

The elements

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

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

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

form an orthonormal basis $\{\hat{H}, \hat{E}, \hat{F}\}$ of $\mathfrak{su}(2)$ with respect to the inner product $(\cdot | \cdot)$, i.e.,

$$(\hat{H} | \hat{H}) = (\hat{E} | \hat{E}) = (\hat{F} | \hat{F}) = 1, \quad (3.18a)$$

$$(\hat{H} | \hat{E}) = (\hat{H} | \hat{F}) = (\hat{E} | \hat{F}) = 0. \quad (3.18b)$$

We can define an \mathbb{R} -linear isomorphism

$$\theta: \mathbb{R}^3 \longrightarrow \mathfrak{su}(2); \quad \theta(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3) = x\hat{H} + y\hat{E} + z\hat{F}, \quad (3.19)$$

which is also an isometry, i.e.,

$$(\theta(\mathbf{x}) | \theta(\mathbf{y})) = \mathbf{x} \cdot \mathbf{y} \quad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^3).$$

Remark 3.25

It would perhaps be more natural to rescale the inner product $(\cdot | \cdot)$ so that H, E, F were all unit vectors. This would certainly make many of the formulæ that follow neater as well as making the Lie bracket in $SU(2)$ correspond exactly with the vector product in \mathbb{R}^3 . However, our choice of $(\cdot | \cdot)$ agrees with the conventional one for $SU(n)$ defined in Chapter 11.

Proposition 3.26

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

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

Proof

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

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

and in particular,

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

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

The invariance property is checked with a straightforward calculation. \square

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

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

hence Ad_A is actually an orthogonal linear transformation with respect to this inner product. Using the orthonormal basis $\hat{H}, \hat{E}, \hat{F}$, we can identify $\mathfrak{su}(2)$ with \mathbb{R}^3 and (\mid) with the usual inner product \cdot, \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

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

is a continuous homomorphism of groups. In fact, $\mathrm{SU}(2)$ is path connected, as is $\mathrm{SO}(3)$. Since $\overline{\mathrm{Ad}}(I) = I$, this implies that $\overline{\mathrm{Ad}}\mathrm{SU}(2) \subseteq \mathrm{SO}(3)$. Because of this, it is convenient to redefine $\overline{\mathrm{Ad}}$ by setting

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

Proposition 3.27

The continuous homomorphism of matrix groups

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

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 Section 9.1. 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 (3.16), the non-trivial brackets of these basis elements are as follows:

$$[\hat{H}, \hat{E}] = \sqrt{2}\hat{F}, \quad [\hat{E}, \hat{F}] = \sqrt{2}\hat{H}, \quad [\hat{F}, \hat{H}] = \sqrt{2}\hat{E}. \quad (3.20)$$

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

Lemma 3.28

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

The formula

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

implies the result. \square

In a similar fashion, we can express a product of elements of $\mathfrak{su}(2)$ in terms of the dot and cross products. Note however, that if $U_1, U_2 \in \mathfrak{su}(2)$ then in general $U_1 U_2 \notin \mathfrak{su}(2)$.

Lemma 3.29

If $U_1, U_2 \in \mathfrak{su}(2)$ with $U_1 = x_1\hat{H} + y_1\hat{E} + z_1\hat{F}$ and $U_2 = x_2\hat{H} + y_2\hat{E} + \hat{F}$, then

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

Proof

Calculation! \square

Notice that when $U_1, U_2 \in \mathfrak{su}(2)$ are *orthogonal* with respect to $(\cdot | \cdot)$ in the sense that $(U_1 | U_2) = 0$, then

$$U_1 U_2 = \frac{1}{2} [U_1, U_2] \in \mathfrak{su}(2). \quad (3.21)$$

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

$$A = \begin{bmatrix} u & v \\ -\bar{v} & \bar{u} \end{bmatrix} \quad (3.22)$$

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 $\sin \theta \geq 0$. A calculation gives

$$S^2 = -((\operatorname{Im} u)^2 + |v|^2)I = -\sin^2 \theta I, \quad (3.23a)$$

$$(S | S) = 2 \sin^2 \theta. \quad (3.23b)$$

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 $(S | U) = 0$, 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

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

By Equation (3.23b), whenever $(S | U) = 0$ 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 $\hat{S} = \frac{1}{\sqrt{2}\sin\theta} S$ 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 $R \in \text{SO}(3)$ has the form Ad_A for some $A \in \text{SU}(2)$. 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 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 v and $(S | S) = 2\sin^2(\varphi/2)$. If we choose $-\varphi$ in place of φ we obtain $-A$ in place of A .

This completes the proof of Proposition 3.27. \square

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

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

we can construct another curve

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

We can differentiate $\bar{\beta}$ at $t = 0$ to obtain an element of $\mathfrak{so}(3)$ given by the formula:

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

For example, when $B = H$ we have

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

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

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

Similarly,

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

giving the matrix

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

and

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

giving

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

The corresponding derivative map is then

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

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

To summarise, we have proved the following important result.

Theorem 3.30

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

In Chapter 5 we will meet the *spinor groups* and some generalisations of this double covering. A related double covering involving $\text{SL}_2(\mathbb{C})$ and the Lorentz group Lor will be discussed in Section 6.3.

3.6 The Complexification of a Real Lie Algebra

The ideas of this section are especially important in the representation theory of Lie groups and Lie algebras (see Serre [25] for example) and for completeness we provide a brief discussion.

Definition 3.31

Given a finite dimensional \mathbb{R} -Lie algebra \mathfrak{g} , a \mathbb{C} -Lie algebra \mathfrak{g}' which contains \mathfrak{g} as an \mathbb{R} -Lie subalgebra and for which $\dim_{\mathbb{C}} \mathfrak{g}' = \dim_{\mathbb{R}} \mathfrak{g}$ is called a *complexification* of \mathfrak{g} .

Theorem 3.32

- i) Every finite dimensional \mathbb{R} -Lie algebra \mathfrak{g} has a complexification.
- ii) If \mathfrak{g}' and \mathfrak{g}'' are two complexifications of \mathfrak{g} then there is an isomorphism of \mathbb{C} -Lie algebras $\mathfrak{g}' \longrightarrow \mathfrak{g}''$ which extends the identity function on $\mathfrak{g} \subseteq \mathfrak{g}'$.

Because of the uniqueness guaranteed by part (ii) this result, we can write g_C for such a complexification of g since this is well defined up to an isomorphism of \mathbb{C} -Lie algebras. Our next result gives a useful criterion for finding complexifications.

Proposition 3.33

Let \mathfrak{h} be a finite dimensional \mathbb{C} -Lie algebra and $\mathfrak{g} \subseteq \mathfrak{h}$ be an \mathbb{R} -Lie subalgebra. If $\mathfrak{g}' \subseteq \mathfrak{h}$ is the smallest \mathbb{C} -vector subspace containing \mathfrak{g} , then \mathfrak{g}' is a complexification of \mathfrak{g} . In particular, if $\{u_1, \dots, u_d\}$ is an \mathbb{R} -basis for \mathfrak{g} which is \mathbb{C} -linearly independent in \mathfrak{h} , then $\{u_1, \dots, u_d\}$ is a basis for \mathfrak{g}' .

Recall that $\dim U(n) = \dim u(n) = n^2$. Then $u(n)$ is an \mathbb{R} -Lie subalgebra of $gl_n(\mathbb{C}) = M_n(\mathbb{C})$ where $\dim_{\mathbb{R}} gl_n(\mathbb{C}) = 2n^2$. Of course, $gl_n(\mathbb{C})$ is also a \mathbb{C} -Lie algebra of dimension $\dim_{\mathbb{C}} gl_n(\mathbb{C}) = n^2$.

Proposition 3.34

The following are complexifications:

$$gl_n(\mathbb{C}) = gl_n(\mathbb{R})_C = u(n)_C, \quad sl_n(\mathbb{C}) = sl_n(\mathbb{R})_C = su(n)_C.$$

Proof

It is easy to check that the following statements are true by explicitly exhibiting \mathbb{R} -bases of $u(n)$ and $su(n)$, then verifying that they are \mathbb{C} -bases for $gl_n(\mathbb{C})$ and $sl_n(\mathbb{C})$.

Let E^{rs} be the $n \times n$ matrix for which

$$E^{rs}_{ij} = \delta_{ir}\delta_{js} = \begin{cases} 1 & \text{if } i = r \text{ and } j = s, \\ 0 & \text{otherwise.} \end{cases} \quad (3.24)$$

$u(n)$ has an \mathbb{R} -basis consisting of the elements

$$\begin{cases} D_k = iE^{kk} & (1 \leq k \leq n), \\ P_{k\ell} = E^{k\ell} - E^{\ell k} & (1 \leq \ell < k \leq n), \\ Q_{k\ell} = iE^{k\ell} + iE^{\ell k} & (1 \leq \ell < k \leq n), \end{cases} \quad (3.25)$$

while $su(n)$ has an \mathbb{R} -basis consisting of the elements

$$\begin{cases} H_k = iE^{kk} - iE^{(k-1)(k-1)} & (1 \leq k \leq n-1), \\ P_{k\ell} = E^{k\ell} - E^{\ell k} & (1 \leq \ell < k \leq n), \\ Q_{k\ell} = iE^{k\ell} + iE^{\ell k} & (1 \leq \ell < k \leq n). \end{cases} \quad (3.26)$$

□

The Lie algebras $\mathfrak{gl}_n(\mathbb{C})$ and $\mathfrak{sl}_n(\mathbb{C})$ have the following \mathbb{C} -bases.

$$\mathfrak{gl}_n(\mathbb{C}): \begin{cases} E^{k k} & (1 \leq k \leq n), \\ E^{k \ell} & (1 \leq \ell, k \leq n, k \neq \ell). \end{cases} \quad (3.27)$$

$$\mathfrak{sl}_n(\mathbb{C}): \begin{cases} E^{k k} - E^{(k-1)(k-1)} & (2 \leq k \leq n), \\ E^{k \ell} & (1 \leq \ell, k \leq n, k \neq \ell). \end{cases} \quad (3.28)$$

Here is a general result on the existence of complexifications.

Proposition 3.35

Every \mathbb{R} -Lie subalgebra $\mathfrak{g} \leq \mathfrak{u}(n)$ has a complexification $\mathfrak{g}_{\mathbb{C}}$ which is a \mathbb{C} -Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$. If $G \leq \mathrm{SU}(n)$ is a Lie subgroup, then its Lie algebra $\mathfrak{g} \leq \mathfrak{su}(n)$ has a complexification $\mathfrak{g}_{\mathbb{C}} \leq \mathfrak{su}(n)_{\mathbb{C}} = \mathfrak{sl}_n(\mathbb{C})$.

We know several 3-dimensional real Lie algebras, namely $\mathfrak{so}(3)$, $\mathfrak{su}(2)$, $\mathfrak{sl}_2(\mathbb{R})$ and (up to isomorphism) a complex one

$$\mathfrak{sl}_2(\mathbb{C}) \cong \mathfrak{sl}_2(\mathbb{R})_{\mathbb{C}} \cong \mathfrak{su}(2)_{\mathbb{C}} \cong \mathfrak{so}(3)_{\mathbb{C}}.$$

We know that $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ as real Lie algebras, but have not yet considered whether $\mathfrak{so}(3) \cong \mathfrak{sl}_2(\mathbb{R})$. To do this, we need to investigate these Lie algebras further.

Since the Lie algebra $\mathfrak{so}(3)$ is essentially the vector space \mathbb{R}^3 with the vector cross product as its bracket, it is easy to see that given a pair of non-zero elements $X, Y \in \mathfrak{so}(3)$ which are linearly independent, their bracket $[X, Y]$ is non-zero and the vectors $X, Y, [X, Y]$ are linearly independent, and hence form a basis for $\mathfrak{so}(3)$. We will show that this is false for the Lie algebra $\mathfrak{sl}_2(\mathbb{R})$, so these Lie algebras cannot be isomorphic.

The pair of matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

in $\mathfrak{sl}_2(\mathbb{R})$ has bracket

$$[A, B] = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix},$$

and the matrices $A, B, [A, B]$ are linearly dependent. As a consequence of this, we can deduce some information about the corresponding matrix groups.

Proposition 3.36

The real Lie algebras $\mathfrak{so}(3)$ and $\mathfrak{sl}_2(\mathbb{R})$ are not isomorphic. Hence there can be no Lie homomorphisms of any of the forms

$$\mathrm{SL}_2(\mathbb{R}) \rightarrow \mathrm{SO}(3), \quad \mathrm{SL}_2(\mathbb{R}) \rightarrow \mathrm{SU}(2), \quad \mathrm{SO}(3) \rightarrow \mathrm{SL}_2(\mathbb{R}), \quad \mathrm{SU}(2) \rightarrow \mathrm{SL}_2(\mathbb{R})$$

with injective derivatives.

EXERCISES

3.1. Let $\mathbf{k} = \mathbb{R}$ or \mathbb{C} .

- a) Consider the 2-dimensional \mathbf{k} -Lie algebras

$$\mathfrak{a} = \mathbf{k}^2, \quad \mathfrak{b} = \left\{ \begin{bmatrix} u & v \\ 0 & 0 \end{bmatrix} : u, v \in \mathbf{k} \right\},$$

with the obvious brackets which make \mathfrak{a} abelian and $\mathfrak{b} \leq M_2(\mathbf{k})$. Show that any 2-dimensional \mathbf{k} -Lie algebra \mathfrak{g} is isomorphic to \mathfrak{a} if it is abelian and \mathfrak{b} otherwise.

- b) Find a matrix group $G \leq \mathrm{GL}_2(\mathbf{k})$ whose Lie algebra is \mathfrak{b} .

3.2. Let G be a matrix group and $U \in G$.

- a) Show that each of the functions

$$\begin{aligned} L_U: G &\longrightarrow G; \quad L_U(A) = UA, \\ R_U: G &\longrightarrow G; \quad R_U(A) = AU, \\ C_U: G &\longrightarrow G; \quad C_U(A) = UAU^{-1}, \end{aligned}$$

is a differentiable map and determine its derivative at I .

- b) Using (a), show that there are \mathbb{R} -linear isomorphisms

$$\lambda_U: T_I G \longrightarrow T_U G, \quad \rho_U: T_I G \longrightarrow T_U G, \quad \chi_U: T_I G \longrightarrow T_I G,$$

such that for all $U, V \in G$,

$$\lambda_{UV} = \lambda_U \circ \lambda_V, \quad \rho_{UV} = \rho_V \circ \rho_U, \quad \chi_{UV} = \chi_U \circ \chi_V.$$

3.3. For each of the following matrix groups G , find the Lie algebra \mathfrak{g} .

$$G_1 = \{A \in \mathrm{GL}_2(\mathbb{R}) : A^T Q_1 A = Q_1\}, \quad Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix};$$

$$G_2 = \{A \in \mathrm{GL}_2(\mathbb{R}) : A^T Q_2 A = Q_2\}, \quad Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix};$$

$$G_3 = \{A \in \mathrm{GL}_3(\mathbb{R}) : A^T Q_3 A = Q_3\}, \quad Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix};$$

$$G_4 = \mathrm{Aff}_n(\mathbb{k}) \quad (n = 1, 2, \dots);$$

$$G_5 = \mathrm{Symp}_{2m}(\mathbb{R}) \quad (m = 1, 2, \dots).$$

3.4. Prove Proposition 3.9.

3.5. Let G be a matrix group with Lie algebra \mathfrak{g} and let $X, Y \in \mathfrak{g}$. Show that $[X, Y] = 0$ if and only if $\exp(sX)\exp(tY) = \exp(tY)\exp(sX)$ for all $s, t \in \mathbb{R}$.

3.6. 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

$$\Phi: U \longrightarrow \mathrm{M}_n(\mathbb{R}); \quad \Phi(A) = (I - A)(I + A)^{-1}.$$

*[\Phi is known as the *real Cayley transform*.]*

- a) Show that $\mathrm{im} \Phi = \mathrm{Sk-Sym}_n(\mathbb{R})$, the set of all $n \times n$ real skew symmetric matrices. Hence we might as well write $\Phi: U \longrightarrow \mathrm{Sk-Sym}_n(\mathbb{R})$.
- b) Find the inverse map $\Phi^{-1}: \mathrm{Sk-Sym}_n(\mathbb{R}) \longrightarrow U$.
- c) Use (b) to determine the dimension of $\mathrm{SO}(n)$.

3.7. 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

$$\Theta: V \longrightarrow \mathrm{M}_n(\mathbb{C}); \quad \Theta(A) = (I - A)(I + A)^{-1}.$$

*[\Theta is known as the *complex Cayley transform*.]*

- a) Show that $\mathrm{im} \Theta = \mathrm{Sk-Herm}_n(\mathbb{C})$, the set of all $n \times n$ skew hermitian matrices. Hence we might as well write $\Theta: V \longrightarrow \mathrm{Sk-Herm}_n(\mathbb{C})$.

- b) Find the inverse map $\Theta^{-1} : \text{Sk-Herm}_n(\mathbb{C}) \longrightarrow V$.
- c) Use (b) to determine the dimension of $U(n)$.
- d) In the case $n = 2$, show that $\Theta(V \cap \text{SU}(2)) \subseteq \text{Sk-Herm}_2^0(\mathbb{C})$ and $\Theta^{-1} \text{Sk-Herm}_2^0(\mathbb{C}) \subseteq \text{SU}(2)$. Is this true for $n > 2$?

3.8. For $n \geq 1$, prove the following:

- a) $O(n)$ is the semi-direct product $\{1, -1\} \ltimes \text{SO}(n)$;
- b) $U(n)$ is the semi-direct product $T \ltimes \text{SU}(n)$, where

$$T = \{z \in \mathbb{C} : |z| = 1\}$$

is the unit circle;

- c) $GL_n(\mathbb{R})$ is the semi-direct product $\mathbb{R}^\times \ltimes \text{SL}_n(\mathbb{R})$;
- d) $GL_n(\mathbb{C})$ is the semi-direct product $\mathbb{C}^\times \ltimes \text{SL}_n(\mathbb{C})$.

3.9. Verify the formula of Lemma 3.29.

Algebras, Quaternions and Quaternionic Symplectic Groups

In this chapter we begin by studying algebras over a field, with their groups of units providing many interesting groups. In particular, we study division algebras and their linear algebra. Then we introduce the *quaternions* which form the only non-commutative example of a real division algebra. There is an associated family of compact connected matrix groups defined using the quaternions, the *quaternionic symplectic groups* which provide another infinite family of compact simply connected matrix groups.

4.1 Algebras

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

Definition 4.1

A *finite dimensional (associative and unital) algebra* A is a finite dimensional \mathbf{k} -vector space which is also an associative and unital ring such that for all $r, s \in \mathbf{k}$ and $a, b \in A$,

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

Here ra and sb are scalar products in the vector space structure, while $(rs)(ab)$ is the scalar product of $rs \in \mathbf{k}$ with the ring product $ab \in A$.

In such a \mathbf{k} -algebra A , if $1 \in \mathbf{k}$ is the unit of A , then for $t \in \mathbf{k}$, the element $t1 \in A$ satisfies

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

If $\dim_{\mathbf{k}} A > 0$, then $1 \neq 0$, and the function

$$\eta: \mathbf{k} \longrightarrow A; \quad \eta(t) = t1$$

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

If A is a commutative ring then A is a *commutative \mathbf{k} -algebra*.

If every non-zero element $u \in A$ is a unit, i.e., is invertible, then A is a **\mathbf{k} -division algebra** or *division algebra over \mathbf{k}* . A commutative division algebra is a *field* while a non-commutative division algebra is called a *skew field*. In French *corps* (\sim field) is often used to refer to a possibly non-commutative division algebra.

Example 4.2

For $n \geq 1$, $M_n(\mathbf{k})$ is a \mathbf{k} -algebra. Here we have $\eta(t) = tI_n$. For $n > 1$, $M_n(\mathbf{k})$ is non-commutative.

Example 4.3

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

Example 4.4

Let G be a finite group. Then for $\mathbf{k} = \mathbb{R}$ or \mathbf{C} , the *group algebra* $\mathbf{k}[G]$ has a basis consisting of the elements g of G , while the addition and multiplication are

$$\begin{aligned} \left(\sum_{g \in G} x_g g \right) + \left(\sum_{g \in G} y_g g \right) &= \sum_{g \in G} (x_g + y_g) g, \\ \left(\sum_{g \in G} x_g g \right) \left(\sum_{h \in G} y_h h \right) &= \sum_{g \in G} \left(\sum_{h \in G} x_{gh^{-1}} y_h \right) g. \end{aligned}$$

The unit is $1 = 1_G$, the identity element of G .

In any \mathbf{k} -algebra A , the set of units of A forms a group A^\times under multiplication, and this contains \mathbf{k}^\times as a central subgroup, $\mathbf{k}^\times \leq A^\times$. If A is non-commutative, we might hope to find interesting subgroups of A^\times . Later we will see that the Clifford algebras provide a good illustration of this idea.

Example 4.5

For $A = M_n(\mathbf{k})$, $M_n(\mathbf{k})^\times = GL_n(\mathbf{k})$, while \mathbf{k}^\times is identified with the subgroup of invertible scalar matrices tI_n ($t \in \mathbf{k}^\times$).

Example 4.6

Let $\mathbf{k}[G]$ be the group algebra of a finite group G . Then the basis elements $g \in G$ form a finite subgroup $G \leq \mathbf{k}[G]^\times$.

Definition 4.7

Let A and B be two \mathbf{k} -algebras. A \mathbf{k} -linear transformation $\varphi: A \rightarrow B$ that is also a ring homomorphism is called a *\mathbf{k} -algebra homomorphism* or *homomorphism of \mathbf{k} -algebras*. A homomorphism of \mathbf{k} -algebras which is also an isomorphism of rings or equivalently of \mathbf{k} -vector spaces is called an *isomorphism of \mathbf{k} -algebras*. An isomorphism $\alpha: A \rightarrow A$ is called an *automorphism of \mathbf{k} -algebras*.

Notice that the unit $\eta: \mathbf{k} \rightarrow A$ is always a homomorphism of \mathbf{k} -algebras. There are obvious notions of kernel and image for such homomorphisms, and of subalgebra.

Definition 4.8

Given two \mathbf{k} -algebras A, B , their *direct product* has underlying set $A \times B$ with sum and product

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

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

It is easy to see that there is an isomorphism of \mathbf{k} -algebras $A \times B \cong B \times A$.

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

$$\dim_{\mathbf{k}} M_m(A) = m^2 \dim_{\mathbf{k}} A.$$

It is often the case that a \mathbf{k} -algebra A contains a subalgebra $\mathbf{k}_1 \subseteq A$ which is also a field. In that case A can be viewed as a over \mathbf{k}_1 in two different ways, corresponding to left and right multiplication by elements of \mathbf{k}_1 . Then for $t \in \mathbf{k}_1$, $a \in A$,

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

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

These give different \mathbf{k}_1 -vector space structures unless all elements of \mathbf{k}_1 commute with all elements of A , in which case \mathbf{k}_1 is said to be a *central subfield* of A . We sometimes write $\mathbf{k}_1 A$ and $A_{\mathbf{k}_1}$ to indicate which structure is being considered. \mathbf{k}_1 is itself a finite dimensional commutative \mathbf{k} -algebra of some dimension $\dim_{\mathbf{k}} \mathbf{k}_1$.

Proposition 4.9

Each of the \mathbf{k}_1 -vector spaces $\mathbf{k}_1 A$ and $A_{\mathbf{k}_1}$ is finite dimensional and in fact

$$\dim_{\mathbf{k}} A = \dim_{\mathbf{k}_1} (\mathbf{k}_1 A) \dim_{\mathbf{k}} \mathbf{k}_1 = \dim_{\mathbf{k}_1} (A_{\mathbf{k}_1}) \dim_{\mathbf{k}} \mathbf{k}_1.$$

Example 4.10

Let $\mathbf{k} = \mathbb{R}$ and $A = M_2(\mathbb{R})$ where $\dim_{\mathbb{R}} A = 4$. If

$$\mathbf{k}_1 = \left\{ \begin{bmatrix} x & y \\ -y & x \end{bmatrix} : x, y \in \mathbb{R} \right\} \subseteq M_2(\mathbb{R}),$$

then $\mathbf{k}_1 \cong \mathbb{C}$ so is a subfield of $M_2(\mathbb{R})$, but it is not a central subfield. We also have $\dim_{\mathbf{k}_1} A = 2$.

Example 4.11

Let $\mathbf{k} = \mathbb{R}$ and $A = M_2(\mathbb{C})$, so $\dim_{\mathbb{R}} A = 8$. Let

$$\mathbf{k}_1 = \left\{ \begin{bmatrix} x & y \\ -y & x \end{bmatrix} : x, y \in \mathbb{R} \right\} \subseteq M_2(\mathbb{C}).$$

Then $\mathbf{k}_1 \cong \mathbb{C}$ so is subfield of $M_2(\mathbb{C})$, but it is not a central subfield. Here $\dim_{\mathbf{k}_1} A = 4$.

Given a \mathbf{k} -algebra A and a subfield $\mathbf{k}_1 \subseteq A$ containing \mathbf{k} (possibly equal to \mathbf{k}), an element $a \in A$ acts on A by left multiplication:

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

This is always a \mathbf{k} -linear transformation of A , and if we view A as the \mathbf{k}_1 -vector space $A_{\mathbf{k}_1}$, it is always a \mathbf{k}_1 -linear transformation. Given a \mathbf{k}_1 -basis $\{v_1, \dots, v_m\}$ for $A_{\mathbf{k}_1}$, there is an $m \times m$ matrix $\lambda(a)$ with entries in \mathbf{k}_1 defined by

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

It is easy to check that

$$\lambda: A \longrightarrow M_m(\mathbf{k}_1); \quad a \longmapsto \lambda(a)$$

is a homomorphism of \mathbf{k} -algebras, called the *left regular representation* of A over \mathbf{k}_1 with respect to the basis $\{v_1, \dots, v_m\}$.

Lemma 4.12

$\lambda: A \longrightarrow M_m(\mathbf{k}_1)$ 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 $a = 0$. \square

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

$$\lambda^\times: A^\times \longrightarrow GL_m(\mathbf{k}_1); \quad \lambda^\times(a)(u) = au,$$

where $\mathbf{k}_1 \subseteq A$ is a subfield containing \mathbf{k} and we have chosen a \mathbf{k}_1 -basis of $A_{\mathbf{k}_1}$. Because

$$A^\times \cong \text{im } \lambda^\times \leq GL_m(\mathbf{k}_1),$$

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

Given a \mathbf{k} -basis of A , we obtain a group homomorphism

$$\rho^\times: A^\times \longrightarrow GL_n(\mathbf{k}); \quad \rho^\times(a)(u) = ua^{-1}.$$

We can combine λ^\times and ρ^\times to obtain two further group homomorphisms

$$\begin{aligned} \lambda^\times \times \rho^\times: A^\times \times A^\times &\longrightarrow GL_n(\mathbf{k}); \quad \lambda^\times \times \rho^\times(a, b)(u) = aub^{-1}, \\ \Delta: A^\times &\longrightarrow GL_n(\mathbf{k}); \quad \Delta(a)(u) = aua^{-1}. \end{aligned}$$

Notice that these homomorphisms have non-trivial kernels,

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

Definition 4.13

A \mathbf{k} -algebra A is *simple* if it has only one proper two-sided ideal, namely (0) , hence every non-trivial \mathbf{k} -algebra homomorphism $\theta: A \rightarrow B$ is an injection.

Example 4.14

Let \mathbf{k} be a field and $p(x) \in \mathbf{k}[x]$ be an irreducible polynomial of degree $d > 0$. Then the quotient ring $\mathbf{k}[x]/(p(x))$ is a field and so is a simple \mathbf{k} -algebra of dimension $\dim_{\mathbf{k}} \mathbf{k}[x]/(p(x)) = d$.

Proposition 4.15

Let \mathbf{k} be a field.

- i) For a division algebra \mathbf{D} over \mathbf{k} , \mathbf{D} is simple.
- ii) For a simple \mathbf{k} -algebra A , $M_n(A)$ is simple. In particular, $M_n(\mathbf{k})$ is a simple \mathbf{k} -algebra.

In fact, there is a classification of all such simple \mathbf{k} -algebras.

Theorem 4.16 (Wedderburn's Theorem)

Let A be a finite dimensional simple algebra over \mathbf{k} .

- i) A is isomorphic to some $M_n(\mathbf{D})$ where \mathbf{D} is a finite dimensional division algebra over \mathbf{k} , hence its dimension is $\dim_{\mathbf{k}} A = n^2 \dim_{\mathbf{k}} \mathbf{D}$.
- ii) If \mathbf{k} is algebraically closed then A is isomorphic to some $M_n(\mathbf{k})$, hence it has dimension $\dim_{\mathbf{k}} A = n^2$.
- iii) Every finite dimensional division algebra \mathbf{D} over \mathbf{k} has dimension of the form $\dim_{\mathbf{k}} \mathbf{D} = d^2$ for some natural number d .

When $\mathbf{k} = \mathbb{R}$, it turns out that there is only one non-commutative division algebra, namely the division algebra of quaternions, \mathbb{H} , described in Section 4.4.

Definition 4.17

Let A be a \mathbf{k} -algebra.

- An element $e \in A$ is an *idempotent* if $e^2 = e$.
- An idempotent $e \in A$ is *central* if $ae = ea$ for every $a \in A$.
- A collection of idempotents e_1, \dots, e_ℓ is *orthogonal* if $e_i e_j = 0$ when $i \neq j$.

- An idempotent $e \in A$ is *indecomposable* if whenever $e = e_1 + e_2$ for idempotents e_1, e_2 , then $e_1 = 0$ or $e_2 = 0$.

Proposition 4.18

A finite dimensional \mathbf{k} -algebra A is a product of ℓ algebras if and only if there are orthogonal central idempotents e_1, \dots, e_ℓ such that

$$e_1 + \cdots + e_\ell = 1.$$

Proof

If $A = A_1 \times \cdots \times A_\ell$ then

$$e_k = (0, \dots, 0, 1, 0, \dots, 0)$$

(with a single 1 in the k th place) is a central idempotent and $e_1 + \cdots + e_\ell = 1$; furthermore, $e_i e_j = 0$ whenever $i \neq j$.

Conversely, if such central orthogonal idempotents exist, we can set

$$A_k = Ae_k = \{ae_k : a \in A\}$$

and this is easily seen to be an algebra. Then the correspondence

$$a \longleftrightarrow (ae_1, \dots, ae_\ell)$$

shows that A is the product of the algebras A_k . Notice that

$$1 \longleftrightarrow (e_1^2, \dots, e_\ell^2) = (e_1, \dots, e_\ell)$$

under this identification. □

Lemma 4.19

Let A be a \mathbf{k} -algebra and $e \in A$ be a central idempotent. Then $(1 - e)$ is an idempotent and the idempotents $e, (1 - e)$ are orthogonal. Moreover, Ae is an ideal in A which is trivial if and only if $e = 0$.

Proof

We have

$$(1 - e)^2 = 1 + e^2 - 2e = 1 + e - 2e = 1 - e,$$

and

$$e(1-e) = (1-e)e = e - e^2 = e - e = 0.$$

If $a, b \in A$, then we have

$$b(ae) = (ba)e \in Ae, \quad (ae)b = aeb = abe = (ab)e \in Ae,$$

showing that Ae is an ideal. \square

Definition 4.20

A \mathbf{k} -algebra A is *semi-simple* if it is a product of simple algebras A_k ,

$$A = A_1 \times \cdots \times A_\ell.$$

Proposition 4.21

A finite dimensional \mathbf{k} -algebra A is semi-simple if and only if there are indecomposable orthogonal central idempotents e_1, \dots, e_ℓ such that

$$e_1 + \cdots + e_\ell = 1.$$

Proof

If $A = A_1 \times \cdots \times A_\ell$ with each A_k simple, then each of the corresponding central orthogonal idempotents e_i must be indecomposable since otherwise A_i would have a proper ideal. \square

Example 4.22

Consider the commutative 2-dimensional \mathbf{C} -algebra

$$A = \{(u, v) : u, v \in \mathbf{C}\}$$

equipped with the obvious addition but multiplication

$$(u_1, v_1)(u_2, v_2) = (u_1 u_2 - v_1 v_2, u_1 v_2 + v_1 u_2).$$

The unit is $1 = (1, 0)$. Notice that if we restrict attention to

$$A_0 = \{(u, v) : u, v \in \mathbb{R}\},$$

this is isomorphic to the \mathbb{R} -algebra of complex numbers \mathbf{C} , using the correspondence

$$(u, v) \longleftrightarrow u + vi.$$

However, A contains the orthogonal idempotents

$$e = (1/2, -i/2), \quad 1 - e = (1/2, i/2).$$

Hence,

$$\begin{aligned} A &= Ae \times A(1 - e) \\ &= \left\{ \left(\frac{u + vi}{2}, \frac{-ui + v}{2} \right) : u, v \in \mathbb{C} \right\} \times \left\{ \left(\frac{u - vi}{2}, \frac{ui + v}{2} \right) : u, v \in \mathbb{C} \right\}. \end{aligned}$$

Each of the factors here is isomorphic to \mathbb{C} as an \mathbb{R} -algebra.

Example 4.23

Let \mathbf{k} be a field and $p(x), q(x) \in \mathbf{k}[x]$ be coprime polynomials of positive degree. Consider the \mathbf{k} -algebra $A = \mathbf{k}[x]/(p(x)q(x))$ of dimension

$$\dim_{\mathbf{k}} A = \deg p(x) + \deg q(x).$$

Then there are polynomials $u(x), v(x) \in \mathbf{k}[x]$ for which

$$u(x)p(x) + v(x)q(x) = 1.$$

The polynomials $v(x)q(x)$ and $u(x)p(x)$ satisfy

$$\begin{aligned} (v(x)q(x))^2 &= v(x)q(x)(1 - u(x)p(x)) \\ &\equiv v(x)q(x) \pmod{(p(x)q(x))} \end{aligned}$$

and similarly

$$\begin{aligned} (u(x)p(x))^2 &= u(x)p(x)(1 - v(x)q(x)) \\ &\equiv u(x)p(x) \pmod{(p(x)q(x))}, \end{aligned}$$

so their residue classes

$$e_1 = v(x)q(x) + (p(x)q(x)), \quad e_2 = u(x)p(x) + (p(x)q(x))$$

are idempotents in A for which $e_1 + e_2 = 1$. Thus A can be expressed as a product

$$\mathbf{k}[x]/(p(x)q(x)) = \mathbf{k}[x]/(p(x)) \times \mathbf{k}[x]/(q(x)).$$

By Example 4.14, if $p(x)$ or $q(x)$ is irreducible then the corresponding factor $\mathbf{k}[x]/(p(x))$ or $\mathbf{k}[x]/(q(x))$ is simple. If both polynomials are irreducible then A is semi-simple.

Example 4.24

Let $\mathbf{k}[G]$ be the group algebra of a finite group G . Then for $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , $\mathbf{k}[G]$ is semi-simple. When $\mathbf{k} = \mathbb{C}$, indecomposable orthogonal idempotents can be written down in terms of the *irreducible complex characters* of G . These are important in the study of finite dimensional representations of G .

Example 4.25

Consider $\mathbb{C}[S_3]$, the group algebra of S_3 , the symmetric group on three objects. There are three central orthogonal idempotents here and

$$\mathbb{C}[S_3] = A_1 \times A_2 \times A_3,$$

where $A_1 \cong A_2 \cong \mathbb{C}$ and $A_3 \cong M_2(\mathbb{C})$. These idempotents can be found using character theory and turn out to be

$$\begin{aligned} e_1 &= \frac{1}{6} [1 + (1\ 2\ 3) + (1\ 3\ 2) + (1\ 2) + (1\ 3) + (2\ 3)], \\ e_2 &= \frac{1}{6} [1 + (1\ 2\ 3) + (1\ 3\ 2) - (1\ 2) - (1\ 3) - (2\ 3)], \\ e_3 &= \frac{1}{6} [2 - (1\ 2\ 3) - (1\ 3\ 2)]. \end{aligned}$$

In fact this also gives a decomposition of $\mathbb{R}[S_3]$ since each of these idempotents is an element of this \mathbb{R} -algebra. We obtain

$$\mathbb{R}[S_3] = A'_1 \times A'_2 \times A'_3,$$

where $A'_1 \cong A'_2 \cong \mathbb{R}$ but this time A'_3 has to be one of the \mathbb{R} -algebras $M_2(\mathbb{R})$ or \mathbb{H} (it is in fact $M_2(\mathbb{R})$).

Sometimes it is useful to have methods of deciding whether a \mathbf{k} -algebra can be a division algebra. Here is one useful criterion that can be applied to show that the factor A'_3 in Example 4.25 cannot be \mathbb{H} .

Lemma 4.26

Let \mathbb{D} be a \mathbf{k} -division algebra. Suppose that $\alpha \in \mathbb{D}$ is a root of a polynomial $f(X) \in \mathbf{k}[X]$ which factorises completely into linear factors over \mathbf{k} . Then $\alpha \in \mathbf{k}$.

Proof

Let $r_1, \dots, r_d \in \mathbf{k}$ be the roots of $f(X)$ in \mathbf{k} (with multiplicity), so that

$$f(X) = (X - r_1) \cdots (X - r_d).$$

Then since $f(\alpha) = 0$ and α commutes with elements of \mathbf{k} ,

$$(\alpha - r_1) \cdots (\alpha - r_d) = 0.$$

For each $j = 1, \dots, d$, either $\alpha = r_j$ or we can multiply by $(\alpha - r_j)^{-1}$. If $\alpha \neq r_j$ for all j , then we obtain $1 = 0$, hence for at least one value of j we must have $\alpha = r_j$. \square

Remark 4.27

If a \mathbf{k} -algebra A is a product $A = A_1 \times \cdots \times A_\ell$, then its group of units is also a product of groups,

$$A^\times = A_1^\times \times \cdots \times A_\ell^\times.$$

So if looking for interesting groups of units in algebras, we only need to consider simple algebras. It turns out that the real Clifford algebras introduced in Section 5.1 occur as factors of finite group algebras, for example the spinor groups could be found in this way.

Not all algebras are semi-simple as the following examples show. To understand them, notice that if $A = A_1 \times \cdots \times A_\ell$ is a product of simple algebras A_j , then for each $k = 1, \dots, \ell$, there is a two-sided ideal $J_k \triangleleft A$ with quotient algebra $A/J_k \cong A_k$. This provides a useful criterion for determining when an algebra is *not* semi-simple.

Lemma 4.28

Let A be a finite dimensional \mathbf{k} -algebra and suppose that A has only one proper ideal $J \triangleleft A$ for which the quotient A/J is simple. If $J \neq (0)$ then A is not semi-simple.

Proof

If A is semi-simple then it must be simple, hence its only proper ideal is (0) . \square

A very general class of examples is furnished by the next result.

Proposition 4.29

Let $p(X) \in \mathbf{k}[x]$ be an irreducible polynomial of positive degree. Then the commutative \mathbf{k} -algebra $A = \mathbf{k}[x]/(p(x)^m)$ is

- simple if $m = 1$;

- not semi-simple if $m > 1$.

Proof

To see this, first notice that if $m = 1$, then A is actually a field, hence is a simple \mathbf{k} -algebra. If $m > 1$, then the ideal $\overline{J} = (\overline{p(x)}) \triangleleft \mathbf{k}[x]/(p(x)^m)$ is the unique maximal ideal in $A = \mathbf{k}[x]/(p(x)^m)$, where $\overline{f(x)} = f(x) + (p(x)^m)$ is the residue class of $f(x)$ modulo $(p(x)^m)$. As $A/(\overline{p(x)}) \cong \mathbf{k}[x]/(p(x))$ is a field it is simple, allowing us to apply Lemma 4.28.

When $m > 1$, the group of units can be shown to be

$$A^\times = \mathbf{k}[x]/(p(x)^m)^\times \cong (\mathbf{k}[x]/(p(x)))^\times \times (1 + (\overline{p(x)})),$$

where

$$1 + (\overline{p(x)}) = \{\overline{1 + f(x)p(x)} \in \mathbf{k}[x]/(p(x)^m) : f(x) \in \mathbf{k}[x]\}$$

is a subgroup of $(\mathbf{k}[x]/(p(x)^m))^\times$. There is an isomorphism of groups

$$(1 + (\overline{p(x)})) \cong \mathbf{k}[x]/(p(x)^{m-1}); \quad \overline{1 + f(x)p(x)} \longleftrightarrow f(x) + (p(x)^{m-1}),$$

giving rise to an isomorphism of groups

$$\mathbf{k}[x]/(p(x)^m)^\times \cong (\mathbf{k}[x]/(p(x)))^\times \times \mathbf{k}[x]/(p(x)^{m-1}),$$

where $\mathbf{k}[x]/(p(x)^{m-1})$ is a group under addition. □

Here is an explicit example of Proposition 4.29.

Example 4.30

Let $\mathbf{k} = \mathbb{R}$ and $A = \mathbb{R}[x]/((x^2 + 1)^2)$. Then

$$\begin{aligned} A^\times &\cong (\mathbb{R}[x]/(x^2 + 1))^\times \times (1 + (\overline{x^2 + 1})) \\ &\cong (\mathbb{R}[x]/(x^2 + 1))^\times \times \mathbb{R}[x]/(x^2 + 1). \end{aligned}$$

Since $\mathbb{R}[x]/(x^2 + 1) \cong \mathbb{C}$ as \mathbb{R} -algebras, we have

$$\mathbb{R}[x]/((x^2 + 1)^2)^\times \cong \mathbb{C}^\times \times \mathbb{C}.$$

4.2 Real and Complex Normed Algebras

In this section, $\mathbf{k} = \mathbb{R}$ or \mathbb{C} . For a detailed look at normed algebras, see [21, 22].

Definition 4.31

Let A be a finite dimensional \mathbf{k} -algebra and $\nu: A \rightarrow \mathbb{R}$ be a function. Then ν is a **\mathbf{k} -norm** and the pair (A, ν) is called a **normed \mathbf{k} -algebra**, if $\nu(a) \geq 0$ for all $a \in A$ and ν satisfies the conditions

- i) for $t \in \mathbf{k}$, $a \in A$, $\nu(ta) = |t|\nu(a)$;
- ii) for $a, b \in A$, $\nu(ab) \leq \nu(a)\nu(b)$;
- iii) for $a, b \in A$, $\nu(a+b) \leq \nu(a) + \nu(b)$;
- iv) for $a \in A$, $\nu(a) = 0$ if and only if $a = 0$;
- v) $\nu(1) = 1$.

Actually, there is a certain amount of variation in the definition of a normed algebra, with some authors using the phrase to denote an algebra with a norm satisfying $\nu(ab) \leq \nu(a)\nu(b)$ for all $a, b \in A$, rather than the weaker requirement of Definition 4.31(ii), while others do not insist on condition (v).

Example 4.32

For $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , let $\mathbf{k}[G]$ be the group algebra of a finite group G . Define $\nu: \mathbf{k}[G] \rightarrow \mathbb{R}$ by

$$\nu\left(\sum_{g \in G} x_g g\right) = \sqrt{\sum_{g \in G} |x_g|^2}.$$

Then it is easy to verify that ν is a norm on $\mathbf{k}[G]$, the most interesting condition being the submultiplicative identity of Definition 4.31(ii) which follows from the calculation

$$\begin{aligned} \nu\left(\left(\sum_{g \in G} x_g g\right)\left(\sum_{h \in G} y_h h\right)\right)^2 &= \nu\left(\sum_{\substack{g \in G \\ h \in G}} x_{gh^{-1}} y_h g\right)^2 \\ &= \sum_{g \in G} \left| \sum_{h \in G} x_{gh^{-1}} y_h \right|^2 \\ &\leq \sum_{\substack{g \in G \\ h \in G}} |x_{gh^{-1}}|^2 |y_h|^2 \\ &= \nu\left(\sum_{g \in G} x_g g\right)^2 \nu\left(\sum_{h \in G} y_h h\right)^2. \end{aligned}$$

Notice that each $u \in G$ acts on $\mathbf{k}[G]$ by left multiplication preserving ν :

$$\begin{aligned}\nu(u \sum_{g \in G} x_g g) &= \nu(\sum_{g \in G} x_g (ug)) \\ &= \nu(\sum_{g \in G} x_{u^{-1}g} g) \\ &= \nu(\sum_{g \in G} x_g g).\end{aligned}$$

Similarly the action by right multiplication preserves ν . Finally, conjugation by an element $u \in G$ preserves ν and also gives a \mathbf{k} -algebra isomorphism.

By forgetting the multiplication, a normed \mathbf{k} -algebra (A, ν) can be viewed as a normed vector space in the sense of Definition 1.51. There is a natural metric ρ_ν on A for which

$$\rho_\nu(a, b) = \nu(a - b)$$

which allows us to introduce topological and analytic ideas when studying (A, ν) . In particular, A is complete with respect to this metric, i.e., all Cauchy sequences in A converge. In particular, for each $a \in A$, the exponential series

$$\exp(a) = \sum_{k=0}^{\infty} \frac{1}{k!} a^k$$

converges. Other analytic constructions generalise to this setting. For example, given a function $\alpha: (r, s) \rightarrow A$, where $r, s \in \mathbb{R}$ with $r < s$, we define its derivative at $t \in (r, s)$ by

$$\alpha'(t) = \lim_{h \rightarrow 0} \frac{1}{h} (\alpha(t + h) - \alpha(t)),$$

provided this limit is defined. It is straightforward to carry over the ideas of Chapters 1 and 3 to this situation. In particular, the group of units $A^\times \subseteq A$ is an open subset and the Lie algebra is equal to A with the Lie bracket defined by taking commutators in the algebra A , i.e.,

$$[x, y] = xy - yx \quad (x, y \in A).$$

A closed subgroup $G \leqslant A^\times$ has a Lie algebra which is an \mathbb{R} -Lie subalgebra of A . We leave the reader to work through the details.

If (A, ν) is a normed \mathbf{k} -algebra then we can define a version of the operator norm $\|\cdot\|_\nu$ on A by

$$\|a\|_\nu = \sup\{\nu(au) : u \in A, \nu(u) = 1\}.$$

In general this does not agree with the norm ν . It does provide another normed algebra $(A, \|\cdot\|_\nu)$ with a metric defined by

$$\rho'_\nu(a, b) = \|a - b\|_\nu.$$

Proposition 4.33

Let (A, ν) be a normed \mathbf{k} -algebra.

- i) The identity map $\text{Id}_A: A \rightarrow A$ is a homeomorphism of metric spaces $(A, \rho_\nu) \rightarrow (A, \rho'_\nu)$.
- ii) The left regular representation $\lambda: A \rightarrow M_n(\mathbf{k})$ with respect to any \mathbf{k} -basis $\{v_1, \dots, v_n\}$ of A gives rise to continuous injections of metric spaces $(A, \rho_\nu) \rightarrow (A, \|\cdot\|)$ and $(A, \rho'_\nu) \rightarrow (A, \|\cdot\|)$.

Proof

These results follow from Theorem 1.53 and Corollary 1.54. \square

4.3 Linear Algebra over a Division Algebra

Throughout this section, let \mathbf{D} be a finite dimensional division algebra over a field \mathbf{k} .

Definition 4.34

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

$$\begin{aligned} v(xy) &= (vx)y, \\ v(x+y) &= vx + vy, \\ (u+v)x &= ux + vx, \\ v1 &= v. \end{aligned}$$

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

Theorem 4.35

Let V be a \mathbf{D} -vector space.

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

Definition 4.36

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 $n \geq 1$, 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

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} w = \begin{bmatrix} z_1 w \\ z_2 w \\ \vdots \\ z_n w \end{bmatrix} \quad (z_1, \dots, z_n, w \in \mathbb{D}).$$

Proposition 4.37

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 $\{v_1, \dots, v_m\}$, $\{w_1, \dots, w_n\}$. Then a \mathbb{D} -linear transformation $\varphi: V \rightarrow W$ is given by

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

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

$$\varphi \left(\sum_{s=1}^m v_s x_s \right) = \sum_{s=1}^m w_s y_s,$$

then

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

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.

Of course, this is 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.

Our main interest is in linear transformations which we will identify with the corresponding matrices. If $\theta: \mathbb{D}^k \rightarrow \mathbb{D}^m$ and $\varphi: \mathbb{D}^m \rightarrow \mathbb{D}^n$ are \mathbb{D} -linear transformations with corresponding matrices $[\theta], [\varphi]$, then

$$[\theta][\varphi] = [\theta \circ \varphi]. \tag{4.1}$$

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

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

$$\begin{aligned} v_r b_t & \quad (r = 1, \dots, m, t = 1, \dots, d), \\ w_s b_t & \quad (s = 1, \dots, n, t = 1, \dots, d) \end{aligned}$$

form \mathbf{k} -bases for V, W as \mathbf{k} -vector spaces.

We denote the set of all $m \times n$ matrices with entries in \mathbb{D} by $M_{m,n}(\mathbb{D})$ and $M_n(\mathbb{D}) = M_{n,n}(\mathbb{D})$. Then $M_n(\mathbb{D})$ is a \mathbf{k} -algebra of dimension

$$\dim_{\mathbf{k}} M_n(\mathbb{D}) = n^2 \dim_{\mathbf{k}} \mathbb{D}.$$

The group of units $M_n(\mathbb{D})^\times$ is usually denoted $GL_n(\mathbb{D})$. However, for non-commutative \mathbb{D} there is no determinant function so we cannot define an analogue of the special linear group. However, we can use the left regular representation to circumvent this problem.

Proposition 4.38

Let A be an algebra over a field \mathbf{k} and $B \subseteq A$ be 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 ($k \geq 0$) are linearly dependent over \mathbf{k} , so for some $t_r \in \mathbf{k}$ ($r = 0, \dots, \ell$) with $t_\ell \neq 0$ and $\ell \geq 1$, there is a non-trivial relation of the form

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

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

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

hence u has an inverse in B . \square

For a division algebra D , each matrix $A \in M_n(D)$ acts by multiplication on the left of D^n . For any subfield $k_1 \subseteq D$ containing k , A induces a (right) k_1 -linear transformation,

$$D^n \longrightarrow D^n; \quad x \mapsto Ax.$$

If we choose a k_1 -basis for D , A gives rise to a matrix $\Lambda_A \in M_{nd}(k_1)$ where $d = \dim_{k_1} D_{k_1}$. It is easy to see that the function

$$\Lambda: M_n(D) \longrightarrow M_{nd}(k_1); \quad \Lambda(A) = \Lambda_A,$$

is a ring homomorphism with $\ker \Lambda = 0$. This allows us to identify $M_n(D)$ with the subring $\text{im } \Lambda \subseteq M_{nd}(k_1)$.

Applying Proposition 4.38 we see that A is invertible in $M_n(D)$ if and only if Λ_A is invertible in $M_{nd}(k_1)$. But the latter is true if and only if $\det \Lambda_A \neq 0$.

Hence to determine invertibility of $A \in M_n(D)$, it suffices to consider $\det \Lambda_A$ using a subfield k_1 . The resulting function

$$\text{Rdet}_{k_1}: M_n(D) \longrightarrow k_1; \quad \text{Rdet}_{k_1}(A) = \det \Lambda_A,$$

is called the k_1 -reduced determinant of $M_n(D)$ and is a group homomorphism. It is actually true that $\det \Lambda_A \in k$, not just in k_1 , although we will not prove this here.

Proposition 4.39

$A \in M_n(D)$ is invertible if and only if $\text{Rdet}_{k_1}(A) \neq 0$ for some subfield $k_1 \subseteq D$ containing k .

4.4 The Quaternions

Proposition 4.40

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 $A \cong \mathbf{C}$.

Proof

Let $\alpha \in A$. Since A is a finite dimensional \mathbb{R} -vector space, the powers of α must be linearly dependent, say

$$t_0 + t_1\alpha + \cdots + t_m\alpha^m = 0 \quad (4.2)$$

for some $t_i \in \mathbb{R}$ with $m \geq 1$ 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 the minimality of m ; 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. To see this, suppose that $p(X) = p_1(X)p_2(X)$; then as A is a division algebra, either $p_1(\alpha) = 0$ or $p_2(\alpha) = 0$, contradicting minimality of m if both $\deg p_1(X) > 0$ and $\deg p_2(X) > 0$.

Consider the \mathbb{R} -subspace

$$\mathbb{R}(\alpha) = \left\{ \sum_{j=0}^k s_j\alpha^j : s_j \in \mathbb{R} \right\} \subseteq A.$$

Then $\mathbb{R}(\alpha)$ is easily seen to be an \mathbb{R} -subalgebra of A and the elements $1, \alpha, \dots, \alpha^{m-1}$ form a basis by Equation (4.2), hence $\dim_{\mathbb{R}} \mathbb{R}(\alpha) = m$.

Let $\gamma \in \mathbb{C}$ be any complex root of the irreducible polynomial

$$t_0 + t_1X + \cdots + t_mX^m \in \mathbb{R}[X]$$

(such a root certainly exists by the Fundamental Theorem of Algebra). There is an \mathbb{R} -linear transformation which is actually an injection,

$$\varphi: \mathbb{R}(\alpha) \longrightarrow \mathbb{C}; \quad \varphi\left(\sum_{j=0}^{m-1} s_j\alpha^j\right) = \sum_{j=0}^{m-1} s_j\gamma^j.$$

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 $m = \dim_{\mathbb{R}} \mathbb{R}(\alpha) \leq 2$. If $m = 1$, then by Equation (4.2), $\alpha \in \mathbb{R}$. 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

$$q(X) = u_0 + u_1X + \cdots + u_\ell X^\ell \in \mathbb{C}[X]$$

with $\ell \geq 1$ 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 $\dim_{\mathbb{R}} A > 1$. \square

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, up to isomorphism there is only one other, the skew field of *quaternions* (or *Hamiltonians* after their discoverer William Rowan Hamilton), usually denoted \mathbb{H} . We will construct \mathbb{H} as a ring of 2×2 complex matrices. They provide the first example of a non-commutative *Clifford algebra* which we will define later in Chapter 5.

Let

$$\mathbb{H} = \left\{ \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} : z, w \in \mathbb{C} \right\} \subseteq M_2(\mathbb{C}).$$

It is easy to see that \mathbb{H} is a subring of $M_2(\mathbb{C})$ and is in fact an \mathbb{R} -subalgebra where we view $M_2(\mathbb{C})$ as an \mathbb{R} -algebra of dimension 8. It also contains a copy of \mathbb{C} , namely the \mathbb{R} -subalgebra

$$\left\{ \begin{bmatrix} z & 0 \\ 0 & \bar{z} \end{bmatrix} : z \in \mathbb{C} \right\} \subseteq \mathbb{H}.$$

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

$$\begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} \begin{bmatrix} \bar{z} & -w \\ \bar{w} & z \end{bmatrix} = \begin{bmatrix} |z|^2 + |w|^2 & 0 \\ 0 & |z|^2 + |w|^2 \end{bmatrix},$$

hence $\begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix}$ is invertible if and only if $\begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} \neq O$; furthermore in that case,

$$\begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{\bar{z}}{|z|^2 + |w|^2} & \frac{-w}{|z|^2 + |w|^2} \\ \frac{\bar{w}}{|z|^2 + |w|^2} & \frac{z}{|z|^2 + |w|^2} \end{bmatrix}$$

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

$$SU(2) = \{A \in \mathbb{H} : \det A = 1\} \leqslant \mathbb{H}^\times.$$

It is useful to define on \mathbb{H} a norm in the sense of Proposition 1.5:

$$\left\| \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} \right\| = \det \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} = |z|^2 + |w|^2.$$

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

$$1 = I, \quad i = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad j = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad k = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

These satisfy the equations

$$i^2 = j^2 = k^2 = -1, \quad ij = k = -ji, \quad k = i = -kj, \quad ki = j = -ik.$$

This should be compared with the vector product on \mathbb{R}^3 as discussed in Example 3.2. From now on we will write quaternions in the form

$$q = xi + yj + zk + t1 \quad (x, y, z, t \in \mathbb{R}).$$

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 $xi + yj + zk$ with the element $xe_1 + ye_2 + ze_3 \in \mathbb{R}^3$. Using this identification we see that the scalar and vector products on \mathbb{R}^3 are related to quaternion multiplication as in the following result.

Proposition 4.41

For two pure quaternions $q_1 = x_1i + y_1j + z_1k$, $q_2 = x_2i + y_2j + z_2k$,

$$q_1 q_2 = -(x_1i + y_1j + z_1k) \cdot (x_2i + y_2j + z_2k) + (x_1i + y_1j + z_1k) \times (x_2i + y_2j + z_2k).$$

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 next result describes the general solution of the equation $X^2 + 1 = 0$ in \mathbb{H} .

Proposition 4.42

The quaternion $q = xi + yj + zk + t1$ satisfies $q^2 + 1 = 0$ if and only if $t = 0$ and $x^2 + y^2 + z^2 = 1$.

Proof

This easily follows from Proposition 4.41. □

There is a quaternionic analogue of complex conjugation, namely

$$q = xi + yj + zk + t1 \mapsto \bar{q} = q^* = -xi - yj - zk + t1.$$

This is 'almost' a ring homomorphism $\mathbb{H} \rightarrow \mathbb{H}$, in fact it satisfies

$$\overline{(q_1 + q_2)} = \bar{q}_1 + \bar{q}_2; \quad (4.3a)$$

$$\overline{(q_1 q_2)} = \bar{q}_2 \bar{q}_1; \quad (4.3b)$$

$$\bar{q} = q \iff q \text{ is a real quaternion}; \quad (4.3c)$$

$$\bar{q} = -q \iff q \text{ is a pure quaternion}. \quad (4.3d)$$

Because of Equation (4.3b) this is called an *anti-homomorphism of skew rings* or *ring anti-homomorphism*. The inverse of a non-zero quaternion q can be written as

$$q^{-1} = \frac{1}{(q\bar{q})} \bar{q} = \frac{\bar{q}}{(q\bar{q})}. \quad (4.4)$$

The real quantity $q\bar{q}$ is the square of the length of the corresponding vector,

$$|q| = \sqrt{q\bar{q}} = \sqrt{x^2 + y^2 + z^2 + t^2}.$$

For $z = u + vi$ with $u, v \in \mathbb{R}$, $\bar{z} = u1 - vi$ is the usual complex conjugation.

In terms of the matrix description of \mathbb{H} , quaternionic conjugation is given by hermitian conjugation,

$$\begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} \mapsto \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix}^* = \begin{bmatrix} \bar{z} & -w \\ \bar{w} & z \end{bmatrix}.$$

To simplify notation, from now on we will write

$$1 = 1, \quad i = i, \quad j = j, \quad k = k.$$

4.5 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 a quaternionic inner product on \mathbb{H} by

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^* \mathbf{y} = \sum_{r=1}^n \bar{x}_r y_r,$$

where we define the *quaternionic conjugate* of a vector by

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^* = [\bar{x}_1 \quad \bar{x}_2 \quad \cdots \quad \bar{x}_n].$$

Similarly, for any matrix $[a_{ij}]$ over \mathbb{H} we can define $[a_{ij}]^* = [\bar{a}_{ji}]$.

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 $M_n(\mathbb{H})$ by the method used in Section 1.2, i.e., for $A \in M_n(\mathbb{H})$,

$$\|A\| = \sup \left\{ \frac{|Ax|}{|\mathbf{x}|} : \mathbf{0} \neq \mathbf{x} \in \mathbb{H}^n \right\}.$$

The analogue of Proposition 1.5 holds for $\|\cdot\|$ and the norm $|\cdot|$ on \mathbb{H} , although statements involving scalar multiplication need to be formulated with scalars on the right. There is also a resulting metric on $M_n(\mathbb{H})$,

$$(A, B) \mapsto \|A - B\|,$$

and we can use this to do analysis on $M_n(\mathbb{H})$. The multiplication map

$$M_n(\mathbb{H}) \times M_n(\mathbb{H}) \rightarrow M_n(\mathbb{H})$$

is again continuous, and the group of invertible elements $GL_n(\mathbb{H}) \subseteq M_n(\mathbb{H})$ is actually an open subset. This can be proved using either of the reduced determinants

$$Rdet_{\mathbb{R}}: M_n(\mathbb{H}) \rightarrow \mathbb{R}, \quad Rdet_{\mathbb{C}}: M_n(\mathbb{H}) \rightarrow \mathbb{C},$$

each of which is continuous. By Proposition 4.39,

$$GL_n(\mathbb{H}) = M_n(\mathbb{H}) - Rdet_{\mathbb{R}}^{-1} 0, \tag{4.5a}$$

$$GL_n(\mathbb{H}) = M_n(\mathbb{H}) - Rdet_{\mathbb{C}}^{-1} 0. \tag{4.5b}$$

In either case we see that $GL_n(\mathbb{H})$ is an open subset of $M_n(\mathbb{H})$. It is also possible to show that the two embeddings

$$GL_n(\mathbb{H}) \rightarrow GL_{4n}(\mathbb{R}), \quad GL_n(\mathbb{H}) \rightarrow GL_{2n}(\mathbb{C}),$$

have closed images. So $GL_n(\mathbb{H})$ and its closed subgroups are real and complex matrix groups.

Definition 4.43

The $n \times n$ quaternionic symplectic group is

$$Sp(n) = \{A \in GL_n(\mathbb{H}) : A^* A = I\} \leqslant GL_n(\mathbb{H}).$$

Then $\mathrm{Sp}(n)$ is easily seen to satisfy

$$\mathrm{Sp}(n) = \{A \in \mathrm{GL}_n(\mathbb{H}) : \forall \mathbf{x}, \mathbf{y} \in \mathbb{H}^n, A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}\}. \quad (4.6)$$

Hence we have the following proposition.

Proposition 4.44

The Lie algebra of $\mathrm{Sp}(n)$ is

$$\mathfrak{sp}(n) = \{Q \in \mathrm{M}_n(\mathbb{H}) : Q^* = -Q\}$$

and this has dimension

$$\dim \mathrm{Sp}(n) = \dim_{\mathbb{R}} \mathfrak{sp}(n) = 2n^2 + n.$$

Proof

The equation $Q^* = -Q$ has solution space of dimension

$$3n + 4 \binom{n}{2} = 3n + 2(n^2 - n) = 2n^2 + n$$

in $\mathrm{M}_n(\mathbb{H})$. □

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 .

4.6 Automorphism Groups of Algebras

Let A be a finite dimensional \mathbf{k} -algebra. We can view A as a \mathbf{k} -vector space and consider its general linear group $\mathrm{GL}_{\mathbf{k}}(A)$.

Definition 4.45

The group of \mathbf{k} -algebra automorphisms of A is

$$\mathrm{Aut}_{\mathbf{k}}(A) = \{\alpha \in \mathrm{GL}_{\mathbf{k}}(A) : \alpha \text{ is an algebra automorphism}\}.$$

It is easy to see that $\text{Aut}_{\mathbf{k}}(A)$ is a group and $\text{Aut}_{\mathbf{k}}(A) \leqslant \text{GL}_{\mathbf{k}}(A)$.

If A has an algebra norm ν then by Theorem 1.53, we can give $\text{GL}_{\mathbf{k}}(A)$ and hence $\text{Aut}_{\mathbf{k}}(A)$, the metric associated to the operator norm $\|\cdot\|_{\nu}$, introduced in Section 1.8. Using these metrics we obtain the following propositions.

Proposition 4.46

$\text{Aut}_{\mathbf{k}}(A)$ is a closed subgroup of $\text{GL}_{\mathbf{k}}(A)$, hence is a matrix subgroup.

Proof

Here is one approach. Choose a basis $\{v_1, \dots, v_n\}$ for A and use this to identify A with \mathbf{k}^n and then produce an isomorphism of matrix groups $\text{GL}_{\mathbf{k}}(A) \rightarrow \text{GL}_n(\mathbf{k})$ as in Proposition 1.56.

Let $\mu: \mathbf{k}^n \times \mathbf{k}^n \rightarrow \mathbf{k}^n$ be the \mathbf{k} -bilinear map corresponding to the product A . Then for $\alpha \in \text{GL}_n(\mathbf{k})$, the condition that it corresponds to an element of $\text{Aut}_{\mathbf{k}}(A)$ is that

$$\mu \circ (\alpha \times \alpha) = \alpha \circ \mu,$$

and writing $\alpha = [a_{ij}]$, this amounts to n polynomial equations for the a_{ij} . By the sort of considerations we used to identify matrix subgroups in Chapter 1, this shows that $\text{Aut}_{\mathbf{k}}(A) \leqslant \text{GL}_n(\mathbf{k})$ is a closed subgroup. \square

Example 4.47

Consider $A = \mathbf{C}$ as an \mathbb{R} -algebra. Then $\text{Aut}_{\mathbb{R}}(\mathbf{C}) = \{\text{Id}_{\mathbf{C}}, \overline{(\cdot)}\}$, where $\overline{(\cdot)}$ denotes complex conjugation.

Proof

Let $\alpha: \mathbf{C} \rightarrow \mathbf{C}$ be an \mathbb{R} -algebra automorphism. Every complex number z has the form $z = x + yi$ with $x, y \in \mathbb{R}$, hence

$$\alpha(z) = \alpha(x + yi) = \alpha(x) + \alpha(y)\alpha(i) = x + y\alpha(i).$$

So it suffices to determine $\alpha(i)$. But applying α to the equation $i^2 + 1 = 0$, we obtain

$$\alpha(i)^2 + 1 = 0,$$

so $\alpha(i) = \pm i$. Hence either $\alpha = \text{Id}_{\mathbf{C}}$ or $\alpha = \overline{(\cdot)}$. \square

This example is rather simple because the algebra is commutative. Non-commutative algebras have lots of automorphisms obtained using conjugation

in the group theoretic sense. Let $u \in A^\times$ be a unit in the \mathbf{k} -algebra A . The function

$$\chi_u: A \longrightarrow A; \quad \chi_u(a) = uau^{-1}$$

is clearly a \mathbf{k} -linear isomorphism. It is also easy to check that for $a, b \in A$ and $t \in \mathbf{k}$,

$$\chi_u(ab) = \chi_u(a)\chi_u(b).$$

Hence χ_u is a \mathbf{k} -algebra automorphism of A .

Definition 4.48

An automorphism α of the \mathbf{k} -algebra A is called an *inner automorphism* of A if it has the form χ_u for some unit $u \in A^\times$. The inner automorphisms form a subgroup $\text{Inn}_\mathbf{k}(A) \leqslant \text{Aut}_\mathbf{k}(A)$. Automorphisms which are not inner are called *outer automorphisms*.

Of course, if A is commutative, the only inner automorphism is the identity Id_A . For the \mathbb{R} -algebra of complex numbers, complex conjugation $(\bar{})$ is an outer automorphism. The construction of inner automorphisms defines a function

$$\chi: A^\times \longrightarrow \text{Aut}_\mathbf{k}(A); \quad u \longmapsto \chi_u.$$

The proof of the next result is left as an exercise.

Proposition 4.49

Let A be a finite dimensional \mathbf{k} -algebra.

i) The function $\chi: A^\times \longrightarrow \text{Aut}_\mathbf{k}(A)$ is a continuous group homomorphism, with kernel and image

$$\ker \chi = Z(A^\times) = \text{the centre of } A^\times, \quad \text{im } \chi = \text{Inn}_\mathbf{k}(A) \leqslant \text{Aut}_\mathbf{k}(A).$$

ii) The subgroup of inner automorphisms $\text{Inn}_\mathbf{k}(A) \leqslant \text{Aut}_\mathbf{k}(A)$ is a closed normal subgroup, hence $\text{Inn}_\mathbf{k}(A)$ is a matrix group.

Definition 4.50

For a finite dimensional \mathbf{k} -algebra A , the *outer automorphism group* is the quotient group

$$\text{Out}_\mathbf{k}(A) = \text{Aut}_\mathbf{k}(A)/\text{Inn}_\mathbf{k}(A).$$

This has a natural *quotient topology*, which we discuss in Section 8.1.

Example 4.51

Let $A = M_n(\mathbf{k})$. Then

$$\text{Inn}_{\mathbf{k}}(M_n(\mathbf{k})) = \text{GL}_n(\mathbf{k}) / \{zI_n : z \in \mathbf{k}^\times\} = \text{PGL}_n(\mathbf{k}),$$

the $n \times n$ projective linear group.

Example 4.52

For the \mathbb{R} -algebra of quaternions \mathbb{H} ,

$$\text{Aut}_{\mathbb{R}}(\mathbb{H}) = \text{Inn}_{\mathbb{R}}(\mathbb{H}) \cong \text{SO}(3).$$

Proof

Let $\alpha \in \text{Aut}_{\mathbb{R}}(\mathbb{H})$. Then by Propositions 4.41 and 4.42, $\alpha(i), \alpha(j), \alpha(k)$ are pure quaternions satisfying

$$\alpha(i)^2 = \alpha(j)^2 = \alpha(k)^2 = -1, \quad \alpha(i)\alpha(j) = \alpha(k).$$

In particular, viewed as elements of the real vector space \mathbb{R}^3 with usual basis $e_1 = i, e_2 = j$ and $e_3 = k$, these vectors form an orthonormal basis, hence there is an orthogonal matrix $A \in O(3)$ for which

$$\alpha(i) = Ai, \quad \alpha(j) = A, \quad \alpha(k) = Ak.$$

It is even true that $\det A = 1$ since in terms of the vector product \times we have

$$\alpha(i)\alpha(j) = \alpha(i) \times \alpha(j),$$

and it is standard that for a pair of orthonormal vectors u, v , the triple $u, v, u \times v$ is a left-handed orthonormal basis. Thus α is equivalent to a special orthogonal matrix $A \in SO(3)$ acting on \mathbb{R}^3 viewed as the pure quaternions.

To see that such an automorphism is actually inner, recall that in Section 3.5 we defined a surjective group homomorphism $\overline{\text{Ad}}: \text{SU}(2) \rightarrow \text{SO}(3)$. Careful comparison of the definitions of $\overline{\text{Ad}}$ and χ restricted to the subgroup of \mathbb{H}^\times consisting of unit quaternions shows that these are essentially the same mapping. Hence $\alpha \in \text{Inn}_{\mathbb{R}}(\mathbb{H})$.

So $\text{Aut}_{\mathbb{R}}(\mathbb{H}) = \text{Inn}_{\mathbb{R}}(\mathbb{H}) \cong \text{SO}(3)$ as claimed. □

Remark 4.53

Suppose that A is a finite dimensional simple \mathbf{k} -algebra where \mathbf{k} is actually the centre of A , i.e.,

$$\mathbf{k} = \{z \in A : \forall a \in A, az = za\}.$$

Then the important *Noether–Skolem Theorem* asserts that $\text{Aut}_{\mathbf{k}}(A) = \text{Inn}_{\mathbf{k}}(A)$, or equivalently that $\text{Out}_{\mathbf{k}}(A) = \{1\}$. This is true for all division algebras and matrix algebras over a division algebra, which gives a more theoretical explanation for part of Example 4.52 and also shows that in Example 4.51 we actually have $\text{Aut}_{\mathbf{k}}(\text{M}_n(\mathbf{k})) \cong \text{PGL}_n(\mathbf{k})$.

EXERCISES

- 4.1. Recall Example 4.25. Show that in $\mathbb{R}[\text{S}_3]$, the summand

$$A'_3 = \mathbb{R}[\text{S}_3]e_3 = e_3\mathbb{R}[\text{S}_3]$$

contains at least 4 elements of order 2. Deduce that A'_3 cannot be isomorphic to \mathbf{H} as an \mathbb{R} -algebra.

- 4.2. If $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , let (A, ν) be a normed \mathbf{k} -algebra. Consider the set

$$A_1^\times = \{u \in A^\times : \nu(u) = 1 = \nu(u^{-1})\}.$$

- a) Show that $A_1^\times \subseteq A$ is a compact subset with respect to the metric associated with ν .
- b) Show that A_1^\times is a subgroup of A^\times .
- c) Find A_1^\times in each of the following cases and determine its Lie algebra as a Lie subalgebra of A with its usual bracket: $A = \mathbb{C}$, $A = \mathbf{H}$, $A = \text{M}_n(\mathbf{k})$.

- 4.3. Show that for $r \in \mathbb{R}$, in the real division algebra of quaternions \mathbf{H} , the equation $x^2 = r$ has

- two solutions if $r > 0$;
- infinitely many solutions if $r < 0$ and these are all pure quaternions.

A good question to consider after reading Chapter 5 is whether this result generalises to an arbitrary Clifford algebra Cl_n .

- 4.4. Using the bases $\{1, i, j, k\}$ of $\mathbf{H}_{\mathbb{R}}$ over \mathbb{R} and $\{1, j\}$ of $\mathbf{H}_{\mathbb{C}}$ over \mathbb{C} , determine the reduced determinants $\text{Rdet}_{\mathbb{R}}: \text{GL}_n(\mathbf{H}) \rightarrow \mathbb{R}^\times$ and $\text{Rdet}_{\mathbb{C}}: \text{GL}_n(\mathbf{H}) \rightarrow \mathbb{C}^\times$ for small values of n .
- 4.5. a) Verify that $\text{M}_n(\mathbf{H})$ is complete with respect to the norm $\| \cdot \|$. Use this to define an exponential function $\exp: \text{M}_n(\mathbf{H}) \rightarrow \text{GL}_n(\mathbf{H})$ with properties analogous to those for the exponential functions on $\text{M}_n(\mathbb{R})$ and $\text{M}_n(\mathbb{C})$.
- b) 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 $|u| = 1$.

- 4.6. For each of the following matrix groups G , determine the Lie algebra \mathfrak{g} and dimension $\dim G$.

- $G = \mathrm{GL}_n(\mathbb{H})$.
- $G = \mathrm{Sp}_n(\mathbb{H})$.
- $G = \ker \mathrm{Rdet}_{\mathbf{k}}: \mathrm{GL}_n(\mathbb{H}) \rightarrow \mathbf{k}^\times$, where $\mathbf{k} = \mathbb{R}, \mathbb{C}$ and n is small.
- $G = \ker \mathrm{Rdet}_{\mathbf{k}}: \mathrm{Sp}(n) \rightarrow \mathbf{k}^\times$, where $\mathbf{k} = \mathbb{R}, \mathbb{C}$ and n is small.

4.7. 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

$$q \cdot x = qxq^{-1} = qx\bar{q} \quad (x \in \mathbb{H}).$$

- By identifying \mathbb{H} with \mathbb{R}^4 using the basis $\{i, j, k, 1\}$, show that this defines a Lie homomorphism $\mathrm{Sp}(1) \rightarrow \mathrm{SO}(4)$.
- 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 $\alpha: \mathrm{Sp}(1) \rightarrow \mathrm{SO}(3)$. Show that $\ker \alpha = \{1, -1\}$.

- 4.8. Δ Using the surjective homomorphism $\alpha: \mathrm{Sp}(1) \rightarrow \mathrm{SO}(3)$ of the previous question, for a subgroup $G \leqslant \mathrm{SO}(3)$ set

$$\tilde{G} = \alpha^{-1}G = \{g \in \mathrm{Sp}(1) : \alpha(g) \in G\} \leqslant \mathrm{Sp}(1).$$

From now on assume that G is finite.

- Determine the order of \tilde{G} .
- Show that the order of the centre of \tilde{G} , $Z(\tilde{G})$, is even.
- If G contains an element of order 2, show that the group homomorphism $\alpha: \tilde{G} \rightarrow G$ is not split in the sense that there is no group homomorphism $\beta: \tilde{G} \rightarrow G$ for which $\alpha \circ \beta = \mathrm{Id}_G$.

$$\begin{array}{ccc} G & \xrightarrow{\beta} & \tilde{G} \\ & \searrow \mathrm{Id}_G & \downarrow \alpha \\ & & G \end{array}$$

- d) 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 $\alpha Q_8 \leqslant \mathrm{SO}(3)$. Generalise this by considering for each $n \geqslant 2$,

$$Q_{2n} = \left\{ e^{2\pi ir/n} : r = 0, \dots, n-1 \right\} \cup \left\{ e^{2\pi ir/n} j : r = 0, \dots, n-1 \right\}.$$

e) Show that the set T_{24} consisting of the 24 elements

$$\pm 1, \pm i, \pm j, \pm k, \frac{1}{2}(\pm 1 \pm i \pm j \pm k),$$

is a subgroup of $\mathrm{Sp}(1)$ and find a geometric interpretation for the group $\alpha T_{24} \leq \mathrm{SO}(3)$.

f) $\Delta\Delta$ 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, $\mathrm{Symm}^+(\text{Icos}) \cong A_5$. Determine $\alpha^{-1} \mathrm{Symm}^+(\text{Icos}) \leq \mathrm{Sp}(1)$.

[This requires a good way to view the icosahedron relative to the x, y, z -axes. The resulting subgroup of $\mathrm{Sp}(1)$ is called the *binary icosahedral group* since it provides a double covering of the symmetry group $\mathrm{Symm}^+(\text{Icos})$. It also provides a non-split double covering $\tilde{A}_5 \rightarrow A_5$ of the simple group A_5 .]

4.9. $\Delta\Delta$ a) Show that for each $n \geq 1$,

$$\dim \mathrm{Sp}(n) = \mathrm{Symp}_{2n}(\mathbb{R}).$$

b) Notice that there are embeddings $\mathrm{Symp}_{2n}(\mathbb{R}) \leq \mathrm{GL}_{2n}(\mathbb{C})$ and $\mathrm{Sp}(n) \leq \mathrm{GL}_{2n}(\mathbb{C})$, hence $\mathfrak{sp}(n)_{\mathbb{C}} \leq \mathfrak{gl}_{2n}(\mathbb{C})$ and $\mathfrak{symp}_{2n}(\mathbb{R})_{\mathbb{C}} \leq \mathfrak{gl}_{2n}(\mathbb{C})$. Show that as \mathbb{C} -Lie algebras

$$\mathfrak{sp}(n)_{\mathbb{C}} \cong \mathfrak{symp}_{2n}(\mathbb{R})_{\mathbb{C}}.$$

4.10. Prove Proposition 4.49.

Clifford Algebras and Spinor Groups

In this chapter, we generalise the quaternions by studying the *real Clifford algebras*, and our account of these is heavily influenced by the classic paper of Atiyah, Bott & Shapiro [3]; Porteous [23, 24] also provides an accessible description, as does Curtis [7] but there are some errors and omissions in that account. Lawson & Michelsohn [19] provides a more sophisticated introduction which shows how central Clifford algebras have become to modern geometry and topology; they also appear in Quantum Theory in connection with the Dirac operator. There is also a theory of *Clifford Analysis* in which the field of complex numbers is replaced by a Clifford algebra and a suitable class of *Clifford analytic functions* generalising complex analytic functions is studied; motivation for this is provided by the above applications. The groups of units in Clifford algebras contain the *spinor groups* which we define and also show how they provide double coverings of the special orthogonal groups.

We restrict attention to real Clifford algebras associated to positive definite inner products on \mathbb{R}^n . There are also complex and indefinite Clifford algebras discussed in [3, 19, 23, 24].

5.1 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 for $n \geq 0$, where the first few satisfy

$$\text{Cl}_0 = \mathbb{R}, \quad \text{Cl}_1 \cong \mathbb{C}, \quad \text{Cl}_2 \cong \mathbb{H},$$

as \mathbb{R} -algebras.

We begin by describing Cl_n as an \mathbb{R} -vector space and define the product in terms of a particular basis. There are elements $e_1, e_2, \dots, e_n \in \text{Cl}_n$ for which

$$\begin{cases} e_s e_r = -e_r e_s & \text{if } s \neq r, \\ e_r^2 = -1. \end{cases} \quad (5.1)$$

Moreover, Cl_n has an \mathbb{R} -basis consisting of the elements $e_{i_1} e_{i_2} \cdots e_{i_r}$ corresponding to increasing sequences $1 \leq i_1 < i_2 < \cdots < i_r \leq n$ with $0 \leq r \leq n$. Thus

$$\dim_{\mathbb{R}} \text{Cl}_n = 2^n. \quad (5.2)$$

When $r = 0$, the element $e_{i_1} e_{i_2} \cdots e_{i_r}$ is taken to be 1.

Doing calculations with these basis elements is easy as the following example illustrates.

Example 5.1

Let i, j, k, ℓ be distinct numbers in the range 1 to n . Then

$$e_i e_j e_k e_\ell = e_k e_\ell e_i e_j.$$

Proof

Repeatedly using the relations of (5.1) we obtain

$$e_i e_j e_k e_\ell = -e_i e_k e_j e_\ell = e_k e_i e_j e_\ell = -e_k e_i e_\ell e_j = e_k e_\ell e_i e_j.$$

□

Proposition 5.2

There are isomorphisms of \mathbb{R} -algebras

$$\text{Cl}_1 \cong \mathbb{C}, \quad \text{Cl}_2 \cong \mathbb{H}.$$

Proof

For Cl_1 , the function

$$\text{Cl}_1 \longrightarrow \mathbb{C}; \quad x + ye_1 \longmapsto x + yi \quad (x, y \in \mathbb{R}),$$

is an \mathbb{R} -linear ring isomorphism.

Similarly, for Cl_2 , the function

$$\text{Cl}_2 \longrightarrow \mathbb{H}; \quad t1 + xe_1 + ye_2 + ze_1e_2 \longmapsto t1 + xi + yj + zk \quad (t, x, y, z \in \mathbb{R}),$$

is an \mathbb{R} -linear ring isomorphism. \square

We can order the basis monomials in the e_r by declaring $e_{i_1}e_{i_2}\cdots e_{i_r}$ to be numbered

$$1 + 2^{i_1-1} + 2^{i_2-1} + \cdots + 2^{i_r-1},$$

which should be interpreted as 1 when $r = 0$. Every integer k for which $1 \leq k \leq 2^n$ has a unique *binary expansion*

$$k = k_0 + 2k_1 + \cdots + 2^jk_j + \cdots + 2^nk_n,$$

where each $k_j = 0, 1$. This provides a one to one correspondence between such numbers k and the basis monomials of Cl_n . Here are the orderings of the bases for the first few Clifford algebras.

Cl_1	1, e_1
Cl_2	1, e_1, e_2, e_1e_2
Cl_3	1, $e_1, e_2, e_1e_2, e_3, e_1e_3, e_2e_3, e_1e_2e_3$

Using the left regular representation over \mathbb{R} associated with this basis of Cl_n , we can realise Cl_n as a subalgebra of $M_{2^n}(\mathbb{R})$.

Example 5.3

For Cl_1 we have the basis $\{1, e_1\}$ and find that

$$\rho(1) = I_2, \quad \rho(e_1) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

while the general formula is

$$\rho(x + ye_1) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \quad (x, y \in \mathbb{R}).$$

Applying ρ to the basis $\{1, e_1, e_2, e_1e_2\}$ we obtain the following elements in $M_4(\mathbb{R})$:

$$\begin{aligned}\rho(1) &= I_4, & \rho(e_1) &= \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\ \rho(e_2) &= \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, & \rho(e_1e_2) &= \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.\end{aligned}$$

In all cases the matrices $\rho(e_{i_1}e_{i_2}\cdots e_{i_r})$ are *generalised permutation matrices* all of whose entries are $0, \pm 1$ and exactly one non-zero entry in each row and column. These are always orthogonal matrices of determinant 1.

The Clifford algebras Cl_n are characterised by an important *universal property*. First notice that there is an \mathbb{R} -linear transformation

$$j_n: \mathbb{R}^n \longrightarrow Cl_n; \quad j_n \left(\sum_{r=1}^n x_r e_r \right) = \sum_{r=1}^n x_r e_r,$$

for which

$$j_n \left(\sum_{r=1}^n x_r e_r \right)^2 = - \sum_{r=1}^n x_r^2 = - \left| \sum_{r=1}^n x_r e_r \right|^2. \quad (5.3)$$

Theorem 5.4 (Universal property of a Clifford algebra)

Let A be an \mathbb{R} -algebra. If $f: \mathbb{R}^n \longrightarrow A$ is an \mathbb{R} -linear transformation for which

$$f(x)^2 = -|x|^2 1 \quad (x \in \mathbb{R}^n),$$

then there is a unique homomorphism of \mathbb{R} -algebras $F: Cl_n \longrightarrow A$ for which $F \circ j_n = f$, i.e., for all $x \in \mathbb{R}^n$, $F(j_n(x)) = f(x)$.

This can be indicated in the following commutative diagram

$$\begin{array}{ccc} \mathbb{R}^n & & \\ j_n \swarrow & & \searrow f \\ Cl_n & \xrightarrow{\exists! F} & A \end{array}$$

in which ' $\exists!$ ' stands for 'there exists a unique F ' and the dotted arrow indicates a function F which solves the following equation amongst functions, $F \circ j_n = f$.

Proof

The homomorphism $F: \text{Cl}_n \rightarrow A$ is defined by setting $F(e_r) = f(e_r)$ and showing that it extends it to a ring homomorphism given on basis monomials by

$$F(e_{i_1} \cdots e_{i_r}) = f(e_{i_1}) \cdots f(e_{i_r}).$$

The details are left as an exercise. \square

Corollary 5.5

Let U be an \mathbb{R} -algebra and $j: \mathbb{R}^n \rightarrow U$ be an \mathbb{R} -linear transformation for which

$$j(x)^2 = -|x|^2 1 \quad (x \in \mathbb{R}^n).$$

Suppose that U and j have the universal property enjoyed by Cl_n .

If A is an \mathbb{R} -algebra and $h: \mathbb{R}^n \rightarrow A$ is an \mathbb{R} -linear transformation for which

$$h(x)^2 = -|x|^2 1 \quad (x \in \mathbb{R}^n),$$

then there is a unique homomorphism of \mathbb{R} -algebras $H: U \rightarrow A$ satisfying $H \circ j = h$, i.e., for all $x \in \mathbb{R}^n$, $H(j(x)) = h(x)$.

Then there is a unique \mathbb{R} -algebra isomorphism $\Phi: \text{Cl}_n \rightarrow U$ which satisfies $\Phi \circ j_n = j$.

Proof

The algebra homomorphism $\Phi: \text{Cl}_n \rightarrow U$ is constructed using the universal property Cl_n applied to the linear transformation $j: \mathbb{R}^n \rightarrow U$. On the other hand, applying the universal property of U to $j_n: \mathbb{R}^n \rightarrow \text{Cl}_n$, we obtain an algebra homomorphism $\Theta: U \rightarrow \text{Cl}_n$. These homomorphisms satisfy the equations

$$\Phi \circ j_n = j, \quad \Theta \circ j = j_n.$$

Notice that the following equations also hold:

$$\Theta \circ \Phi \circ j_n = \Theta \circ j = j_n, \quad \Phi \circ \Theta \circ j = \Phi \circ j_n = j.$$

By the uniqueness part of the universality of Cl_n applied to the linear transformation $j_n: \mathbb{R}^n \rightarrow \text{Cl}_n$ together with the identity $\text{Id}_{\text{Cl}_n} \circ j_n = j_n$, we have $\Theta \circ \Phi = \text{Id}_{\text{Cl}_n}$. Similarly, combining the uniqueness part of the universality of U

applied to the linear transformation $j: \mathbb{R}^n \rightarrow U$ with the identity $\text{Id}_U \circ j = j$, we obtain $\Phi \circ \Theta = \text{Id}_U$.

$$\begin{array}{ccc} & \mathbb{R}^n & \\ j_n \swarrow & \downarrow j & \searrow j_n \\ \text{Cl}_n & \xrightarrow{\exists! \Phi} & U \xrightarrow{\exists! \Theta} \text{Cl}_n \\ & \text{Id}_{\text{Cl}_n} & \\ & \searrow & \swarrow \\ & U & \xrightarrow{\exists! \Theta} \text{Cl}_n \xrightarrow{\exists! \Phi} U \end{array}$$

Hence Φ is an algebra isomorphism with inverse $\Phi^{-1} = \Theta$. \square

This uniqueness up to isomorphism is an important characteristic of objects satisfying such universal properties. Another example we discuss is that of the quotient topology, see Proposition 8.3 and Corollary 8.4.

Example 5.6

There is an \mathbb{R} -linear transformation

$$\alpha_0: \mathbb{R}^n \rightarrow \text{Cl}_n; \quad \alpha_0(\mathbf{x}) = -j_n(\mathbf{x}) = j_n(-\mathbf{x}).$$

Then in Cl_n ,

$$\alpha_0(\mathbf{x})^2 = j_n(-\mathbf{x})^2 = -|\mathbf{x}|^2,$$

so by Theorem 5.4 there is a unique algebra homomorphism $\alpha: \text{Cl}_n \rightarrow \text{Cl}_n$ for which

$$\alpha(j_n(\mathbf{x})) = \alpha_0(\mathbf{x}).$$

Since $j_n(\mathbf{e}_r) = \mathbf{e}_r$, this implies that

$$\alpha(\mathbf{e}_r) = -\mathbf{e}_r.$$

Notice that for $1 \leq i_1 < i_2 < \dots < i_k \leq n$,

$$\alpha(e_{i_1} e_{i_2} \cdots e_{i_k}) = (-1)^k e_{i_1} e_{i_2} \cdots e_{i_k} \begin{cases} e_{i_1} e_{i_2} \cdots e_{i_k} & \text{if } k \text{ is even,} \\ -e_{i_1} e_{i_2} \cdots e_{i_k} & \text{if } k \text{ is odd.} \end{cases}$$

It is easy to see that α is an isomorphism and hence an automorphism, often called the *canonical automorphism* of Cl_n .

We record the explicit forms of the next few Clifford algebras. Consider the \mathbb{R} -algebra $M_2(\mathbb{H})$ of dimension 16. There is an \mathbb{R} -linear transformation

$$\theta_4: \mathbb{R}^4 \rightarrow M_2(\mathbb{H});$$

$$\theta_4(x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3 + x_4 \mathbf{e}_4) = \begin{bmatrix} x_1 i + x_2 j + x_3 k & x_4 k \\ x_4 k & x_1 i + x_2 j - x_3 k \end{bmatrix}.$$

Direct calculation shows that θ_4 satisfies the condition of Theorem 5.4, hence there is a unique \mathbb{R} -algebra homomorphism $\Theta_4: \text{Cl}_4 \rightarrow M_2(\mathbb{H})$ for which $\Theta_4 \circ j_4 = \theta_4$. It can be shown that this is an isomorphism of \mathbb{R} -algebras, so

$$\text{Cl}_4 \cong M_2(\mathbb{H}).$$

Since $\mathbb{R} \subseteq \mathbb{R}^2 \subseteq \mathbb{R}^3 \subseteq \mathbb{R}^4$, we obtain compatible injective homomorphisms

$$\Theta_1: \text{Cl}_1 \rightarrow M_2(\mathbb{H}), \quad \Theta_2: \text{Cl}_2 \rightarrow M_2(\mathbb{H}), \quad \Theta_3: \text{Cl}_3 \rightarrow M_2(\mathbb{H}),$$

which turn out to have images

$$\begin{aligned}\text{im } \Theta_1 &= \{zI_2 : z \in \mathbb{C}\}, \\ \text{im } \Theta_2 &= \{qI_2 : q \in \mathbb{H}\}, \\ \text{im } \Theta_3 &= \left\{ \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} : q_1, q_2 \in \mathbb{H} \right\}.\end{aligned}$$

This shows that there is an isomorphism of \mathbb{R} -algebras

$$\text{Cl}_3 \cong \mathbb{H} \times \mathbb{H},$$

where the latter is the direct product of Definition 4.8. By direct calculation, we also obtain

$$\text{Cl}_5 \cong M_4(\mathbb{C}), \quad \text{Cl}_6 \cong M_8(\mathbb{R}), \quad \text{Cl}_7 \cong M_8(\mathbb{R}) \times M_8(\mathbb{R}).$$

These results are summarised in Table 5.1, whose last column gives their dimensions.

Cl_0	\mathbb{R}	1
Cl_1	\mathbb{C}	2
Cl_2	\mathbb{H}	4
Cl_3	$\mathbb{H} \times \mathbb{H}$	8
Cl_4	$M_2(\mathbb{H})$	16
Cl_5	$M_4(\mathbb{C})$	32
Cl_6	$M_8(\mathbb{R})$	64
Cl_7	$M_8(\mathbb{R}) \times M_8(\mathbb{R})$	128
Cl_8	$M_{16}(\mathbb{R})$	256

Table 5.1 The first 8 Clifford algebras

To go beyond this we use the following periodicity result, in which $M_m(\text{Cl}_n)$ denotes the ring of $m \times m$ matrices with entries in Cl_n .

Theorem 5.7

For $n \geq 0$,

$$\mathrm{Cl}_{n+8} \cong M_{16}(\mathrm{Cl}_n).$$

We will not give the proof but note that it involves the useful observation that for a unital ring R , there is an isomorphism of rings

$$M_m(M_n(R)) \cong M_{mn}(R), \quad (5.4)$$

obtained by expressing each $m \times m$ matrix of $n \times n$ matrices as an $mn \times mn$ matrix. Thus we obtain for example

$$\begin{aligned}\mathrm{Cl}_{10} &\cong M_{16}(\mathbb{H}), \\ \mathrm{Cl}_{12} &\cong M_{16}(M_2(\mathbb{H})) \cong M_{32}(\mathbb{H}), \\ \mathrm{Cl}_{14} &\cong M_{16}(M_8(\mathbb{R})) \cong M_{128}(\mathbb{R}).\end{aligned}$$

In the next section we will make use of some more structure in Cl_n . First there is a *conjugation* $(\overline{})$: $\mathrm{Cl}_n \rightarrow \mathrm{Cl}_n$ 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 $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, and satisfying

$$\begin{aligned}\overline{x+y} &= \overline{x} + \overline{y}, \\ \overline{txt} &= \overline{x},\end{aligned}$$

for $x, y \in \mathrm{Cl}_n$ and $t \in \mathbb{R}$. Notice that if $n > 1$, $(\overline{})$ is not a ring homomorphism since whenever $r < s$,

$$\overline{e_r e_s} = e_s e_r = -e_r e_s = -\overline{e_r e_s} \neq \overline{e_r} \overline{e_s}.$$

However, it is a ring *anti-homomorphism* in the sense that for all $x, y \in \mathrm{Cl}_n$,

$$\overline{xy} = \overline{y} \overline{x} \quad (x, y \in \mathrm{Cl}_n). \quad (5.5)$$

When $n = 1$ or 2 , $(\overline{})$ agrees with the conjugations already defined in \mathbf{C} and \mathbb{H} .

Second there is the canonical automorphism $\alpha: \mathrm{Cl}_n \rightarrow \mathrm{Cl}_n$ defined in Example 5.6. 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 5.8

- i) Every element $v \in Cl_n$ can be uniquely expressed as $v = v^+ + v^-$ with $v^+ \in Cl_n^+$ and $v^- \in Cl_n^-$. Hence $Cl_n = Cl_n^+ \oplus Cl_n^-$ as an \mathbb{R} -vector space.
ii) This decomposition is multiplicative in the sense that

$$\begin{cases} uv \in Cl_n^+ & \text{if } u, v \in Cl_n^+ \text{ or } u, v \in Cl_n^-, \\ uv, vu \in Cl_n^+ & \text{if } u \in Cl_n^+ \text{ and } v \in Cl_n^-. \end{cases}$$

Proof

- (i) The elements

$$v^+ = \frac{1}{2}(v + \alpha(v)), \quad v^- = \frac{1}{2}(v - \alpha(v)), \quad (5.6)$$

satisfy $\alpha(v^+) = v^+$, $\alpha(v^-) = -v^-$ and $v = v^+ + v^-$, hence $v^+ \in Cl_n^+$ and $v^- \in Cl_n^-$. On the other hand, if $v = v' + v''$ with $v' \in Cl_n^+$ and $v'' \in Cl_n^-$, then $\alpha(v) = v' - v''$ and so

$$\frac{1}{2}(v + \alpha(v)) = v', \quad \frac{1}{2}(v - \alpha(v)) = v''.$$

So this expression is the only one with such properties and therefore defines the stated vector space direct sum decomposition.

- (ii) This is easily checked using the fact that α is a ring homomorphism. \square

Notice that for bases of Cl_n^\pm we have the monomials

$$\begin{cases} e_{j_1} \cdots e_{j_{2m}} \in Cl_n^+ & (1 \leq j_1 < \cdots < j_{2m} \leq n), \\ e_{j_1} \cdots e_{j_{2m+1}} \in Cl_n^- & (1 \leq j_1 < \cdots < j_{2m+1} \leq n). \end{cases} \quad (5.7)$$

For later use, it is useful to record the following identity which holds for all $u \in Cl_n$:

$$\alpha(\bar{u}) = \overline{\alpha(u)}. \quad (5.8)$$

This is verified by checking that it holds on the monomial basis in the e_i .

Remark 5.9

It is worth noting that the composite of α and $\overline{(\cdot)}$ (in either order) is another anti-homomorphism of Cl_n ; then Id, α , $\overline{(\cdot)}$ and $\overline{\alpha} = \alpha \circ \overline{(\cdot)}$ form a finite group of order 4 which is not cyclic since the square of every element is the identity.

Finally, we introduce an inner product \cdot and a norm $\| \cdot \|$ on Cl_n by defining the distinct monomials $e_{i_1}e_{i_2}\cdots e_{i_k}$ with $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ 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

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

We leave it to the reader to check that $\| \cdot \|$ is a norm. Perhaps a more illuminating way to define the inner product \cdot is by using the formula

$$u \cdot v = \frac{1}{2} \operatorname{Re}(\bar{u}v + \bar{v}u), \quad (5.9)$$

where for $w \in \text{Cl}_n$ 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 $e_{i_1}\cdots e_{i_r}$. It is easily verified that for any $u, v \in \text{Cl}_n$ and $w \in j_n\mathbb{R}^n$,

$$(wu) \cdot (wv) = |w|^2(u \cdot v). \quad (5.10)$$

In particular, when $|w| = 1$ left multiplication by w defines an \mathbb{R} -linear transformation on Cl_n which is an isometry. The norm $\| \cdot \|$ gives rise to a metric on Cl_n^\times which makes the group of units Cl_n^\times into a topological group and the above embeddings of Cl_n into matrix rings are then continuous, so Cl_n^\times is a matrix group. Unfortunately, these embeddings are not norm preserving in general. For example, $2 + e_1e_2e_3 \in \text{Cl}_3$ has $|2 + e_1e_2e_3| = \sqrt{5}$, but the corresponding matrix in $M_8(\mathbb{R})$ has norm $\sqrt{3}$. However, by defining for each $w \in \text{Cl}_n$

$$\|w\| = \sup \{|wx| : x \in \text{Cl}_n, |x| = 1\} = \max \{|wx| : x \in \text{Cl}_n, |x| = 1\},$$

we obtain another equivalent norm on Cl_n for which the above embedding $\text{Cl}_n \rightarrow M_{2^n}(\mathbb{R})$ does preserve norms. For $w \in j_n\mathbb{R}^n$ we do have $\|w\| = |w|$ and more generally, for $w_1, \dots, w_k \in j_n\mathbb{R}^n$,

$$\|w_1 \cdots w_k\| = |w_1 \cdots w_k| = |w_1| \cdots |w_k|.$$

For $x, y \in \text{Cl}_n$,

$$\|xy\| \leq \|x\| \|y\|,$$

without equality in general.

5.2 Clifford Groups

Using the injective linear transformation $j_n : \mathbb{R}^n \rightarrow \text{Cl}_n$ we can identify \mathbb{R}^n with a subspace of Cl_n , i.e.,

$$\sum_{r=1}^n x_r e_r \longleftrightarrow j_n\left(\sum_{r=1}^n x_r e_r\right) = \sum_{r=1}^n x_r e_r.$$

From now on we do this without further comment, writing elements as $x \in \mathbb{R}^n$ rather than \mathbf{x} .

Notice that $\mathbb{R}^n \subseteq \text{Cl}_n^-$, so for $x \in \mathbb{R}^n$, $u \in \text{Cl}_n^+$ and $v \in \text{Cl}_n^-$,

$$xu, ux \in \text{Cl}_n^-, \quad xv, vx \in \text{Cl}_n^+. \quad (5.11)$$

Definition 5.10

For $n \geq 1$, the *Clifford group* Γ_n is the subgroup

$$\Gamma_n = \{u \in \text{Cl}_n^\times : \forall x \in \mathbb{R}^n, \alpha(u)xu^{-1} \in \mathbb{R}^n\} \leq \text{Cl}_n^\times.$$

Notice that $\mathbb{R}^\times \leq \Gamma_n$ is a normal subgroup and is the centre of Γ_n . Also for any non-zero $u \in \mathbb{R}^n$ and $x \in \mathbb{R}^n$ it is easily seen that $\alpha(u)xu^{-1} \in \mathbb{R}^n$, hence $u \in \Gamma_n$.

Proposition 5.11

Γ_n is a closed subgroup of Cl_n^\times .

Proof

There is a continuous action

$$\text{Cl}_n^\times \times \text{Cl}_n \longrightarrow \text{Cl}_n; \quad (u, v) = \alpha(u)vu^{-1}.$$

Notice that for each $u \in \text{Cl}_n^\times$, the function

$$\text{Cl}_n \longrightarrow \text{Cl}_n; \quad v \longmapsto \alpha(u)vu^{-1},$$

is a linear isomorphism. Since $\mathbb{R}^n \subseteq \text{Cl}_n$ is an \mathbb{R} -subspace, it is closed and so by Section 1.9, its stabiliser

$$\text{Stab}_{\text{Cl}_n^\times}(\mathbb{R}^n) \leq \text{Cl}_n^\times$$

is a closed subgroup of Cl_n^\times . But this stabiliser is clearly Γ_n . \square

Since we already know that Cl_n^X is a matrix group we have the following corollary.

Corollary 5.12

For $n \geq 1$, Γ_n is a matrix group.

Proposition 5.13

For $u \in \Gamma_n$, $\alpha(u)$ and \bar{u} are also in Γ_n .

Proof

For $u \in \Gamma_n$ and $x \in \mathbb{R}^n$, $\alpha(x) = -x = \bar{x} \in \mathbb{R}^n$. By Equations (5.8) and (5.5) together with the fact that $\alpha \circ \alpha = \text{Id}$, we have

$$\begin{aligned}\alpha(\alpha(u))x\alpha(u)^{-1} &= \alpha(\alpha(u))\alpha(-x)\alpha(u)^{-1} \\ &= \alpha(\alpha(u)(-x)u^{-1}) \in \mathbb{R}^n,\end{aligned}$$

and

$$\begin{aligned}\alpha(\bar{u})x\bar{u}^{-1} &= \alpha(\bar{u})(\overline{-x})\overline{u^{-1}} \\ &= \overline{\alpha(u)(-x)u^{-1}} \\ &= \overline{u^{-1}(-x)\alpha(u)} \\ &= \overline{\alpha(\alpha(u^{-1})\alpha(-x)u)} \in \mathbb{R}^n,\end{aligned}$$

since \mathbb{R}^n is closed under α , $(\bar{})$ and $\rho_{\alpha(u^{-1})}$. □

For each $u \in \text{Cl}_n^X$, we can restrict the function $\text{Cl}_n \rightarrow \text{Cl}_n$ of this proof to

$$\rho_u: \mathbb{R}^n \rightarrow \mathbb{R}^n; \quad \rho_u(v) = \alpha(u)vu^{-1}.$$

Proposition 5.14

For each $u \in \Gamma_n$, the function $\rho_u: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an \mathbb{R} -linear isometry.

Proof

For any $x \in \mathbb{R}^n$,

$$\begin{aligned} |\rho_u(x)|^2 &= -\rho_u(x)^2 \\ &= -((-u)xu^{-1})^2 \\ &= -(-u)xu^{-1}(-u)xu^{-1} \\ &= -ux^2u^{-1} \\ &= u|x|^2u^{-1} \\ &= |x|^2. \end{aligned}$$

Hence $|\rho_u(x)| = |x|$ for all $x \in \mathbb{R}^n$. \square

This means that for $u \in \Gamma_n$, if we express ρ_u in terms of the standard basis $\{e_1, \dots, e_n\}$, then we have $\rho_u \in O(n)$. Hence there is a group homomorphism

$$\rho: \Gamma_n \longrightarrow O(n); \quad \rho(u) = \rho_u.$$

In fact, ρ is also continuous. We also need to identify the kernel of ρ .

Proposition 5.15

$$\ker \rho = \mathbb{R}^\times = \{t1 : t \in \mathbb{R}, t \neq 0\}.$$

Proof

Suppose that $u \in \ker \rho$. Then for every $x \in \mathbb{R}^n$, $\alpha(u)xu^{-1} = x$, i.e., $\alpha(u)x = xu$. Writing $u = u^+ + u^-$ with $u^\pm \in Cl_n^\pm$, we have

$$u^+x = xu^+, \tag{5.12a}$$

$$-u^-x = xu^-. \tag{5.12b}$$

For each $r = 1, \dots, n$, we can write

$$u^+ = a_r^+ + e_r b_r^-, \quad u^- = a_r^- + e_r b_r^+,$$

where $a_r^\pm, b_r^\pm \in Cl_n^\pm$ do not involve e_r in their expansions in terms of the monomial bases of Equation (5.7).

On taking $x = e_r$, by Equation (5.12a) together with the identities

$$e_r a_r^+ = a_r^+ e_r, \quad e_r b_r^- = -b_r^- e_r,$$

we obtain

$$a_r^+ e_r + b_r^- = a_r^+ e_r - b_r^-.$$

Now by comparing the parts involving or not involving e_r we find that $b_r^- = 0$. Similarly, Equation (5.12b) gives

$$-a_r^- e_r + b_r^+ = -a_r^- e_r - b_r^+,$$

from which we obtain $b_r^+ = 0$. So we have $u^+ = a_r^+$ and $u^- = a_r^-$, showing that neither of these involves e_r . But as this is true for all values of r , we must have $u^- = 0$ and $u^+ = t1$ for some $t \in \mathbb{R}$, hence $u = t1$. \square

Proposition 5.16

For $u \in \Gamma_n$, $u\bar{u} \in \mathbb{R}^\times$ and $\bar{u}u = u\bar{u}$. If $v \in \Gamma_n$ as well, then

$$uv\bar{u}\bar{v} = u\bar{u}v\bar{v}.$$

Proof

For $u \in \Gamma_n$, $u\bar{u} \in \Gamma_n$ by Proposition 5.13. We will show that $u\bar{u} \in \ker \rho$.

Let $x \in \mathbb{R}^n$. Then using the notation $\bar{\alpha} = \alpha \circ (\bar{}) = (\bar{}) \circ \alpha$ of Remark 5.9 we have

$$\begin{aligned} \rho_{u\bar{u}}(x) &= \alpha(u\bar{u})x(u\bar{u})^{-1} \\ &= \alpha(u)\alpha(\bar{u})x\bar{u}^{-1}u^{-1} \\ &= \alpha(u)\left(\alpha(\bar{u})x\bar{u}^{-1}\right)u^{-1} \\ &= \alpha(u)\alpha\left(\bar{u}\bar{x}\alpha(u^{-1})\right)u^{-1}, \end{aligned}$$

since $\alpha(\bar{x}) = x$. Hence

$$\begin{aligned} \rho_{u\bar{u}}(x) &= \alpha(u)\alpha\left(\overline{\alpha(u^{-1})xu}\right)u^{-1} \\ &= \alpha(u)\alpha(-\alpha(u^{-1})xu)u^{-1} \\ &\quad (\text{since } \bar{y} = -y \text{ for } y \in \mathbb{R}^n) \\ &= \alpha(u)(\alpha(u^{-1})xu)u^{-1} \\ &\quad (\text{since } \alpha(y) = -y \text{ for } y \in \mathbb{R}^n) \\ &= \alpha(uu^{-1})x(uu^{-1}) = x. \end{aligned}$$

Thus $\rho_{u\bar{u}}(x) = x$. We also have

$$\bar{u}u = u^{-1}u\bar{u}u = u\bar{u}u^{-1}u$$

since $u\bar{u} \in \mathbb{R}^\times$, hence $\bar{u}u = u\bar{u}$.

If $v \in \Gamma_n$, then

$$uv\bar{w}\bar{v} = uv\bar{v}\bar{u} = u\bar{w}v\bar{v},$$

since $v\bar{v} \in \mathbb{R}^\times$ and so commutes with every element of Cl_n . \square

Proposition 5.17

The function

$$\nu: \Gamma_n \longrightarrow \mathbb{R}^\times; \quad \nu(u) = u\bar{u}$$

is a continuous group homomorphism.

Proof

Notice that if $u \in \Gamma_n$,

$$|u|^2 = \text{Re}(\bar{u}u) = \text{Re}(u\bar{u}) = u\bar{u} = \nu(u),$$

hence ν is continuous and always takes *positive* values. Proposition 5.16 implies that it is a group homomorphism. \square

We will study the kernel of ν more thoroughly in the next section.

5.3 Pinor and Spinor Groups

In this section we will describe for each $n \geq 1$, the compact connected *spinor group* $\text{Spin}(n)$ which is a group of units in the Clifford algebra Cl_n . Moreover, there is a surjective Lie homomorphism $\text{Spin}(n) \rightarrow \text{SO}(n)$ whose kernel has two elements. We will first introduce the *pinor group* $\text{Pin}(n)$ which has two connected components, one of them being $\text{Spin}(n)$.

Definition 5.18

For $n \geq 1$, the *pinor group* $\text{Pin}(n)$ is the kernel of $\nu: \Gamma_n \rightarrow \mathbb{R}^\times$.

Notice that $\text{Pin}(n)$ is a closed subgroup of $\Gamma_n \subseteq \text{Cl}_n$ and is a bounded subset with respect to the metric induced from the norm $| |$, hence it is compact.

Definition 5.19

For $n \geq 1$, the *spinor group* $\text{Spin}(n)$ is

$$\text{Spin}(n) = \text{Pin}(n) \cap \text{Cl}_n^+ \leq \text{Pin}(n).$$

The restriction of α to a function $\alpha: \text{Pin}(n) \rightarrow \text{Pin}(n)$ is a continuous group homomorphism for which

$$\{u \in \text{Pin}(n) : \alpha(u) = u\} = \text{Pin}(n) \cap \text{Cl}_n^+,$$

hence $\text{Spin}(n) \leq \text{Pin}(n)$ is a closed subgroup. Later we will see that it is also a normal subgroup.

Our principal goal in this section is to show that the restricted homomorphism $\rho: \text{Pin}(n) \rightarrow \text{O}(n)$ is surjective and $\text{Spin}(n) = \rho^{-1}(\text{SO}(n))$. We will do this by showing that $\text{Pin}(n)$ is generated by certain elements $u \in \mathbb{R}^n$ for which ρ_u is a reflection.

Within $\mathbb{R}^n \subseteq \text{Cl}_n$ is the *unit sphere*

$$\mathbf{S}^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| = 1\} = \left\{ \sum_{r=1}^n x_r e_r : \sum_{r=1}^n x_r^2 = 1 \right\}.$$

Lemma 5.20

Let $u \in \mathbf{S}^{n-1} \subseteq \text{Cl}_n$. Then u is a unit in Cl_n , i.e., $u \in \text{Cl}_n^\times$, and $u^{-1} \in \mathbf{S}^{n-1}$.

Proof

Since $u \in \mathbb{R}^n$,

$$(-u)u = u(-u) = -u^2 = -(-|u|^2) = 1,$$

so $-u$ is the inverse of u . Of course $-u \in \mathbf{S}^{n-1}$. □

More generally, for $u_1, \dots, u_k \in \mathbf{S}^{n-1}$ we have

$$(u_1 \cdots u_k)^{-1} = (-1)^k u_k \cdots u_1 = \overline{u_1 \cdots u_k}. \quad (5.13)$$

Recall that in a group G , the subgroup *generated* by the subset $S \subseteq G$ is the smallest subgroup of G containing S , often denoted $\langle S \rangle \leq G$. A typical element of $\langle S \rangle$ is a product of elements of the form s or s^{-1} where $s \in S$, i.e.,

$$s_1^{\epsilon_1} \cdots s_k^{\epsilon_k} \quad (\epsilon_1, \dots, \epsilon_k = \pm 1).$$

Furthermore, if $H \leq G$ and $S \subseteq H$ then $\langle S \rangle \leq H$.

We will show that $\mathbf{S}^{n-1} \subseteq \text{Pin}(n)$ and therefore $\langle \mathbf{S}^{n-1} \rangle \leq \text{Pin}(n)$. We begin with the following useful result.

Lemma 5.21

Let $u, v \in \mathbb{R}^n \subseteq \text{Cl}_n$. 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

$$\begin{aligned} vu &= \sum_{s=1}^n \sum_{r=1}^n y_s x_r e_s e_r = \sum_{r=1}^n y_r x_r e_r^2 + \sum_{r < s} (x_s y_r - x_r y_s) e_r e_s \\ &= - \sum_{r=1}^n y_r x_r - \sum_{r < s} (x_r y_s - x_s y_r) e_r e_s \\ &= -u \cdot v - \sum_{r < s} (x_r y_s - x_s y_r) e_r e_s \\ &= - \sum_{r < s} (x_r y_s - x_s y_r) e_r e_s \\ &= v \cdot u - \sum_{r < s} (x_r y_s - x_s y_r) e_r e_s \\ &= - \sum_{r=1}^n \sum_{s=1}^n x_r y_s e_r e_s \\ &= -uv. \end{aligned}$$

□

Notice that for $u \in \mathbf{S}^{n-1}$ and $x \in \mathbb{R}^n$, by Equation (5.13),

$$\alpha(u)xu^{-1} = (-u)x(-u) = uxu. \quad (5.14)$$

If $u \cdot x = 0$, then Lemma 5.21 implies that

$$\alpha(u)xu^{-1} = -u^2x = -(-1)x = x, \quad (5.15a)$$

since $u^2 = -|u|^2 = -1$. On the other hand, if $x = tu$ for some $t \in \mathbb{R}$, then

$$\alpha(u)xu^{-1} = tu^3 = -tu. \quad (5.15b)$$

So in particular $\alpha(u)x^{-1} \in \mathbb{R}^n$ which yields the following result.

Proposition 5.22

The subgroup of Cl_n^\times generated by \mathbf{S}^{n-1} is contained in $\text{Pin}(n)$, $\langle \mathbf{S}^{n-1} \rangle \leq \text{Pin}(n)$.

Remark 5.23

By Lemma 5.20, every element of $\langle S^{n-1} \rangle$ is a product of elements of S^{n-1} .

Next we will investigate the restriction of the homomorphism $\rho: \text{Pin}(n) \rightarrow O(n)$ to $\rho: \langle S^{n-1} \rangle \rightarrow O(n)$. For each $u \in S^{n-1}$, we have an \mathbb{R} -linear isometry

$$\rho_u: \mathbb{R}^n \rightarrow \mathbb{R}^n; \quad \rho_u(x) = \alpha(u)xu^{-1} = uxu.$$

More generally, for $u \in \langle S^{n-1} \rangle$, suppose that $u = u_1 \cdots u_r$ for $u_1, \dots, u_r \in S^{n-1}$. Then

$$\begin{aligned} \rho_u(x) &= \alpha(u_1 \cdots u_r)x(u_1 \cdots u_r)^{-1} \\ &= ((-1)^r u_1 \cdots u_r)x((-1)^r u_r \cdots u_1) \\ &= u_1 \cdots u_r xu_r \cdots u_1. \end{aligned} \tag{5.16}$$

Proposition 5.24

For $u \in S^{n-1}$, $\rho_u: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is reflection in the hyperplane orthogonal to u .

Proof

Recalling the defining equation (1.4) of a hyperplane reflection, Equations (5.15) show that

$$\rho_u(x) = \begin{cases} x & \text{if } u \cdot x = 0, \\ -x & \text{if } x = tu \text{ for some } t \in \mathbb{R}. \end{cases}$$

So $\rho_u: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is indeed reflection in the hyperplane orthogonal to u . \square

Proposition 5.25

- i) $\rho: \langle S^{n-1} \rangle \rightarrow O(n)$ is surjective and $\ker \rho = \{1, -1\}$.
- ii) $\rho: \text{Pin}(n) \rightarrow O(n)$ is surjective with $\ker \rho = \{1, -1\}$.
- iii) $\langle S^{n-1} \rangle = \text{Pin}(n)$.

Proof

- (i) The observation in the proof of Proposition 5.24 shows that reflection in the hyperplane orthogonal to $u \in S^{n-1}$ has the form ρ_u . Surjectivity is a consequence of Proposition 1.41.
- (ii) If $t \in \ker \rho = \text{Pin}(n) \cap \mathbb{R}^\times$, then $|t| = 1$ then $|t|^2 = t\bar{t} = t^2$, so $t = \pm 1$. On the other hand, $\pm 1 \in \ker \rho$.

(iii) Suppose that $v \in \text{Pin}(n)$. Then $\rho_v \in O(n)$ can be expressed as a product of hyperplane reflections, each of which has the form ρ_w for some vector $w \in S^{n-1}$. Hence $\rho_v = \rho_{w_1 \cdots w_k}$ for suitable $w_1, \dots, w_k \in S^{n-1}$. But then

$$(-1)^k v w_k \cdots w_1 \in \ker \rho = \{\pm 1\}.$$

Thus $v = \pm w_1 \cdots w_k \in \langle S^{n-1} \rangle$. □

At this point we have a great deal of information about elements of $\text{Pin}(n)$ which we collect into a theorem.

Theorem 5.26

i) $\text{Pin}(n)$ is the disjoint union of open subsets

$$\text{Pin}(n) = \text{Pin}(n) \cap Cl_n^+ \cup \text{Pin}(n) \cap Cl_n^- = \text{Spin}(n) \cup \text{Pin}(n) \cap Cl_n^-.$$

ii) $\text{Spin}(n)$ is a normal subgroup of $\text{Pin}(n)$ and every element $v \in \text{Spin}(n)$ can be expressed as an even-length product $v = v_1 \cdots v_{2\ell}$ where $v_1, \dots, v_{2\ell} \in S^{n-1}$.

iii) For any $w \in S^{n-1}$,

$$\text{Pin}(n) \cap Cl_n^- = w \text{Spin}(n),$$

and every element $v \in \text{Pin}(n) \cap Cl_n^-$ can be expressed as an odd-length product $v = v_1 \cdots v_{2\ell+1}$ where $v_1, \dots, v_{2\ell+1} \in S^{n-1}$.

Proof

(i) Since $\text{Pin}(n) = \langle S^{n-1} \rangle$, every element $u \in \text{Pin}(n)$ can be expressed as $u = u_1 \cdots u_k$ for some $u_1, \dots, u_k \in S^{n-1}$. Since

$$\alpha(u_1 \cdots u_k) = (-1)^k u_1 \cdots u_k,$$

we find that $u \in \text{Pin}(n) \cap Cl_n^+$ if k is even, while $u \in \text{Pin}(n) \cap Cl_n^-$ if k is odd. The function

$$\tilde{\alpha}: \text{Pin}(n) \longrightarrow \{+1, -1\}; \quad \tilde{\alpha}(u) = \alpha(u)u^{-1},$$

is continuous and

$$\tilde{\alpha}^{-1}\{1\} = \text{Pin}(n) \cap Cl_n^+, \quad \tilde{\alpha}^{-1}\{-1\} = \text{Pin}(n) \cap Cl_n^-,$$

showing that these are disjoint open subsets.

(ii) If $u \in \text{Spin}(n) = \text{Pin}(n) \cap Cl_n^+$ and $v \in \text{Pin}(n)$, then writing $u = u_1 \cdots u_{2k}$ and $v = v_1 \cdots v_{2\ell+1}$ with $u_i, v_j \in S^{n-1}$ we find

$$vuv^{-1} = -v_1 \cdots v_{2\ell+1} u_1 \cdots u_{2k} v_{2\ell+1} \cdots v_1 \in \text{Pin}(n) \cap Cl_n^+.$$

(iii) If $u \in \text{Pin}(n) \cap \text{Cl}_n^-$ then for each $w \in \mathbb{S}^{n-1}$ we have

$$w^{-1}u = (-w)u \in \text{Pin}(n) \cap \text{Cl}_n^+.$$

□

From now on we will make use of the notation

$$\text{Pin}(n)^+ = \text{Spin}(n) = \text{Pin}(n) \cap \text{Cl}_n^+, \quad \text{Pin}(n)^- = \text{Pin}(n) \cap \text{Cl}_n^-.$$

We can also characterise $\text{Spin}(n)$ and $\text{Pin}(n)^-$ using the surjective homomorphism $\rho: \text{Pin}(n) \rightarrow \text{O}(n)$. If $u_1, \dots, u_k \in \mathbb{S}^{n-1}$, then

$$\det \rho_{u_1 \dots u_k} = \det \rho_{u_1} \cdots \det \rho_{u_k} = (-1)^k,$$

since each ρ_{u_r} is a hyperplane reflection for which $\det \rho_{u_r} = -1$. Since

$$\text{SO}(n) = \text{O}(n)^+ = \{A \in \text{O}(n) : \det A = 1\},$$

$$\text{O}(n)^- = \{A \in \text{O}(n) : \det A = -1\},$$

this means that for $u \in \text{Pin}(n)$,

$$u \in \text{Spin}(n) \iff \rho_u \in \text{SO}(n),$$

$$u \in \text{Pin}(n)^- \iff \rho_u \in \text{O}(n)^-.$$

Theorem 5.27

The continuous homomorphism $\rho: \text{Pin}(n) \rightarrow \text{O}(n)$ is surjective and

$$\rho^{-1}\text{SO}(n) = \text{Spin}(n), \quad \rho^{-1}\text{O}(n)^- = \text{Pin}(n)^-.$$

Hence the restriction of ρ to $\rho^+: \text{Spin}(n) \rightarrow \text{SO}(n)$ is also surjective and the kernels of these homomorphisms are $\ker \rho = \rho^+ = \{+1, -1\}$.

We can also say more about the topology of $\text{Spin}(n)$ and $\text{Pin}(n)$.

Theorem 5.28

$\text{Spin}(n)$ is a compact, path connected, closed normal subgroup of $\text{Pin}(n)$. Furthermore, if $n \geq 3$ the fundamental group of $\text{Spin}(n)$ is trivial, $\pi_1 \text{Spin}(n) = \{1\}$.

Proof

We only discuss connectivity. Recall that the sphere $S^{n-1} \subseteq \mathbb{R}^n \subseteq Cl_n$ is path connected. Choose a base point $u_0 \in S^{n-1}$. Now for an element $u = u_1 \cdots u_k \in \text{Spin}(n)$ with $u_1, \dots, u_k \in S^{n-1}$, as noted in the proof of Proposition 5.25, 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 $w \in S^{n-1}$. Then there are continuous paths

$$p_r: [0, 1] \longrightarrow S^{n-1} \quad (r = 1, \dots, 2m),$$

for which $p_r(0) = u_0$ and $p_r(1) = u_r$. Then

$$p: [0, 1] \longrightarrow S^{n-1}; \quad p(t) = p_1(t) \cdots p_{2m}(t)$$

is a continuous path in $\text{Pin}(n)$ with

$$p(0) = u_0^{2m} = (-1)^m = 1, \quad p(1) = u.$$

But $t \mapsto \rho(p(t))$ is a continuous path in $O(n)$ with $\rho(p(0)) \in SO(n)$, hence $\rho(p(t)) \in SO(n)$ for all t . This shows that p is a path in $\text{Spin}(n)$. So every element $u \in \text{Spin}(n)$ can be connected to 1 and therefore $\text{Spin}(n)$ is path connected.

The final statement involves homotopy theory and is not proved here. It should be compared with the fact that for $n \geq 3$, $\pi_1 SO(n) \cong \{1, -1\}$ and in fact the map is an example of a *universal covering*. \square

The double covering maps $\rho: \text{Spin}(n) \longrightarrow SO(n)$ generalise the case of $SU(2) \longrightarrow SO(3)$ discussed in Section 3.5. In fact, around each element $u \in \text{Spin}(n)$ there is an open neighbourhood $N_u \subseteq \text{Spin}(n)$ for which $\rho: N_u \longrightarrow \rho N_u$ is a homeomorphism, and actually a diffeomorphism. This implies the following.

Proposition 5.29

The derivative $d\rho: \mathfrak{spin}(n) \longrightarrow \mathfrak{so}(n)$ is an isomorphism of \mathbb{R} -Lie algebras and

$$\dim \text{Spin}(n) = \dim SO(n) = \binom{n}{2}.$$

The Lie algebra of $\text{Spin}(n)$ can be described as a Lie subalgebra of Cl_n .

Proposition 5.30

For $n \geq 2$, the Lie algebra of $\text{Spin}(n)$ is

$$\mathfrak{spin}(n) = \left\{ \sum_{1 \leq i < j \leq n} t_{ij} e_i e_j : t_{ij} \in \mathbb{R} \right\} \subseteq Cl_n.$$

Proof

For $1 \leq i < j \leq n$ and $t \in \mathbb{R}$, consider

$$\exp(te_i e_j) = \sum_{r=0}^{\infty} \frac{t^r}{r!} (e_i e_j)^r,$$

which converges to an element of Cl_n^{\times} . Notice that

$$(e_i e_j)^r = (-1)^{\binom{r}{2}} e_i^r e_j^r = \begin{cases} 1 & \text{if } r \equiv 0 \pmod{4}, \\ -1 & \text{if } r \equiv 2 \pmod{4}, \\ e_i e_j & \text{if } r \equiv 1 \pmod{4}, \\ -e_i e_j & \text{if } r \equiv 3 \pmod{4}. \end{cases}$$

Hence we have

$$\exp(te_i e_j) = \cos t + \sin t e_i e_j. \quad (5.17)$$

We also obtain

$$\begin{aligned} e_i (e_i e_j)^r &= (-e_i e_j)^r e_i, \\ e_j (e_i e_j)^r &= (-e_i e_j)^r e_j, \end{aligned}$$

and if $k \neq i, j$,

$$e_k (e_i e_j)^r = (e_i e_j)^r e_k.$$

Then since $(e_i e_j)^2 = -1$,

$$\begin{aligned} \alpha(\exp(te_i e_j)) e_i \exp(te_i e_j)^{-1} &= \exp(te_i e_j) e_i \exp(-te_i e_j) \\ &= (\cos t + \sin t e_i e_j) e_i (\cos t - \sin t e_i e_j) \\ &= (\cos t + \sin t e_i e_j)^2 e_i \\ &= (\cos 2t + \sin 2t e_i e_j) e_i \\ &= \cos 2t e_i + \sin 2t e_j, \\ \alpha(\exp(te_i e_j)) e_j \exp(te_i e_j)^{-1} &= (\cos t + \sin t e_i e_j) e_j (\cos t - \sin t e_i e_j) \\ &= (\cos t + \sin t e_i e_j)^2 e_j \\ &= (\cos 2t + \sin 2t e_i e_j) e_j \\ &= \cos 2t e_j - \sin 2t e_i, \end{aligned}$$

and if $k \neq i, j$,

$$\begin{aligned} \alpha(\exp(te_i e_j)) e_k \exp(te_i e_j)^{-1} &= \exp(te_i e_j) \exp(te_i e_j)^{-1} e_k \\ &= e_k. \end{aligned}$$

Since the e_i form a basis of \mathbb{R}^n , this shows that $\exp(te_i e_j) \in \Gamma_n \cap \text{Cl}_n^+$. But also

$$\begin{aligned}\exp(te_i e_j) \overline{\exp(te_i e_j)} &= \exp(te_i e_j) \exp((-1)^2 te_j e_i) \\ &= \exp(te_i e_j) \exp(-te_i e_j) \\ &= 1,\end{aligned}$$

hence $\exp(te_i e_j) \in \text{Spin}(n)$.

Taking the derivative at 0 of the curve

$$\alpha_{ij}: \mathbb{R} \longrightarrow \text{Spin}(n); \quad \alpha_{ij}(t) = \exp(te_i e_j),$$

we obtain $e_i e_j$, so $e_i e_j \in \text{spin}(n)$. As these $\binom{n}{2}$ elements $e_i e_j$ are clearly linearly independent, by Proposition 5.29 they must form a basis. \square

With the aid of the calculations in this proof, the action of the derivative $d\rho: \text{spin}(n) \longrightarrow \mathfrak{so}(n)$ on the basis elements $e_i e_j$ can be determined, namely

$$d\rho(e_i e_j) = 2E^{ji} - 2E^{ij}, \quad (5.18a)$$

where we use the basis for $M_n(\mathbb{R})$ provided by Equation (3.24). Similarly, the effect of the homomorphism $\rho: \text{Spin}(n) \longrightarrow \text{SO}(n)$ on $\exp(te_i e_j)$ is

$$\rho(\exp(te_i e_j)) = \cos 2t I_n + \sin 2t (E^{ji} - E^{ij}). \quad (5.18b)$$

5.4 The Centres of Spinor Groups

Recall that for a group G the *centre* of G is

$$Z(G) = \{c \in G : \forall g \in G, gc = cg\}.$$

Then $Z(G) \triangleleft G$. The centres of the special orthogonal groups are easily found and described.

Proposition 5.31

For $n \geq 3$,

$$Z(\text{SO}(n)) = \{tI_n : t = \pm 1, t^n = 1\} = \begin{cases} \{I_n\} & \text{if } n \text{ is odd,} \\ \{\pm I_n\} & \text{if } n \text{ is even.} \end{cases}$$

Before stating our main result on the centres of spinor groups, we note that $\text{Spin}(1)$ and $\text{Spin}(2)$ are both abelian.

Proposition 5.32

For $n \geq 3$,

$$\begin{aligned} Z(\text{Spin}(n)) &= \begin{cases} \{\pm 1\} & \text{if } n \text{ is odd,} \\ \{\pm 1, \pm e_1 \cdots e_n\} & \text{if } n \equiv 2 \pmod{4}, \\ \{\pm 1, \pm e_1 \cdots e_n\} & \text{if } n \equiv 0 \pmod{4}. \end{cases} \\ &\cong \begin{cases} \mathbb{Z}/2 & \text{if } n \text{ is odd,} \\ \mathbb{Z}/4 & \text{if } n \equiv 2 \pmod{4}, \\ \mathbb{Z}/2 \times \mathbb{Z}/2 & \text{if } n \equiv 0 \pmod{4}. \end{cases} \end{aligned}$$

Proof

If $g \in Z(\text{Spin}(n))$, then since $\rho: \text{Spin}(n) \rightarrow \text{SO}(n)$, $\rho(g) \in Z(\text{SO}(n))$. As $\pm 1 \in Z(\text{Spin}(n))$, this gives $|Z(\text{Spin}(n))| = 2|Z(\text{SO}(n))|$ and indeed

$$Z(\text{Spin}(n)) = \rho^{-1} Z(\text{SO}(n)).$$

For n even,

$$\begin{aligned} (\pm e_1 \cdots e_n)^2 &= e_1 \cdots e_n e_1 \cdots e_n \\ &= (-1)^{\binom{n}{2}} e_1^2 \cdots e_n^2 \\ &= (-1)^{\binom{n}{2}+n} = (-1)^{\binom{n+1}{2}}. \end{aligned}$$

Since

$$\binom{n+1}{2} = \frac{(n+1)n}{2} \equiv \begin{cases} 1 \pmod{2} & \text{if } n \equiv 2 \pmod{4}, \\ 0 \pmod{2} & \text{if } n \equiv 0 \pmod{4}, \end{cases}$$

this implies that

$$(\pm e_1 \cdots e_n)^2 = \begin{cases} -1 & \text{if } n \equiv 2 \pmod{4}, \\ 1 & \text{if } n \equiv 0 \pmod{4}. \end{cases}$$

So for n even, the multiplicative order of $\pm e_1 \cdots e_n$ is 2 or 4 depending on the congruence class of n modulo 4. This gives the stated groups. \square

5.5 Finite Subgroups of Spinor Groups

Each orthogonal group $O(n)$ and $\text{SO}(n)$ contains finite subgroups. For example, when $n = 2, 3$, these correspond to symmetry groups of compact plane figures

and solids. The case of $n = 3$ was explored in the exercises of Chapter 4. Here we make some remarks about the symmetric and alternating groups.

Recall that for each $n \geq 1$ the *symmetric group* S_n is the group of all permutations of the set $\mathbf{n} = \{1, \dots, n\}$. The corresponding *alternating group* $A_n \leq S_n$ is the subgroup consisting of all *even* permutations, i.e., the elements $\sigma \in S_n$ for which $\text{sgn}(\sigma) = 1$ where $\text{sgn}: S_n \rightarrow \{\pm 1\}$ is the sign homomorphism.

For a field \mathbf{k} , we can make S_n act on \mathbf{k}^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 $\mathbf{k} = \mathbb{R}$ each of these matrices is orthogonal, while when $\mathbf{k} = \mathbb{C}$ it is unitary. For a given n we can view S_n as the subgroup of $O(n)$ or $U(n)$ consisting of all such matrices which are usually called *permutation matrices*.

Proposition 5.33

For each $\sigma \in S_n$ we have $\text{sgn}(\sigma) = \det([\sigma])$. Hence

$$A_n = \begin{cases} SO(n) \cap S_n & \text{if } \mathbf{k} = \mathbb{R}, \\ SU(n) \cap S_n & \text{if } \mathbf{k} = \mathbb{C}. \end{cases}$$

Recall that if $n \geq 5$, A_n is a simple group.

As $\rho: \text{Pin}(n) \rightarrow O(n)$ is onto, there are finite subgroups $\tilde{S}_n = \rho^{-1}S_n \leq \text{Pin}(n)$ and $\tilde{A}_n = \rho^{-1}A_n \leq \text{Spin}(n)$ for which there are surjective homomorphisms $\rho: \tilde{S}_n \rightarrow S_n$ and $\rho: \tilde{A}_n \rightarrow A_n$ whose kernels contain the elements ± 1 . Note that $|\tilde{S}_n| = 2 \cdot n!$, while $|\tilde{A}_n| = n!$. However, for $n \geq 4$, there are no homomorphisms $\tau: S_n \rightarrow \tilde{S}_n$, $\tau: A_n \rightarrow \tilde{A}_n$ for which $\rho \circ \tau = \text{Id}$.

$$\begin{array}{ccc} S_n & \xrightarrow{\tau} & \tilde{S}_n \\ & \searrow \text{Id}_{S_n} & \downarrow \rho \\ & & S_n \end{array} \qquad \begin{array}{ccc} A_n & \xrightarrow{\tau} & \tilde{A}_n \\ & \searrow \text{Id}_{A_n} & \downarrow \rho \\ & & A_n \end{array}$$

Similar considerations apply to other finite subgroups of $O(n)$.

In Cl_n^\times we have a subgroup E_n consisting of all the elements

$$\pm e_{i_1} \cdots e_{i_r} \quad (1 \leq i_1 < \cdots < i_r \leq n, 0 \leq r).$$

The order of this group is $|E_n| = 2^{n+1}$ and as it contains ± 1 , its image under $\rho: \text{Pin}(n) \rightarrow O(n)$ is $\bar{E}_n = \rho E_n$ of order $|\bar{E}_n| = 2^n$. In fact, $\{\pm 1\} = Z(E_n)$ is also the commutator subgroup since $e_i e_j e_i^{-1} e_j^{-1} = -1$ and so \bar{E}_n is abelian. Every non-trivial element in \bar{E}_n has order 2 since $e_i^2 = -1$, hence $\bar{E}_n \leq O(n)$ is an *elementary 2-group*, i.e., it is isomorphic to $(\mathbb{Z}/2)^n$. Each element $\rho(e_r) \in O(n)$ is a generalised permutation matrix with all its non-zero entries on the main diagonal. There is also a subgroup $\bar{E}_n^0 = \rho E_n^0 \leq SO(n)$ of order 2^{n-1} , where

$$E_n^0 = E_n \cap \text{Spin}(n).$$

In fact \bar{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

$$1 \rightarrow \mathbb{Z}/2 \rightarrow E_n \rightarrow (\mathbb{Z}/2)^n \rightarrow 1, \quad 1 \rightarrow \mathbb{Z}/2 \rightarrow E_n^0 \rightarrow (\mathbb{Z}/2)^{n-1} \rightarrow 1,$$

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 examples of *extraspecial 2-groups*.

EXERCISES

5.1. In the Clifford algebra Cl_n , let $u, v \in \mathbb{R}^n \subseteq \text{Cl}_n$.

- a) 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 uvv .
- b) Let $\{u_1, \dots, u_n\}$ be an 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 if i, j, k, ℓ are distinct numbers in the range 1 to n then

$$u_i u_j u_k u_\ell = u_k u_\ell u_i u_j.$$

- c) Show that each element $A \in O(n)$ induces an automorphism $A_*: \text{Cl}_n \rightarrow \text{Cl}_n$ for which $A_*x = Ax$ if $x \in \mathbb{R}^n$. Determine $A_*(e_1 \cdots e_n)$.

5.2. Use the Universal Property of Theorem 5.4 in the following.

a) Show that the natural embedding

$$i_n: \mathbb{R}^n \longrightarrow \mathbb{R}^{n+1}; \quad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \longmapsto \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 0 \end{bmatrix}$$

induces an \mathbb{R} -algebra homomorphism $i'_n: \text{Cl}_n \longrightarrow \text{Cl}_{n+1}$ for which $i'_n(x) = i_n(x)$ whenever $x \in \mathbb{R}^n$. Show that i'_n is injective and determine its image $i'_n \text{Cl}_n \subseteq \text{Cl}_{n+1}$.

b) Show that the \mathbb{R} -linear transformation

$$k_n: \mathbb{R}^n \longrightarrow \text{Cl}_{n+1}; \quad k_n(x) = xe_{n+1},$$

induces an \mathbb{R} -algebra homomorphism $k'_n: \text{Cl}_n \longrightarrow \text{Cl}_{n+1}$ for which $k'_n(x) = k_n(x)$ whenever $x \in \mathbb{R}^n$. Show that k'_n is injective with image $k'_n \text{Cl}_n = \text{Cl}_{n+1}^+$.

5.3. Let $n \geq 1$ and $n \equiv 3 \pmod{4}$. In the Clifford algebra Cl_n , consider the element $\omega_n = e_1 \cdots e_n$.

- a) Show that ω_n is central in Cl_n .
- b) Show that in Cl_n , the elements

$$\Omega_+ = \frac{1}{2}(1 + \omega_n), \quad \Omega_- = \frac{1}{2}(1 - \omega_n),$$

are central orthogonal idempotents for which $\Omega_+ + \Omega_- = 1$.

- c) Decompose Cl_n as a product $\text{Cl}_n = A_+ \times A_-$.
- d) Use the decomposition of (c) to show that $\text{Cl}_3 \cong \mathbb{H} \times \mathbb{H}$ and $\text{Cl}_7 \cong M_8(\mathbb{R}) \times M_8(\mathbb{R})$.

5.4. For $n \geq 1$, define the *finite Clifford group* ClGp_n to be generated by $(n+1)$ elements $\varepsilon_1, \dots, \varepsilon_n, \gamma$ with γ central and subject to the relations

$$\varepsilon_1^2 = \cdots = \varepsilon_n^2 = \gamma, \quad \varepsilon_i \varepsilon_j = \gamma \varepsilon_j \varepsilon_i \text{ if } i \neq j, \quad \gamma^2 = 1.$$

- a) Show that the order of ClGp_n is 2^{n+1} , hence the real group algebra $\mathbb{R}[\text{ClGp}_n]$ has dimension 2^{n+1} .
- b) Show that ClGp_n is isomorphic to a subgroup of Cl_n^\times .
- c) Find an \mathbb{R} -algebra homomorphism $\varphi: \mathbb{R}[\text{ClGp}_n] \longrightarrow \text{Cl}_n$ which is surjective.
- d) Show that the \mathbb{R} -algebra $\mathbb{R}[\text{ClGp}_n]$ is the product algebra

$$\mathbb{R}[\text{ClGp}_n] = A \times B,$$

where $B = \ker \varphi$ and $A \cong \text{Cl}_n$.

5.5. Δ Let $S = [s_{ij}]$ be a $2m \times 2m$ real skew symmetric matrix. Working in the Clifford algebra Cl_{2m} , consider the element

$$s = \sum_{1 \leq i < j \leq 2m} s_{ij} e_i e_j = \frac{1}{2} \sum_{\substack{1 \leq i, j \leq m \\ i \neq j}} s_{ij} e_i e_j.$$

a) Expressing the element s^m in terms of the basis of monomials $e_{i_1} \cdots e_{i_m}$, show that the coefficient of $e_1 e_2 \cdots e_{2m}$ is the real number

$$c_S = m! \sum_{\sigma \in \Pi(2m)} \operatorname{sgn} \sigma s_{\sigma(1)\sigma(2)} \cdots s_{\sigma(2m-1)\sigma(2m)},$$

where $\Pi(2m)$ consists of all permutations σ of $\{1, 2, \dots, 2m\}$ satisfying the conditions

- $\sigma(2r-1) < \sigma(2s-1)$ if $1 \leq r < s \leq m$;
- $\sigma(2t-1) < \sigma(2t)$ if $1 \leq t \leq m$.

Determine c_S for a few small values of m .

b) Determine c_S when $S = \begin{bmatrix} O_m & I_m \\ -I_m & O_m \end{bmatrix}$.

c) If $\Sigma_{2m} \subseteq M_{2m}(\mathbb{R})$ is the subspace of all non-singular skew symmetric matrices, deduce that the function

$$\operatorname{pf}: \Sigma_{2m} \longrightarrow \mathbb{R}^\times; \quad \operatorname{pf}([s_{ij}]) = \frac{c_S}{m!}$$

is continuous. This function pf is known as a *Pfaffian*.

d) Show that pf takes both positive and negative values and in fact is surjective.

e) Let $P = [p_{ij}] \in \text{GL}_{2m}(\mathbb{R})$. Using the well-known formula

$$\det P = \sum_{\sigma \in S_{2m}} \operatorname{sgn} \sigma \prod_{k=1}^{2m} p_{k\sigma(k)},$$

show that

$$\operatorname{pf}(P^T S P) = \operatorname{pf}(S) \det P.$$

- 5.6. a) Verify Proposition 5.31.
 b) For $n \geq 2$, determine $Z(O(n))$.

Lorentz Groups

We met the Lorentz group Lor in Section 1.5. In this chapter, we will consider this kind of group in greater generality. We merely sketch some of the details, leaving the reader to fill in the more obvious gaps. The most important example is that for which $n = 3$ as this provides the geometric setting for Special Relativity. However, many of the main features can be seen in the cases $n = 1, 2$.

6.1 Lorentz Groups

For $n \geq 1$, consider the non-singular symmetric matrix

$$Q_{n,1} = \text{diag}(\underbrace{1, \dots, 1}_n, -1) \in M_{n+1}(\mathbb{R}),$$

which defines an inner product on \mathbb{R}^{n+1} given by

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T Q_{n,1} \mathbf{y} = x_1 y_1 + \cdots + x_n y_n - x_{n+1} y_{n+1}. \quad (6.1)$$

This is the *Lorentz inner product* on \mathbb{R}^{n+1} and it is standard to denote this inner product space by $\mathbb{R}^{n,1}$. Actually, physicists often adopt different sign conventions and use one of the following formulæ in place of Equation (6.1):

$$\langle \mathbf{x}, \mathbf{y} \rangle = \begin{cases} -x_1 y_1 + \cdots + x_n y_n + x_{n+1} y_{n+1}, \\ -x_1 y_1 + \cdots - x_n y_n + x_{n+1} y_{n+1}, \\ x_1 y_1 + \cdots - x_n y_n - x_{n+1} y_{n+1}. \end{cases}$$

This affects the signs in the following definitions. We will also use t in place of x_{n+1} since in Relativity, this coordinate is related to the time measurements while the others are related to spatial ones.

We will often write elements of $\mathbb{R}^{n,1}$ in the form

$$\mathbf{X} = \begin{bmatrix} \mathbf{x} \\ t \end{bmatrix} \quad (\mathbf{x} \in \mathbb{R}^n, t \in \mathbb{R}).$$

Then

$$\langle \mathbf{X}_1, \mathbf{X}_2 \rangle = \mathbf{x}_1 \cdot \mathbf{x}_2 - t_1 t_2.$$

Definition 6.1

A non-zero vector $\mathbf{X} \in \mathbb{R}^{n,1}$ is called

- *spacelike* if $\langle \mathbf{X}, \mathbf{X} \rangle > 0$ or equivalently $|\mathbf{x}|^2 > t^2$;
- *timelike* if $\langle \mathbf{X}, \mathbf{X} \rangle < 0$ or equivalently $|\mathbf{x}|^2 < t^2$;
- a *null vector* if $\langle \mathbf{X}, \mathbf{X} \rangle = 0$ or equivalently $|\mathbf{x}|^2 = t^2$.

We can view \mathbb{R}^n as sitting inside $\mathbb{R}^{n,1}$ as the subset of all spacelike vectors of the form

$$\mathbf{X} = \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 0 \end{bmatrix}$$

and we usually identify $\mathbf{x} \in \mathbb{R}^n$ with the element $\mathbf{X} = \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} \in \mathbb{R}^{n,1}$. Of course we then have $\langle \mathbf{X}, \mathbf{X} \rangle = \mathbf{x} \cdot \mathbf{x}$, so this embedding of \mathbb{R}^n into $\mathbb{R}^{n,1}$ is an *isometry*.

If we represent $\mathbb{R}^{n,1}$ by the (x, t) -plane with \mathbb{R}^n as the x -axis, then the t -axis is timelike, while the null vectors correspond to the points lying on the line $x = t$. In fact, the null vectors in $\mathbb{R}^{n,1}$ lie in the set

$$\mathcal{H}_{n,1}(0) = \{\mathbf{X} \in \mathbb{R}^{n,1} : \mathbf{X} \neq 0, \langle \mathbf{X}, \mathbf{X} \rangle = \mathbf{x} \cdot \mathbf{x} - t^2 = 0\}.$$

For $n > 1$, $\mathcal{H}_{n,1}(0)$ has two path connected components,

$$\mathcal{H}_{n,1}^+(0) = \{\mathbf{X} \in \mathcal{H}_{n,1}(0) : t > 0\}, \quad \mathcal{H}_{n,1}^-(0) = \{\mathbf{X} \in \mathcal{H}_{n,1}(0) : t < 0\}.$$

These are often referred to as the *positive* or *future pointing light cone* and the *negative* or *past pointing light cone* respectively since in Relativity these represent points moving at the speed of light.

For each positive real number r , the hyperboloid

$$\mathcal{H}_{n,1}(r) = \{\mathbf{X} \in \mathbb{R}^{n,1} : \langle \mathbf{X}, \mathbf{X} \rangle = \mathbf{x} \cdot \mathbf{x} - t^2 = -r\}$$

has two path connected components

$$\mathcal{H}_{n,1}^+(r) = \{\mathbf{X} \in \mathcal{H}_{n,1}(r) : t > 0\}, \quad \mathcal{H}_{n,1}^-(r) = \{\mathbf{X} \in \mathcal{H}_{n,1}(r) : t < 0\}.$$

For later use we extend the Lorentz inner product to an \mathbb{R} -bilinear inner product on \mathbb{C}^{n+1} by setting

$$\langle \mathbf{U}, \mathbf{V} \rangle = \mathbf{U}^* Q_{n,1} \mathbf{V} = \mathbf{u}^* \mathbf{v} - \bar{u}v = \mathbf{u} \cdot \mathbf{v} - \bar{u}v$$

for each pair of vectors

$$\mathbf{U} = \begin{bmatrix} \mathbf{u} \\ u \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{v} \\ v \end{bmatrix} \in \mathbb{C}^{n,1}.$$

We will denote the \mathbb{C} -vector space \mathbb{C}^{n+1} with this inner product by $\mathbb{C}^{n,1}$. Then $\langle \cdot, \cdot \rangle$ has all the properties of a hermitian inner product except that the real number $\langle \mathbf{U}, \mathbf{U} \rangle$ is not necessarily non-negative. In particular, for $z, w \in \mathbb{C}$,

$$(z\mathbf{U}, w\mathbf{V}) = \bar{z}w \langle \mathbf{U}, \mathbf{V} \rangle. \quad (6.2)$$

We will refer to a non-zero $\mathbf{U} \in \mathbb{C}^{n,1}$ as spacelike, timelike or null according to whether $\langle \mathbf{U}, \mathbf{U} \rangle > 0$, $\langle \mathbf{U}, \mathbf{U} \rangle < 0$ or $\langle \mathbf{U}, \mathbf{U} \rangle = 0$ as we do for vectors in $\mathbb{R}^{n,1}$.

We will make repeated use of the following result.

Proposition 6.2

Let $\mathbf{U} = \begin{bmatrix} \mathbf{u} \\ u \end{bmatrix}$ and $\mathbf{X} = \begin{bmatrix} \mathbf{x} \\ t \end{bmatrix} \in \mathbb{C}^{n,1}$ be non-zero vectors with $\langle \mathbf{U}, \mathbf{X} \rangle = 0$.

- i) If \mathbf{U} is timelike then $\langle \mathbf{X}, \mathbf{X} \rangle > 0$.
- ii) If \mathbf{U} is null and \mathbf{X} is not a complex multiple of \mathbf{U} , then $\langle \mathbf{X}, \mathbf{X} \rangle > 0$.

Proof

When (i) holds, then $|\mathbf{u}|^2 - |u|^2 < 0$, so $u \neq 0$; when (ii) holds then $\mathbf{U} \neq 0$ and $|\mathbf{u}|^2 = |u|^2$, also giving $u \neq 0$. Since

$$\mathbf{u} \cdot \mathbf{x} - \bar{u}t = 0, \quad (6.3)$$

in either case we have

$$\langle \mathbf{X}, \mathbf{X} \rangle = |\mathbf{x}|^2 - |t|^2 = |\mathbf{x}|^2 - \frac{|\mathbf{u} \cdot \mathbf{x}|^2}{|\mathbf{u}|^2}.$$

In case (i), this gives

$$\langle \mathbf{X}, \mathbf{X} \rangle > |\mathbf{x}|^2 - \frac{|\mathbf{u} \cdot \mathbf{x}|^2}{|\mathbf{u}|^2} \geq 0,$$

since $|\mathbf{u} \cdot \mathbf{x}| \leq |\mathbf{u}||\mathbf{x}|$. In case (ii),

$$\langle \mathbf{X}, \mathbf{X} \rangle = |\mathbf{x}|^2 - \frac{|\mathbf{u} \cdot \mathbf{x}|^2}{|\mathbf{u}|^2} \geq 0,$$

with equality if and only if $|\mathbf{u} \cdot \mathbf{x}| = |\mathbf{u}||\mathbf{x}|$, i.e., $\mathbf{x} = z\mathbf{u}$ for some $z \in \mathbb{C}$ with $|z| = 1$. In the latter case, from Equation (6.3) we would have

$$t\bar{u} = z|\mathbf{u}|^2 = z\mathbf{u}\bar{u},$$

so $t = z\mathbf{u}$, in turn giving $\mathbf{X} = z\mathbf{U}$ and so $\langle \mathbf{X}, \mathbf{X} \rangle = 0$. \square

We can consider the closed subgroup

$$O(n, 1) = \{A \in GL_{n+1}(\mathbb{R}) : A^T Q_{n,1} A = Q_{n,1}\} \leqslant GL_{n+1}(\mathbb{R}).$$

Note that $A \in O(n, 1)$ if and only if for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n,1}$, $\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$, so $O(n, 1)$ consists of all the Lorentzian isometries of $\mathbb{R}^{n,1}$. Notice that for $A \in O(n, 1)$ we have $(\det A)^2 = 1$, so $\det A = \pm 1$. Also $A \in O(n, 1)$ preserves each of the sets of spacelike, timelike and null vectors; it also either preserves or interchanges the positive and negative light cones. It is standard to set

$$SO(n, 1) = \{A \in O(n, 1) : \det A = 1\} = O(n, 1) \cap SL_{n+1}(\mathbb{R}) \leqslant O(n, 1).$$

However, $SO(n, 1)$ is not connected since its elements can interchange the sets $\mathcal{H}_{n,1}^\pm(r)$ for $r \geq 0$. We define the Lorentz group of $\mathbb{R}^{n,1}$ to be the closed subgroup of $SO(n, 1)$ preserving each of the connected sets $\mathcal{H}_{n,1}^\pm(1)$,

$$Lor(n, 1) = \{A \in SO(n, 1) : A\mathcal{H}_{n,1}^\pm(1) = \mathcal{H}_{n,1}^\pm(1)\} \leqslant SO(n, 1).$$

Proposition 6.3

For $r \geq 0$,

$$Lor(n, 1) = \{A \in SO(n, 1) : A\mathcal{H}_{n,1}^\pm(r) = \mathcal{H}_{n,1}^\pm(r)\} \leqslant SO(n, 1).$$

Proof

It is easy to verify this for $r > 0$. For $r = 0$, notice that every null vector $\mathbf{U} = \begin{bmatrix} \mathbf{u} \\ u \end{bmatrix}$ can be expressed as a limit

$$\mathbf{U} = \lim_{k \rightarrow \infty} \mathbf{V}_k$$

with each \mathbf{V}_k timelike. We can even arrange that for every k , $\mathbf{V}_k = \begin{bmatrix} \mathbf{v}_k \\ u_k \end{bmatrix}$ with $u\mathbf{v}_k > 0$ (or $u\mathbf{v}_k < 0$). Since the action of each $A \in \text{Lor}(n, 1)$ on $\mathbb{R}^{n,1}$ is continuous,

$$A\mathbf{U} = \lim_{k \rightarrow \infty} A\mathbf{V}_k,$$

from which we see that A preserves each of the sets $\mathcal{H}_{n,1}^{\pm}(0)$.

On the other hand, suppose that $A \in \text{SO}(n, 1)$ preserves each of the connected sets $\mathcal{H}_{n,1}^{\pm}(0)$. If A fails to preserve one (and hence all) of the sets $\mathcal{H}_{n,1}^{+}(s)$ for $s > 0$, then as any $\mathbf{U} \in \mathcal{H}_{n,1}^{+}(0)$ is a limit

$$\mathbf{U} = \lim_{k \rightarrow \infty} \mathbf{V}_k$$

with $\mathbf{V}_k \in \mathcal{H}_{n,1}^{+}(s_k)$ for $s_k > 0$, we have

$$A\mathbf{U} = \lim_{k \rightarrow \infty} A\mathbf{V}_k$$

where $A\mathbf{V}_k \in \mathcal{H}_{n,1}^{-}(s'_k)$ for $s'_k > 0$. This implies that $A\mathbf{U} \in \mathcal{H}_{n,1}^{-}(0)$, contradicting the original assumption on A . \square

Notice that $A^* = A^T$ if $A \in \text{Lor}(n, 1)$, so if $\mathbf{U}, \mathbf{V} \in \mathbb{C}^{n,1}$, we have

$$(A\mathbf{U}, A\mathbf{V}) = \mathbf{U}^* A^* Q_{n,1} A \mathbf{V} = \mathbf{U}^* Q_{n,1} \mathbf{V} = (\mathbf{U}, \mathbf{V}). \quad (6.4)$$

When $n = 3$, $\text{Lor}(3, 1) = \text{Lor}$ in the notation of Section 1.5.

Example 6.4

We have

$$\text{Lor}(1, 1) = \left\{ \begin{bmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{bmatrix} : s \in \mathbb{R} \right\},$$

$$\text{SO}(1, 1) = \text{Lor}(1, 1) \cup \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \text{Lor}(1, 1),$$

$$\text{O}(1, 1) = \text{SO}(1, 1) \cup \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{SO}(1, 1).$$

Proof

If $A \in \text{O}(1, 1)$ has $\det A = -1$ then

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} A \in \text{SO}(1, 1),$$

so it suffices to determine $\text{SO}(1, 1)$. Similarly, if $B \in \text{SO}(1, 1)$ with $B\mathcal{H}_{n,1}^+(0) = \mathcal{H}_{n,1}^-(0)$ then

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} B \in \text{Lor}(1, 1),$$

so it suffices to determine $\text{Lor}(1, 1)$.

Let $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \text{Lor}(1, 1)$. Since

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix},$$

we must have $d > 0$. Also, the simultaneous equations

$$\begin{cases} a^2 - c^2 = 1, \\ d^2 - b^2 = 1, \\ cd - ab = 0, \\ ad - bc = 1, \end{cases}$$

are satisfied. Putting $d = \cosh s$ and $b = \sinh s$ for suitable $s \in \mathbb{R}$, we easily obtain $a = \cosh s$, $c = \sinh s$. Hence every element of $\text{Lor}(1, 1)$ has the form $\begin{bmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{bmatrix}$ for some $s \in \mathbb{R}$. \square

Remark 6.5

The Lie algebra of $\text{Lor}(1, 1)$ is easily determined to be

$$\mathfrak{lor}(1, 1) = \left\{ \begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix} : t \in \mathbb{R} \right\}$$

with trivial Lie bracket. The exponential map $\exp: \mathfrak{lor}(1, 1) \rightarrow \text{Lor}(1, 1)$ is surjective since for $s \in \mathbb{R}$,

$$\exp \left(\begin{bmatrix} 0 & s \\ s & 0 \end{bmatrix} \right) = \begin{bmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{bmatrix}. \quad (6.5)$$

As we will eventually see, the exponential map for $\text{Lor}(n, 1)$ is actually surjective for every $n \geq 1$.

For future use we record the following explicit form for elements of $\text{Lor}(n, 1)$, whose proof is a direct consequence of the definition of the Lorentz group.

Proposition 6.6

A matrix

$$P = [U_1 \ \cdots \ U_{n+1}] \in GL_{n+1}(\mathbb{R})$$

is in $Lor(n, 1)$ if and only if its column vectors $U_k = \begin{bmatrix} u_j \\ u_j \end{bmatrix}$ satisfy the orthonormality equations

$$\langle U_j, U_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k \leq n, \\ -1 & \text{if } j = k = n. \end{cases}$$

In particular U_1, \dots, U_n are spacelike and U_{n+1} is timelike with $u_{n+1} > 0$.

Theorem 6.7

For $n \geq 1$, $Lor(n, 1)$ is path connected.

Proof

We will show that $Lor(n, 1)$ has a connected subgroup H for which the homogeneous space $Lor(n, 1)/H$ is connected. Together with Proposition 9.10, this implies that $Lor(n, 1)$ is connected.

Consider the continuous action of $Lor(n, 1)$ on $\mathbb{R}^{n,1}$ by left multiplication. The stabiliser of the timelike vector e_{n+1} is

$$Stab_{Lor(n,1)}(e_{n+1}) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ O_{1,n-1} & 1 \end{bmatrix} : A \in SO(n) \right\} \leqslant Lor(n, 1)$$

which is a closed subgroup obviously isomorphic to $SO(n)$, hence connected by Example 9.14. Next we need to identify the orbit of e_{n+1} .

Suppose that $U = \begin{bmatrix} u \\ u \end{bmatrix} \in \mathcal{H}_{n,1}^+(1)$, so $|u|^2 - u^2 = -1$ and $u > 0$. We will show that

$$\begin{bmatrix} u \\ u \end{bmatrix} \in Orb_{Lor(n,1)}(e_{n+1}).$$

Consider the equation

$$u \cdot x - ut = 0$$

for the pair $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. The solution set is an n -dimensional vector subspace of $\mathbb{R}^{n,1}$; furthermore, by Proposition 6.2 its non-zero vectors are all

spacelike. Choosing a basis of this spacelike subspace, then applying the Gram-Schmidt process, we can produce a basis consisting of vectors

$$\begin{bmatrix} \mathbf{w}_1 \\ s_1 \end{bmatrix}, \dots, \begin{bmatrix} \mathbf{w}_n \\ s_n \end{bmatrix}$$

for which

$$\begin{cases} \mathbf{w}_i \cdot \mathbf{w}_j - s_i s_j = \delta_{ij} & (i, j = 1, \dots, n), \\ \mathbf{u} \cdot \mathbf{w}_k - u s_k = 0 & (k = 1, \dots, n). \end{cases}$$

The resulting matrix

$$P = \begin{bmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_n & \mathbf{u} \\ s_1 & \cdots & s_n & u \end{bmatrix}$$

is in $\text{Lor}(n, 1)$ and satisfies

$$P \mathbf{e}_{n+1} = \begin{bmatrix} \mathbf{u} \\ u \end{bmatrix}.$$

So

$$\text{Orb}_{\text{Lor}(n, 1)}(\mathbf{e}_{n+1}) = \mathcal{H}_{n, 1}^+(1).$$

This also shows that $\text{Orb}_{\text{Lor}(n, 1)}(\mathbf{e}_{n+1})$ is path connected since we already know that $\mathcal{H}_{n, 1}^+(1)$ is. Thus taking $H = \text{Stab}_{\text{Lor}(n, 1)}(\mathbf{e}_{n+1}) \cong \text{SO}(n)$ we see that $\text{Lor}(n, 1)$ is path connected. \square

This identification of the homogeneous space $\text{Lor}(n, 1)/\text{SO}(n)$ can also be used to calculate the dimension of $\text{Lor}(n, 1)$.

Proposition 6.8

The dimension of the Lorentz group $\text{Lor}(n, 1)$ is

$$\dim \text{Lor}(n, 1) = \binom{n+1}{2}.$$

Proof

Using the Implicit Function Theorem 7.11, it is easy to see that

$$\dim \text{Lor}(n, 1)/\text{SO}(n) = \dim \mathcal{H}_{n, 1}^+(1) = n.$$

Also from Section 3.3 we know that

$$\dim \text{SO}(n) = \binom{n}{2}.$$

Combining these we obtain the result. \square

6.2 A Principal Axis Theorem for Lorentz Groups

We will make use of the Principal Axis Theorem for special orthogonal groups which will be discussed later as part of Theorem 10.13. This says that every matrix $A \in SO(m)$ has the form

$$A = \begin{cases} PR_{2n}(\theta_1, \dots, \theta_n)P^T & \text{if } m = 2n \text{ is even,} \\ PR_{2n+1}(\theta_1, \dots, \theta_n)P^T & \text{if } m = 2n + 1 \text{ is odd,} \end{cases}$$

where for $\theta_1, \dots, \theta_n, \theta \in \mathbb{R}$,

$$R_{2n}(\theta_1, \dots, \theta_n) = \begin{bmatrix} R(\theta_1) & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & R(\theta_2) & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & R(\theta_n) & \end{bmatrix},$$

$$R_{2n+1}(\theta_1, \dots, \theta_n) = \begin{bmatrix} R(\theta_1) & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & R(\theta_2) & 0 & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & R(\theta_n) & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix},$$

$$R_{2n+1}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

The following result leads to the analogue for Lorentz groups.

Theorem 6.9

For $n \geq 1$, every Lorentz matrix $A \in \text{Lor}(n, 1)$ has one of the forms

$$A = P \begin{bmatrix} B & O_{n-2,2} \\ O_{2,n-2} & \begin{matrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{matrix} \end{bmatrix} P^{-1}, \quad (\text{I})$$

$$A = P \begin{bmatrix} C & O_{n,1} \\ O_{1,n} & 1 \end{bmatrix} P^{-1}, \quad (\text{II})$$

where $B \in SO(n-1)$, $C \in SO(n)$, $t \in \mathbb{R}$ and $P \in \text{Lor}(n, 1)$.

Proof

We will prove this by induction on $n \geq 1$. The initial case $n = 1$ is dealt with in Example 6.4. So we suppose that the result holds for $\text{Lor}(k, 1)$ whenever $1 \leq k \leq n - 1$.

We begin by analysing the eigenvalues and eigenvectors of A . Suppose that $\lambda \in \mathbb{C}$ is an eigenvalue for $A \in \text{Lor}(n, 1)$ with eigenvector $\mathbf{U} \in \mathbb{C}^{n,1}$. If \mathbf{U} is not a null vector, we have

$$\begin{aligned} |\lambda|^2 \langle \mathbf{U}, \mathbf{U} \rangle &= (\lambda \mathbf{U}, \lambda \mathbf{U}) \\ &= (A\mathbf{U}, A\mathbf{U}) \\ &= \langle \mathbf{U}, \mathbf{U} \rangle, \end{aligned}$$

hence $|\lambda| = 1$. On the other hand, if \mathbf{U} is a null vector then $|\lambda|$ is potentially unrestricted.

Next suppose that λ and μ are distinct eigenvalues for A with eigenvectors \mathbf{U} and \mathbf{V} . Then

$$\begin{aligned} \bar{\lambda}\mu \langle \mathbf{U}, \mathbf{V} \rangle &= (\lambda \mathbf{U}, \mu \mathbf{V}) \\ &= (A\mathbf{U}, A\mathbf{V}) \\ &= \langle \mathbf{U}, \mathbf{V} \rangle, \end{aligned}$$

hence either $\langle \mathbf{U}, \mathbf{V} \rangle = 0$ or $\mu = \bar{\lambda}^{-1}$. In particular, if $|\lambda| = 1$ we must also have $\bar{\lambda}^{-1} = \lambda \neq \mu$, yielding $\langle \mathbf{U}, \mathbf{V} \rangle = 0$.

Next we will investigate in more detail the situation when A has a null eigenvector $\mathbf{U} = \begin{bmatrix} \mathbf{u} \\ u \end{bmatrix}$ for an eigenvalue λ . Since $|\mathbf{u}|^2 = |u|^2 \neq 0$, we can multiply by a suitable complex scalar to ensure that $u = 1$, so we might as well assume that $\mathbf{U} = \begin{bmatrix} \mathbf{u} \\ 1 \end{bmatrix}$. Notice that

$$A \begin{bmatrix} \mathbf{u} \\ 1 \end{bmatrix} = \overline{\left(A \begin{bmatrix} \mathbf{u} \\ 1 \end{bmatrix} \right)} = \bar{\lambda} \begin{bmatrix} \bar{\mathbf{u}} \\ 1 \end{bmatrix},$$

showing that $\bar{\lambda}$ is also an eigenvalue with eigenvector $\bar{\mathbf{U}} = \begin{bmatrix} \bar{\mathbf{u}} \\ 1 \end{bmatrix}$. Also,

$$\begin{aligned} \langle \bar{\mathbf{U}}, \mathbf{U} \rangle &= \langle A\bar{\mathbf{U}}, A\mathbf{U} \rangle \\ &= \langle \bar{\lambda}\bar{\mathbf{U}}, \lambda\mathbf{U} \rangle \\ &= \lambda^2 \langle \bar{\mathbf{U}}, \mathbf{U} \rangle. \end{aligned}$$

As $\bar{\mathbf{U}}$ is a null vector, Proposition 6.2 implies that either $\bar{\mathbf{U}} = \mathbf{U}$ or $\lambda^2 = 1$. Each of these possibilities means that λ is real and in fact positive, since A

preserves the positive light cone. So for a null eigenvector, the corresponding eigenvalue has to be a positive real number.

Continuing with this situation, consider the solution set of the linear equation

$$(\mathbf{U}, \mathbf{X}) = 0. \quad (6.6)$$

This forms a vector subspace of $\mathbb{R}^{n,1}$ of dimension n , containing the subspace of real multiples of \mathbf{U} as well as the $(n - 1)$ -dimensional subspace of spacelike elements together with zero. In fact, the spacelike elements are all the vectors of the form

$$\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} \quad (\mathbf{x} \in \mathbb{R}^n, \mathbf{u} \cdot \mathbf{x} = 0).$$

Thus the general solution of Equation (6.6) is

$$\mathbf{X} = \begin{bmatrix} \mathbf{x} + z\mathbf{u} \\ z \end{bmatrix} \quad (\mathbf{x} \in \mathbb{R}^n, \mathbf{u} \cdot \mathbf{x} = 0, z \in \mathbb{R}).$$

Then there is a unique null vector $\tilde{\mathbf{U}} \in \mathbb{R}^{n,1}$ possessing the two properties

- $\langle \tilde{\mathbf{U}}, \mathbf{X} \rangle = 0$ for all spacelike solutions of Equation (6.6),
- $\langle \mathbf{U}, \tilde{\mathbf{U}} \rangle = -2$.

This vector is easily seen to be $\tilde{\mathbf{U}} = \begin{bmatrix} -\mathbf{u} \\ 1 \end{bmatrix}$. Clearly the spacelike solutions of Equation (6.6) are preserved by A , and indeed A acts as an isometry on this space so relative to a basis it is given by an orthogonal matrix. Also

$$\begin{aligned} \langle \mathbf{U}, A(\lambda \tilde{\mathbf{U}}) \rangle &= \lambda \langle A^{-1}\mathbf{U}, \tilde{\mathbf{U}} \rangle \\ &= \lambda \langle \lambda^{-1}\mathbf{U}, \tilde{\mathbf{U}} \rangle \\ &= \lambda \lambda^{-1} \langle \mathbf{U}, \tilde{\mathbf{U}} \rangle \\ &= -2. \end{aligned}$$

Using the above characterisation of $\tilde{\mathbf{U}}$, we see that

$$A\tilde{\mathbf{U}} = \lambda^{-1}\tilde{\mathbf{U}},$$

i.e., $\tilde{\mathbf{U}}$ is an eigenvector of A for the eigenvalue λ^{-1} .

The vectors

$$\mathbf{U}' = \frac{1}{2}(\mathbf{U} - \tilde{\mathbf{U}}), \quad \mathbf{U}'' = \frac{1}{2}(\mathbf{U} + \tilde{\mathbf{U}}),$$

satisfy

$$\begin{aligned}AU' &= \frac{(\lambda + \lambda^{-1})}{2} \mathbf{U}' + \frac{(\lambda - \lambda^{-1})}{2} \mathbf{U}'', \\AU'' &= \frac{(\lambda - \lambda^{-1})}{2} \mathbf{U}' + \frac{(\lambda + \lambda^{-1})}{2} \mathbf{U}'',\end{aligned}$$

and setting $s = \ln \lambda$, we have

$$\begin{aligned}AU' &= \cosh s \mathbf{U}' + \sinh s \mathbf{U}'', \\AU'' &= \sinh s \mathbf{U}' + \cosh s \mathbf{U}''.\end{aligned}$$

Also

$$(\mathbf{U}', \mathbf{U}'') = 0, \quad (\mathbf{U}', \mathbf{U}') = 1, \quad (\mathbf{U}'', \mathbf{U}'') = -1,$$

so \mathbf{U}' is a unit spacelike vector, while \mathbf{U}'' is timelike and these are orthogonal vectors with respect to the Lorentzian inner product.

Together with an orthonormal basis $\{\mathbf{U}_1, \dots, \mathbf{U}_{n-1}\}$ of the spacelike solutions of Equation (6.6), we see that $\mathbb{R}^{n,1}$ has a basis

$$\{\mathbf{U}_1, \dots, \mathbf{U}_{n-1}, \mathbf{U}', \mathbf{U}''\}$$

which provides the columns of a Lorentz matrix

$$P = [\mathbf{U}_1 \ \cdots \ \mathbf{U}_{n-1} \ \mathbf{U}' \ \mathbf{U}'']$$

for which

$$A = P \begin{bmatrix} B & O_{n-2,2} \\ O_{2,n-2} & \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix} \end{bmatrix} P^{-1}$$

with $B \in O(n-1)$; in fact $\det B = 1$ since $\det A = 1$ and also

$$\det \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix} = \cosh^2 t - \sinh^2 t = 1.$$

Hence A has the form (I) whenever it has a null vector for an eigenvector.

Now we consider the situation where A has a timelike eigenvector \mathbf{U} for the eigenvalue λ . Using similar arguments to those for a null eigenvector, we find that $\lambda = 1$ and $\mathbf{U} \in \mathbb{R}^{n,1}$. Moreover, A acts as an isometry on the solution set of

$$(\mathbf{U}, \mathbf{X}) = 0,$$

which is an n -dimensional subspace of $\mathbb{R}^{n,1}$ with timelike non-zero vectors. Hence A has the form (II).

The remaining case to deal with is that where there are no null or timelike eigenvectors. So suppose that $\mathbf{U} = \begin{bmatrix} \mathbf{u} \\ u \end{bmatrix}$ is a spacelike eigenvector of unit

length with eigenvalue λ of unit modulus. By multiplying by a suitable complex number we can assume that $u \in \mathbb{R}$ and $u \geq 0$.

When $\mathbf{U} \in \mathbb{R}^{n,1}$, we have $\lambda = \pm 1$ and $\det A = 1$.

If $\lambda = 1$ then the solutions of the equation

$$(\mathbf{U}, \mathbf{X}) = 0$$

form a subspace of $\mathbb{R}^{n,1}$ closed under the action of A and indeed, on choosing a basis for this subspace we find that it is isomorphic to $\mathbb{R}^{n-1,1}$ and the action of A is given by an element of $\text{Lor}(n-1, 1)$. Hence there is a $Q \in \text{Lor}(n, 1)$ and $B \in \text{Lor}(n-1, 1)$ for which

$$A = Q \begin{bmatrix} 1 & O_{1,n} \\ O_{n,1} & B \end{bmatrix} Q^{-1}.$$

By the induction hypothesis, B has one of the forms (I) or (II), from which we deduce that A does too.

If $\lambda = -1$, this is an eigenvalue of multiplicity at least 2 since non-real complex eigenvalues come in complex conjugate pairs. With the aid of the Gram-Schmidt process we may produce a second eigenvector \mathbf{V} for this eigenvalue of unit length and orthogonal to \mathbf{U} . The solutions of the pair of simultaneous equations

$$(\mathbf{U}, \mathbf{X}) = 0 = (\mathbf{V}, \mathbf{X})$$

form a subspace of $\mathbb{R}^{n,1}$ closed under the action of A . On choosing a basis for this subspace we find that it is isomorphic to $\mathbb{R}^{n-2,1}$ and the action of A is given by an element of $\text{Lor}(n-2, 1)$. Hence there is a $Q \in \text{Lor}(n, 1)$ and $B' \in \text{Lor}(n-2, 1)$ for which

$$A = Q \begin{bmatrix} -I_2 & O_{2,n-1} \\ O_{n-1,2} & B' \end{bmatrix} Q^{-1}.$$

Again the induction hypothesis implies that B' has one of the forms (I) or (II), and therefore A does too.

When $\mathbf{U} \notin \mathbb{R}^{n,1}$, $\overline{\mathbf{U}} = \begin{bmatrix} \overline{\mathbf{u}} \\ u \end{bmatrix}$ is also a spacelike eigenvector of unit length for the eigenvalue $\bar{\lambda}$. The vectors

$$\mathbf{U}' = \frac{1}{2}(\mathbf{U} + \overline{\mathbf{U}}), \quad \mathbf{U}'' = \frac{1}{2}(\mathbf{U} - \overline{\mathbf{U}}),$$

are orthonormal and satisfy the equations

$$A\mathbf{U}' = \cos \theta \mathbf{U}' + \sin \theta \mathbf{U}'', \quad A\mathbf{U}'' = -\sin \theta \mathbf{U}' + \cos \theta \mathbf{U}'',$$

where $\lambda = e^{\theta i}$. As in the previous case, the set of solutions of the pair of simultaneous equations

$$(\mathbf{U}', \mathbf{X}) = 0 = (\mathbf{U}'', \mathbf{X})$$

forms a Lorentzian subspace which can be identified with $\text{Lor}(n-2, 1)$ after choosing a basis, and A acts on it as an element of $\text{Lor}(n-2, 1)$. Again there is a $Q \in \text{Lor}(n, 1)$ and $B'' \in \text{Lor}(n-2, 1)$ for which

$$A = Q \begin{bmatrix} R(\theta) & O_{2,n-1} \\ O_{n-1,2} & B'' \end{bmatrix} Q^{-1}.$$

By the induction hypothesis, B'' has one of the forms (I) or (II), from which we deduce that A does too. \square

Combining this result with the Principle Axis Theorem 10.13 for $\text{SO}(k)$, we obtain the following.

Theorem 6.10 (Principle Axis Theorem for the Lorentz group)

For $n \geq 1$, each matrix $A \in \text{Lor}(n, 1)$ has one of the forms

$$A = Q \begin{bmatrix} R_{n-1}(\theta_1, \dots, \theta_\ell) & O_{n-2,2} \\ O_{2,n-2} & \begin{matrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{matrix} \end{bmatrix} Q^{-1}, \quad (\text{I}')$$

$$A = Q \begin{bmatrix} R_n(\theta_1, \dots, \theta_m) & O_{n,1} \\ O_{1,n} & 1 \end{bmatrix} Q^{-1}, \quad (\text{II}'')$$

where $t \in \mathbb{R}$, $Q \in \text{Lor}(n, 1)$ and in case (I'), $n = 2\ell + 1$ or $n = 2\ell + 2$, while in case (II''), $n = 2m$ or $n = 2m + 1$.

We can also deduce information on the exponential for Lorentz groups.

Theorem 6.11

For $n \geq 1$, the exponential map $\exp: \text{lor}(n, 1) \rightarrow \text{Lor}(n, 1)$ is surjective.

Proof

When $n = 1$, we know from Remark 6.5 that every element of $\text{Lor}(1, 1)$ has the form given in (I). Also, by Theorem 10.16, $\exp: \text{so}(k) \rightarrow \text{SO}(k)$ is surjective

for $k \geq 1$. Let $A \in \text{Lor}(n, 1)$ for $n \geq 2$. When A has the form (I),

$$A = P \begin{bmatrix} \exp(S) & O_{n-2,2} \\ O_{2,n-2} & \exp\left(\begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix}\right) \end{bmatrix} P^{-1} = \exp\left(P \begin{bmatrix} S & O_{n-2,2} \\ O_{2,n-2} & \begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix} \end{bmatrix} P^{-1}\right)$$

for some $S \in \mathfrak{so}(n-1)$. On the other hand, when A has the form (II),

$$A = P \begin{bmatrix} \exp(S) & O_{n,1} \\ O_{1,n} & 1 \end{bmatrix} P^{-1} = \exp\left(P \begin{bmatrix} S & O_{n,1} \\ O_{1,n} & 0 \end{bmatrix} P^{-1}\right)$$

for some $S \in \mathfrak{so}(n)$. In either case, $A \in \text{explor}(n, 1)$. \square

6.3 $\text{SL}_2(\mathbb{C})$ and the Lorentz Group $\text{Lor}(3, 1)$

Let us now consider the Lie algebra $\text{SL}_2(\mathbb{C})$, $\mathfrak{sl}_2(\mathbb{C})$. By Equation (3.8),

$$\mathfrak{sl}_2(\mathbb{C}) = \ker \text{tr} \subseteq M_2(\mathbb{C})$$

and so $\dim_{\mathbb{C}} \mathfrak{sl}_2(\mathbb{C}) = 3$. 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

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

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 $\mathfrak{ut}_2^0(\mathbb{C})$ of the group of *upper triangular* complex matrices of determinant 1, while H', F' span the Lie algebra of the group of *lower triangular* complex matrices of determinant 1.

Given the existence of the double covering homomorphism $\overline{\text{Ad}}: \text{SU}(2) \rightarrow \text{SO}(3)$ of Section 3.5, it seems reasonable to ask if a similar homomorphism exists for $\text{SL}_2(\mathbb{C})$. It does, but we need to use the Lorentz group $\text{Lor}(3, 1)$ and obtain an important double covering homomorphism $\text{SL}_2(\mathbb{C}) \rightarrow \text{Lor}(3, 1)$ which appears in Physics in connection with spinors and twistors.

Next we will describe the \mathbb{R} -Lie algebra $\text{lor}(3, 1)$ of $\text{Lor}(3, 1) \leq \text{SL}_4(\mathbb{R})$. Let $\alpha: (-\varepsilon, \varepsilon) \rightarrow \text{Lor}(3, 1)$ 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

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

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

$$AQ + QA^T = O \quad (\text{and hence } \text{tr } A = 0)$$

is in $\text{Ior}(3, 1)$. To see this, consider the curve

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

for which $\alpha'(0) = A$. Since $QA^T = -AQ$, we have

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

By Lemma 3.23,

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

Finally, for each $t \in \mathbb{R}$, $\alpha(t)$ preserves each of the components $\mathcal{H}_{3,1}^\pm(1)$ of the hyperboloid

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

All of this shows that $\exp(tA) \in \operatorname{Lor}(3,1)$ so we might as well redefine

$$\alpha: \mathbb{R} \longrightarrow \operatorname{Lor}(3,1); \quad \alpha(t) = \exp(tA).$$

We also have $A \in \operatorname{lor}(3,1)$. So we have

$$\operatorname{lor}(3,1) = \{A \in M_4(\mathbb{R}) : AQ + QA^T = O\}$$

$$= \left\{ A \in M_4(\mathbb{R}) : A = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ a_{14} & a_{24} & a_{34} & 0 \end{bmatrix} \right\} \quad (6.7)$$

and therefore

$$\dim \operatorname{Lor}(3,1) = \dim \operatorname{lor}(3,1) = 6. \quad (6.8)$$

An \mathbb{R} -basis for $\operatorname{lor}(3,1)$ consists of the elements

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

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

The non-trivial brackets for these are

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

We will now define the homomorphism $\operatorname{SL}_2(\mathbb{C}) \rightarrow \operatorname{Lor}(3,1)$. To do this we will identify the 2×2 skew hermitian matrices $\operatorname{Sk-Herm}_2(\mathbb{C})$ with \mathbb{R}^4 using the correspondence

$$\begin{bmatrix} (t+x)i & y+zi \\ -y+zi & (t-x)i \end{bmatrix} \longleftrightarrow xe_1 + ye_2 + ze_3 + te_4.$$

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

$$\langle S_1 | S_2 \rangle = \frac{1}{4} (\det(S_1 + S_2) - \det(S_1 - S_2)). \quad (6.9a)$$

When $S_1 = S_2 = S$ we have

$$\langle S | S \rangle = \frac{1}{4} (\det 2S - \det O) = \det S. \quad (6.9b)$$

It is easy to check that

$$\begin{aligned} & \left\langle \begin{bmatrix} (t_1 + x_1)i & y_1 + z_1i \\ -y_1 + z_1i & (t_1 - x_1)i \end{bmatrix} \mid \begin{bmatrix} (t_2 + x_2)i & y_2 + z_2i \\ -y_2 + z_2i & (t_2 - x_2)i \end{bmatrix} \right\rangle \\ &= x_1 x_2 + y_1 y_2 + z_1 z_2 - t_1 t_2, \end{aligned} \quad (6.10)$$

which is the Lorentzian inner product on $\mathbb{R}^{3,1} = \mathbb{R}^4$ given by

$$x_1 x_2 + y_1 y_2 + z_1 z_2 - t_1 t_2 = [x_1 \ y_1 \ z_1 \ t_1] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \\ t_2 \end{bmatrix}.$$

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

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

so $ASA^* \in \text{Sk-Herm}_2(\mathbb{C})$. By Equation (6.9a), for $S_1, S_2 \in \text{Sk-Herm}_2(\mathbb{C})$, and the fact that $\det A = 1 = \det A^*$,

$$\begin{aligned} \langle AS_1 A^* | AS_2 A^* \rangle &= \frac{1}{4} (\det A (S_1 + S_2) A^* - \det A (S_1 - S_2) A^*) \\ &= \frac{1}{4} (\det A \det(S_1 + S_2) \det A^* - \det A \det(S_1 - S_2) \det A^*) \\ &= \frac{1}{4} (\det(S_1 + S_2) - \det(S_1 - S_2)) \\ &= \langle S_1 | S_2 \rangle. \end{aligned}$$

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 $(|)$. We can identify this with an \mathbb{R} -linear transformation $\widetilde{\text{Ad}}_A: \mathbb{R}^{3,1} \rightarrow \mathbb{R}^{3,1}$ which preserves the Lorentzian inner product. In fact, $\det \widetilde{\text{Ad}}_A = 1$ and $\widetilde{\text{Ad}}_A$ preserves the components of the hyperboloid $x^2 + y^2 + z^2 - t^2 = -1$. Let

$$\widetilde{\text{Ad}}: \text{SL}_2(\mathbb{C}) \longrightarrow \text{Lor}(3,1); \quad \widetilde{\text{Ad}}(A) = \widetilde{\text{Ad}}_A.$$

Then $\widetilde{\text{Ad}}$ is a homomorphism since

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

It is clearly continuous. Also, $A \in \ker \widetilde{\text{Ad}}$ 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 $A = \pm I$. Thus we have $\ker \widetilde{\text{Ad}} = \{\pm I\}$.

Theorem 6.12

- i) The derivative $d \widetilde{\text{Ad}}: \mathfrak{sl}_2(\mathbb{C}) \rightarrow \text{Lor}(3, 1)$ is an isomorphism of \mathbb{R} -Lie algebras.
- ii) $\widetilde{\text{Ad}}: \text{SL}_2(\mathbb{C}) \rightarrow \text{Lor}(3, 1)$ is a continuous surjective Lie homomorphism for which $\ker \widetilde{\text{Ad}} = \{\pm I\}$, hence

$$\text{SL}_2(\mathbb{C})/\{\pm I\} \cong \text{Lor}(3, 1).$$

Proof

(i) As in the case of $\text{SU}(2)$ and $\text{SO}(3)$, we can determine the derivative $d \widetilde{\text{Ad}}$ by considering for each $C \in \mathfrak{sl}_2(\mathbb{C})$, the curve

$$\gamma: \mathbb{R} \rightarrow \text{SL}_2(\mathbb{C}); \quad \gamma(t) = \exp(tC),$$

which gives rise to the curve

$$\bar{\gamma}: \mathbb{R} \rightarrow \text{Lor}(3, 1); \quad \bar{\gamma}(t) = \widetilde{\text{Ad}}_{\gamma(t)}.$$

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 $\widetilde{\text{Ad}}_{\gamma(t)}$ on $\text{Sk-Herm}_2(\mathbb{C})$ and interpret it as an element of $\text{Lor}(3, 1)$. Differentiating at $t = 0$, we obtain the action of C as an element of $\text{Lor}(3, 1)$ and so $d \widetilde{\text{Ad}}(C)$. For $X \in \text{Sk-Herm}_2(\mathbb{C})$ we have

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

hence

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

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

$$H(x_1 V_1 + x_2 V_2 + x_3 V_3 + x_4 V_4) + (x_1 V_1 + x_2 V_2 + x_3 V_3 + x_4 V_4)H^* = x_2 V_3 - x_3 V_2,$$

so

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

Here is the complete list written in terms of the matrices P_{rs} , which we know form an \mathbb{R} -basis of $\text{Lor}(3,1)$:

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

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

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

This shows that $d\widetilde{\text{Ad}}(C)$ maps a basis for $\text{Sk-Herm}_2(\mathbb{C})$ to one for $\text{Lor}(3,1)$, thus it is an isomorphism of Lie algebras.

(ii) The surjectivity of $\widetilde{\text{Ad}}$ follows immediately from the surjectivity of the derivative in (i) together with Theorem 6.11. \square

EXERCISES

6.1. a) For $t \in \mathbb{R}$, show that the matrix

$$A_t = \begin{bmatrix} \cosh t & 0 & 0 & -\sinh t \\ 0 & \cos t & -\sin t & 0 \\ 0 & \sin t & \cos t & 0 \\ -\sinh t & 0 & 0 & \cosh t \end{bmatrix}$$

is in $\text{Lor}(3,1)$.

- b) Express A_t in one of the forms predicted by Theorem 6.10.
- c) Express A_t in the form $\exp(tU)$ for some $U \in \text{Lor}(3,1)$.
- d) Find a differential equation satisfied by the function

$$\alpha: \mathbb{R} \longrightarrow \text{Lor}(3,1); \quad \alpha(t) = A_t.$$

6.2. In this question, we freely use the notation of Section 6.3.

a) Consider the differential equation

$$\alpha'(t) = (P_{24} + P_{34})\alpha(t),$$

where $\alpha: \mathbb{R} \rightarrow \mathrm{GL}_4(\mathbb{R})$. Solve this to obtain a one-parameter subgroup α in $\mathrm{GL}_4(\mathbb{R})$ and show that $\alpha(t) \in \mathrm{Lor}(3, 1)$ for all $t \in \mathbb{R}$.

b) Consider the differential equation

$$\tilde{\alpha}'(t) = -i(E + F)\tilde{\alpha}(t),$$

where $\tilde{\alpha}: \mathbb{R} \rightarrow \mathrm{GL}_2(\mathbb{C})$. Solve this to obtain a one-parameter subgroup $\tilde{\alpha}$ in $\mathrm{GL}_2(\mathbb{C})$ and show that $\tilde{\alpha}(t) \in \mathrm{SL}_2(\mathbb{C})$ for all $t \in \mathbb{R}$.

c) Explain the relationship between these two curves.

6.3. **Δ Δ** For $n \geq 1$, let $\mathbb{R}^{n-1,1}$ be \mathbb{R}^n equipped with the Lorentz inner product $(,)$. Define a *Lorentz Clifford algebra* $\mathrm{Cl}_{n-1,1}$ in a similar fashion to the way Cl_n was defined in Section 5.1, except that the algebra generators e_1, \dots, e_{n-1}, e_n satisfy the relations

$$\begin{cases} e_s e_r = -e_r e_s & \text{if } s \neq r, \\ e_t^2 = -1 & \text{if } 1 \leq t \leq n-1, \\ e_n^2 = 1. \end{cases}$$

Let $j: \mathbb{R}^{n-1,1} \rightarrow \mathrm{Cl}_{n-1,1}$ be the \mathbb{R} -linear transformation embedding $\mathbb{R}^{n-1,1}$ into $\mathrm{Cl}_{n-1,1}$.

a) Show that there is an analogue for $\mathrm{Cl}_{n-1,1}$ of the Universal Property of Theorem 5.4.

b) For $n \geq 2$, define $f_{n-1,1}: \mathbb{R}^{n-1,1} \rightarrow \mathrm{M}_2(\mathrm{Cl}_{n-2})$ by

$$f_{n-1,1}\left(\sum_{1 \leq r \leq n} a_r e_r\right) = \begin{bmatrix} \sum_{1 \leq r \leq n-2} a_r e_r + a_n & -a_{n-1} \\ a_{n-1} & -\sum_{1 \leq r \leq n-2} a_r e_r - a_n \end{bmatrix}.$$

Using the Universal Property of (a), show that $f_{n-1,1}$ induces an isomorphism of \mathbb{R} -algebras

$$F_{n-1,1}: \mathrm{Cl}_{n-1,1} \rightarrow \mathrm{M}_2(\mathrm{Cl}_{n-2})$$

for which $F_{n-1,1} \circ j_{n-1,1} = f_{n-1,1}$.

c) Discuss the analogue $\mathrm{Spin}(n-1, 1) \leq \mathrm{Cl}_{n-1,1}^\times$ of the spinor group $\mathrm{Spin}(n) \leq \mathrm{Cl}_n^\times$. In particular, deduce that there is a connected subgroup $\mathrm{Spin}^0(n-1, 1) \leq \mathrm{Spin}(n-1, 1)$ and a double covering homomorphism

$$\mathrm{Spin}^0(n-1, 1) \rightarrow \mathrm{Lor}(n-1, 1).$$

- 6.4. $\Delta \Delta$ a) Using the previous question in the case where $n = 4$, deduce that there is an isomorphism of \mathbb{R} -algebras

$$\mathrm{Cl}_{3,1} \longrightarrow M_2(\mathbb{H}).$$

Using the regular representation $M_2(\mathbb{H}) \rightarrow M_4(\mathbb{C})$, this provides an action of $\mathrm{Cl}_{3,1}$ by \mathbb{C} -linear transformations on the \mathbb{C} -vector space \mathbb{C}^4 . Identifying $SL_2(\mathbb{C})$ with $\mathrm{Spin}^0(3, 1)$, this gives a continuous \mathbb{C} -linear action of $SL_2(\mathbb{C})$ on \mathbb{C}^4 .

- b) Let $\varphi: \mathbb{R}^{3,1} \rightarrow \mathbb{C}^4$ be a smooth function with

$$\varphi = (\varphi_1, \varphi_2, \varphi_3, \varphi_4).$$

Show that the *D'Alembertian operator* \square^2 given by

$$\square^2 \varphi = \frac{\partial^2 \varphi_1}{\partial x_1^2} + \frac{\partial^2 \varphi_2}{\partial x_2^2} + \frac{\partial^2 \varphi_3}{\partial x_3^2} - \frac{\partial^2 \varphi_4}{\partial t^2}$$

can be expressed as

$$\square^2 = \left(e_1 \frac{\partial}{\partial x_1} + e_2 \frac{\partial}{\partial x_2} + e_3 \frac{\partial}{\partial x_3} + e_4 \frac{\partial}{\partial t} \right)^2,$$

where the terms inside the brackets are to be thought of as differential operators with coefficients in $\mathrm{Cl}_{3,1}$ which act on \mathbb{C}^4 via the matrix representation defined in (a).

[This was originally observed by the physicist Dirac, who used it to factorise a relativistic version of the Schrödinger equation in order to study the electron. The connection with Clifford algebras seems to have been observed by Atiyah, Bott & Shapiro in their work on Index Theory [3].]

Part II

Matrix Groups as Lie Groups

Lie Groups

In this chapter we will introduce some of the basic ideas of *smooth manifolds* and then define *Lie groups*. Full details can be found in books on Differential Geometry such as [6, 8, 29] while [7, 23] contain briefer introductions. One of our main aims is to prove that every matrix subgroup of $\mathrm{GL}_n(\mathbb{R})$ is a *Lie subgroup* and we follow the proof of this result described in Howe [12]. We will also show that not every Lie group is a matrix group by exhibiting the simplest counterexample. In fact every *compact* Lie group is a matrix group, but the proof of this requires the theory of *Haar measure* on a Lie group; we will discuss compact Lie groups in Chapter 10.

7.1 Smooth Manifolds

The concept of a *smooth manifold* is fundamental in much of Mathematics and Physics and provides a natural setting for many geometric concepts, suitable for calculus and analysis involving global topological aspects.

Definition 7.1

A continuous map $g: V_1 \rightarrow V_2$ where each $V_k \subseteq \mathbb{R}^{m_k}$ is open is *smooth* if it is infinitely differentiable. A smooth bijection g is a *diffeomorphism* if its inverse $g^{-1}: V_2 \rightarrow V_1$ is also smooth.

We will require the following topological notion.

Definition 7.2

A topological space X is *separable* if it has a countable open covering, i.e., an open covering of the form $\{U_j\}_{j=1}^{\infty} = \{U_1, U_2, U_3, \dots\}$.

Every compact topological space is separable, as is any subspace of \mathbb{R}^n or \mathbb{C}^n for $n \geq 1$, so all matrix groups are separable. A *quotient space* (in the sense to be introduced in Section 8.1) of a separable space is also separable.

From now on, let M be a separable Hausdorff topological space.

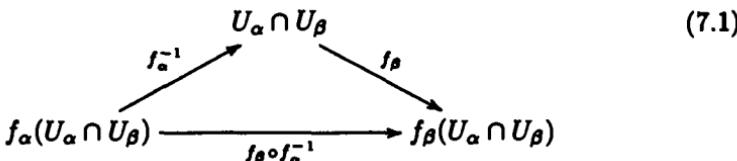
Definition 7.3

If $U \subseteq M$ and $V \subseteq \mathbb{R}^n$ are open subsets, a homeomorphism $f: U \rightarrow V$ is called an *n-chart* for U .

If $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ is an open covering of M and $\mathcal{F} = \{f_\alpha: U_\alpha \rightarrow V_\alpha\}$ is a collection of charts, then \mathcal{F} is called an *atlas* for M if, whenever $U_\alpha \cap U_\beta \neq \emptyset$,

$$f_\beta \circ f_\alpha^{-1}: f_\alpha(U_\alpha \cap U_\beta) \rightarrow f_\beta(U_\alpha \cap U_\beta)$$

is a diffeomorphism.



Sometimes we will 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 7.4

Let $(M, \mathcal{U}, \mathcal{F})$ and $(M', \mathcal{U}', \mathcal{F}')$ be atlases on topological spaces M and M' . A *smooth map* $h: (M, \mathcal{U}, \mathcal{F}) \rightarrow (M', \mathcal{U}', \mathcal{F}')$ is a continuous map $h: M \rightarrow M'$ such that for each pair α, α' with $h(U_\alpha) \cap U'_{\alpha'} \neq \emptyset$, the composite

$$f'_{\alpha'} \circ h \circ f_\alpha^{-1}: f_\alpha(h^{-1}U'_{\alpha'}) \rightarrow V'_{\alpha'}$$

is smooth.

$$\begin{array}{ccc}
 f_\alpha(h^{-1}U'_{\alpha'}) & \xrightarrow{f'_{\alpha'} \circ h \circ f_\alpha^{-1}} & V'_{\alpha'} \\
 \downarrow f_\alpha^{-1} & & \downarrow f'_{\alpha'}^{-1} \\
 h^{-1}U'_{\alpha'} & \xrightarrow{h} & h(U_\alpha) \cap U'_{\alpha'}
 \end{array} \tag{7.2}$$

7.2 Tangent Spaces and Derivatives

Let $(M, \mathcal{U}, \mathcal{F})$ be a smooth n -manifold and $p \in M$. Let $\gamma: (a, b) \rightarrow M$ be a continuous curve with $a < 0 < b$.

Definition 7.5

γ is *differentiable* at $t \in (a, b)$ if for every chart $f: U \rightarrow V$ with $\gamma(t) \in U$, the curve $f \circ \gamma: (a, b) \rightarrow V$ 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 $f \circ \gamma$ exist 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) .

From now on we will often write $fg = f \circ g$ for the composition of functions f and g when no confusion seems likely to result.

Lemma 7.6

Let $f_0: U_0 \rightarrow V_0$ be a chart with $\gamma(t) \in U_0$ and suppose that

$$f_0 \circ \gamma: (a, b) \cap f_0^{-1}V_0 \rightarrow V_0$$

is differentiable (respectively smooth) at t . Then for any chart $f: U \rightarrow V$ with $\gamma(t) \in U$,

$$f \circ \gamma: (a, b) \cap f^{-1}V \rightarrow V$$

is differentiable (respectively smooth) at t .

Proof

This follows using the ideas of Definition 7.3. The composition $f \circ \gamma$ is defined on a subinterval of (a, b) containing t and is smooth. There is the usual *chain rule* or *function of a function rule* for the derivative

$$(f \circ \gamma)'(t) = \text{Jac}_{f \circ \gamma}(f_0 \circ \gamma(t)) (f_0 \circ \gamma)'(t). \tag{7.3}$$

Here, for a differentiable function

$$h: W_1 \rightarrow W_2; \quad h(\mathbf{x}) = \begin{bmatrix} h_1(\mathbf{x}) \\ \vdots \\ h_{m_2}(\mathbf{x}) \end{bmatrix}$$

where $W_1 \subseteq \mathbb{R}^{m_1}$ and $W_2 \subseteq \mathbb{R}^{m_2}$ are open subsets, $\mathbf{x} \in W_1$ and

$$\text{Jac}_h(\mathbf{x}) = \left[\frac{\partial h_i}{\partial x_j}(\mathbf{x}) \right] \in M_{m_2, m_1}(\mathbb{R})$$

is the *Jacobian matrix* of h at \mathbf{x} . □

If $\gamma(0) = p$ and γ is differentiable at 0, then for any (and hence every) chart $f_0: U_0 \rightarrow V_0$ 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 $f: U \rightarrow V$ with $\gamma(0) \in U$ by Equation (7.3) we have

$$(f\gamma)'(0) = \text{Jac}_{f f_0^{-1}}(f_0 \gamma(0))(f_0 \gamma)'(0).$$

In order to define the notion of the *tangent space* $T_p M$ to the manifold M at p , we consider all pairs of the form

$$((f\gamma)'(0), f: U \rightarrow V)$$

where $\gamma(0) = p \in U$, and then impose an equivalence relation \sim under which

$$((f_1 \gamma)'(0), f_1: U_1 \rightarrow V_1) \sim ((f_2 \gamma)'(0), f_2: U_2 \rightarrow V_2).$$

Since

$$(f_2 \gamma)'(0) = \text{Jac}_{f_2 f_1^{-1}}(f_1 \gamma(0))(f_1 \gamma)'(0),$$

we can also write this as

$$(\mathbf{v}, f_1: U_1 \rightarrow V_1) \sim (\text{Jac}_{f_2 f_1^{-1}}(f_1(p))\mathbf{v}, f_2: U_2 \rightarrow V_2),$$

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 $T_p M$ and we will sometimes denote the equivalence class of $(\mathbf{v}, f: U \rightarrow V)$ by $[\mathbf{v}, f: U \rightarrow V]$.

Proposition 7.7

For $p \in M$, $T_p M$ is an \mathbb{R} -vector space of dimension n .

Proof

For any chart $f: U \rightarrow V$ with $p \in U$, we can identify the elements of $T_p M$ with objects of the form $(v, f: U \rightarrow V)$. Every vector $v \in \mathbb{R}^n$ arises as the derivative of a curve $\bar{\gamma}: (-\epsilon, \epsilon) \rightarrow V$ for which $\bar{\gamma}(0) = f(p)$. For example, for small enough ϵ , we could take

$$\bar{\gamma}(t) = f(p) + tv.$$

There is an associated curve in M ,

$$\gamma: (-\epsilon, \epsilon) \rightarrow M; \quad \gamma(t) = f^{-1}\bar{\gamma}(t),$$

for which $\gamma(0) = p$. So using such a chart we can identify $T_p M$ with \mathbb{R}^n by

$$[v, f: U \rightarrow V] \longleftrightarrow v.$$

The same argument as used to prove Proposition 3.16 shows that $T_p M$ is a vector space and that the above correspondence is a linear isomorphism. \square

Let $h: (M, \mathcal{U}, \mathcal{F}) \rightarrow (M', \mathcal{U}', \mathcal{F}')$ be a smooth map between manifolds of dimensions n, n' . We will use the notation of Definition 7.4. For $p \in M$, consider a pair of charts as in Diagram (7.2) 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 $\text{Jac}_{h_{\alpha', \alpha}}(f_\alpha(p))$ has an associated \mathbb{R} -linear transformation

$$dh_{\alpha', \alpha}: \mathbb{R}^n \rightarrow \mathbb{R}^{n'}; \quad dh_{\alpha', \alpha}(x) = \text{Jac}_{h_{\alpha', \alpha}}(f_\alpha(p))x.$$

It can be verified that this passes to equivalence classes to give a well-defined \mathbb{R} -linear transformation

$$dh_p: T_p M \rightarrow T_{h(p)} M'.$$

The following result summarises the properties of the derivative and should be compared with Proposition 3.20.

Proposition 7.8

Let $h: (M, \mathcal{U}, \mathcal{F}) \rightarrow (M', \mathcal{U}', \mathcal{F}')$ and $g: (M', \mathcal{U}', \mathcal{F}') \rightarrow (M'', \mathcal{U}'', \mathcal{F}'')$ be smooth maps between manifolds M, M', M'' of dimensions n, n', n'' . Then for each $p \in M$ we have

- i) $d g_{h(p)} \circ d h_p = d(g \circ h)_p;$
- ii) $d \text{Id}_p = \text{Id}_{T_p M}$, where $\text{Id}: M \rightarrow M$ is the identity map.

Definition 7.9

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 $f: U \rightarrow V$ 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

$$f_0: N \cap U \rightarrow V \cap \mathbb{R}^k; \quad f_0(x) = f(x).$$

We will denote this manifold by $(N, \mathcal{U}_N, \mathcal{F}_N)$. The following result is immediate.

Proposition 7.10

For a submanifold $N \subseteq M$ of dimension k , the inclusion function $\text{incl}: N \rightarrow M$ is smooth and for every $p \in N$, $d \text{incl}_p: T_p N \rightarrow T_p M$ is an injection.

The next result allows us to recognise submanifolds as inverse images of points under smooth mappings.

Theorem 7.11 (Implicit Function Theorem for manifolds)

Let $h: (M, \mathcal{U}, \mathcal{F}) \rightarrow (M', \mathcal{U}', \mathcal{F}')$ be a smooth map between manifolds of dimensions n, n' . Suppose that for some $q \in M'$, $d h_p: T_p M \rightarrow T_{h(p)} M'$ 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 $T_p N = \ker d h_p$.

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 7.12 (Inverse Function Theorem for manifolds)

Let $h: (M, \mathcal{U}, \mathcal{F}) \rightarrow (M', \mathcal{U}', \mathcal{F}')$ be a smooth map between manifolds of dimensions n, n' . Suppose that for some $p \in M$, $d h_p: T_p M \rightarrow T_{h(p)} M'$ 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 $h_1: U \rightarrow V$ is a diffeomorphism. In particular, the derivative $d h_p: T_p \rightarrow T_{h(p)}$ is an \mathbb{R} -linear isomorphism and $n = n'$.

When this occurs we say that h is *locally a diffeomorphism at p*.

Example 7.13

Consider the exponential function $\exp: M_n(\mathbb{R}) \rightarrow GL_n(\mathbb{R})$. Then by Proposition 2.5, $d \exp_O(X) = X$. Hence \exp is locally a diffeomorphism at O .

7.3 Lie Groups

The following should be compared with Definition 1.15.

Definition 7.14

Let G be a smooth manifold which is also a topological group with multiplication map $\text{mult}: G \times G \rightarrow G$ and inverse map $\text{inv}: G \rightarrow G$ and view $G \times G$ as the product manifold. Then G is a *Lie group* if mult, inv are smooth maps.

Definition 7.15

Let G be a Lie group. A closed subgroup $H \leq G$ 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 , and an element $g \in G$ there is a tangent space $T_g G$ and when G is a matrix group this agrees with the tangent space defined in Chapter 3. We will use the notation $\mathfrak{g} = T_1 G$ for the tangent space at the identity of G . A smooth homomorphism of Lie groups $G \rightarrow H$ has the properties of a Lie homomorphism in the sense of Definition 3.19.

For a Lie group G , let $g \in G$. The following three functions are of particular importance:

$$\begin{aligned} L_g: G &\rightarrow G; \quad L_g(x) = gx, && (\text{Left multiplication}) \\ R_g: G &\rightarrow G; \quad R_g(x) = xg, && (\text{Right multiplication}) \\ \chi_g: G &\rightarrow G; \quad \chi_g(x) = gxg^{-1}. && (\text{Conjugation}) \end{aligned}$$

Proposition 7.16

For each $g \in G$, the maps L_g, R_g, χ_g are diffeomorphisms with inverses

$$L_g^{-1} = L_{g^{-1}}, \quad R_g^{-1} = R_{g^{-1}}, \quad \chi_g^{-1} = \chi_{g^{-1}}.$$

Proof

Charts for $G \times G$ have the form

$$\varphi_1 \times \varphi_2: U_1 \times U_2 \longrightarrow V_1 \times V_2,$$

where $\varphi_k: U_k \longrightarrow V_k$ are charts for G . Now suppose that $\mu: U_1 \times U_2 \subseteq W \subseteq G$ where there is a chart $\theta: W \longrightarrow Z$. By assumption, the composition

$$\theta \circ \mu \circ (\varphi_1 \times \varphi_2)^{-1} = \theta \circ \mu \circ (\varphi_1^{-1} \times \varphi_2^{-1}): V_1 \times V_2 \longrightarrow Z$$

is smooth. Then $L_g(x) = \mu(g, x)$, so if $g \in U_1$ and $x \in U_2$, we have

$$L_g(x) = \theta^{-1} \circ (\theta \circ L_g \circ \varphi_2^{-1}) \circ \varphi_2(x).$$

But then it is clear that

$$\theta \circ \varphi_2^{-1}: V_2 \longrightarrow Z$$

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

$$\chi_g = L_g \circ R_g = R_g \circ L_g,$$

and a composite of smooth maps is smooth. \square

The derivatives of these maps at the identity $1 \in G$ are worth studying. Since L_g and R_g are diffeomorphisms with inverses $L_{g^{-1}}$ and $R_{g^{-1}}$, the derivatives

$$d(L_g)_1, d(R_g)_1: \mathfrak{g} = T_1 G \longrightarrow T_g G$$

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

$$\text{Ad}_g = d(\chi_g)_1: \mathfrak{g} \longrightarrow \mathfrak{g}.$$

This is the *adjoint action* of $g \in G$ on \mathfrak{g} . For G a matrix group this agrees with the adjoint action defined in Chapter 3. 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 3.21 is true.

Theorem 7.17

Let G, H be Lie groups and $\varphi: G \rightarrow H$ be a Lie homomorphism. Then the derivative is a homomorphism of Lie algebras. In particular, if $G \leq H$ is a Lie subgroup, the inclusion map $\text{incl}: G \rightarrow H$ induces an injection of Lie algebras $d\text{incl}: \mathfrak{g} \rightarrow \mathfrak{h}$.

7.4 Some Examples of Lie Groups**Example 7.18**

When $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , $\text{GL}_n(\mathbf{k})$ is a Lie group.

Proof

This follows from Proposition 1.14(i) which implies that $\text{GL}_n(\mathbf{k}) \subseteq M_n(\mathbf{k})$ is an open subset, where as usual we identify $M_n(\mathbf{k})$ with \mathbf{k}^{n^2} . For charts we take the open sets $U \subseteq \text{GL}_n(\mathbf{k})$ and the identity function $\text{Id}: U \rightarrow U$. The tangent space at each point $A \in \text{GL}_n(\mathbf{k})$ is just $M_n(\mathbf{k})$. So the notions of tangent space and dimension of Sections 7.1, 7.2 and Chapter 3 agree here. The multiplication and inverse maps are obviously smooth as they are defined by polynomial and rational functions between open subsets of $M_n(\mathbf{k})$. \square

Example 7.19

For $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , $\text{SL}_n(\mathbf{k}) \leq \text{GL}_n(\mathbf{k})$ is a Lie subgroup.

Proof

Following Proposition 1.14(ii), we have

$$\text{SL}_n(\mathbf{k}) = \det^{-1} 1 \leq \text{GL}_n(\mathbf{k}),$$

where $\det: \text{GL}_n(\mathbf{k}) \rightarrow \mathbf{k}$ is smooth. \mathbf{k} is a smooth manifold of dimension $\dim_{\mathbb{R}} \mathbf{k}$ with tangent space the \mathbf{k} -vector space $T_r \mathbf{k} = \mathbf{k}$ at each $r \in \mathbf{k}$. In order to apply Theorem 7.11, first we must show that the derivative $d\det_A: M_n(\mathbf{k}) \rightarrow \mathbf{k}$ is surjective for every $A \in \text{GL}_n(\mathbf{k})$. To do this, consider a smooth curve $\alpha: (-\varepsilon, \varepsilon) \rightarrow \text{GL}_n(\mathbf{k})$ with $\alpha(0) = A$. The derivative $d\det_A$ applied to $\alpha'(0)$ can be found using the formula

$$d\det_A(\alpha'(0)) = \frac{d\det \alpha(t)}{dt} \Big|_{t=0}.$$

The modified curve

$$\alpha_0: (-\varepsilon, \varepsilon) \rightarrow \mathrm{GL}_n(\mathbf{k}); \quad \alpha_0(t) = A^{-1}\alpha(t)$$

satisfies $\alpha_0(0) = I$ and Lemma 3.22 implies that

$$d \det_I(\alpha'_0(0)) = \frac{d \det \alpha_0(t)}{dt} \Big|_{t=0} = \mathrm{tr} \alpha'_0(0).$$

Hence

$$d \det_A(\alpha'(0)) = \frac{d \det(A\alpha_0(t))}{dt} \Big|_{t=0} = \det A \frac{d \det(\alpha_0(t))}{dt} \Big|_{t=0} = \det A \mathrm{tr} \alpha'_0(0).$$

So $d \det_A$ is the \mathbf{k} -linear transformation

$$d \det_A: M_n(\mathbf{k}) \rightarrow \mathbf{k}; \quad d \det_A(X) = (\det A) \mathrm{tr}(A^{-1}X).$$

The kernel of $d \det_A$ is $\ker d \det_A = A \mathrm{Sl}_n(\mathbf{k})$ and it is also surjective since tr is. In particular this is true for $A \in \mathrm{SL}_n(\mathbf{k})$. By Theorem 7.11, $\mathrm{SL}_n(\mathbf{k}) \rightarrow \mathrm{GL}_n(\mathbf{k})$ is a submanifold and so is a Lie subgroup. Again the two notions of tangent space and dimension agree. \square

There are some useful general principles behind this last proof. Although we state the following two results for matrix groups, it is worth noting that they also apply when $\mathrm{GL}_n(\mathbb{R})$ is replaced by an arbitrary Lie group.

Proposition 7.20 (Left Translation Trick)

Let $F: \mathrm{GL}_n(\mathbb{R}) \rightarrow M$ be a smooth function into a smooth manifold M 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 dF_A surjective. Then dF_{BA} is surjective.

Proof

Left multiplication by $B \in G$, $L_B: \mathrm{GL}_n(\mathbb{R}) \rightarrow \mathrm{GL}_n(\mathbb{R})$, is a diffeomorphism and its derivative at $A \in \mathrm{GL}_n(\mathbb{R})$ is

$$d(L_B): M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R}); \quad dL_B(X) = BX.$$

By assumption, $F \circ L_B = F$ as a function on $\mathrm{GL}_n(\mathbb{R})$. Then

$$\begin{aligned} dF_{BA}(X) &= dF_{BA}(B(B^{-1}X)) \\ &= dF_{BA} \circ d(L_B)_A(B^{-1}X) \\ &= d(F \circ L_B)_A(B^{-1}X) \\ &= dF_A(B^{-1}X). \end{aligned}$$

Since left multiplication by B^{-1} on $M_n(\mathbb{R})$ is surjective, the result follows. \square

Proposition 7.21 (Identity Check Trick)

Let $G \leq \mathrm{GL}_n(\mathbb{R})$ be a matrix subgroup and M be a smooth manifold. Let $F: \mathrm{GL}_n(\mathbb{R}) \rightarrow M$ be 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 dF_I is surjective then dF_A is surjective for all $A \in G$ and $\ker dF_A = Ag$.

Example 7.22

$O(n)$ is a Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$.

Proof

Recall from Chapter 3 that we can specify $O(n) \subseteq \mathrm{GL}_n(\mathbb{R})$ 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 equations in the entries of the matrix $A = [a_{ij}]$ are sufficient:

$$\sum_{k=1}^n a_{kr}^2 - 1 = 0 \quad (1 \leq r \leq n), \quad \sum_{k=1}^n a_{kr} a_{ks} = 0 \quad (1 \leq r < s \leq n).$$

Notice that there are

$$n + \binom{n}{2} = \binom{n+1}{2}$$

of these equations. We can combine the left-hand sides of these in some order to give a function $F: \mathrm{GL}_n(\mathbb{R}) \rightarrow \mathbb{R}^{(\frac{n(n+1)}{2})}$, 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 $dF_A: M_n(\mathbb{R}) \rightarrow \mathbb{R}^{(\frac{n(n+1)}{2})}$.

By the Identity Check Trick 7.21, to show that dF_A is surjective for all $A \in O(n)$, 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

$$dF_I = \begin{bmatrix} 2 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 2 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & & & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\ \vdots & & \ddots & & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 & 0 \end{bmatrix}$$

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 dF_I is surjective. It is also true that

$$\ker dF_I = \text{Sk-Sym}_n(\mathbb{R}) = \mathfrak{o}(n).$$

Hence $O(n) \leq GL_n(\mathbb{R})$ is a Lie subgroup and at each element, the tangent space and dimension agree with those obtained using the definitions of Chapter 3. \square

This example is typical of what happens for any matrix group that is a Lie subgroup of $GL_n(\mathbb{R})$. We summarise the situation in the following theorem, whose proof involves a careful comparison between the ideas introduced in Chapter 3 and the definitions involving manifolds.

Theorem 7.23

Let $G \leq GL_n(\mathbb{R})$ 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, $T_A G = Ag$.

In the following sections, our goal will be to prove an important result.

Theorem 7.24

Let $G \leq GL_n(\mathbb{R})$ be a matrix subgroup. Then G is a Lie subgroup of $GL_n(\mathbb{R})$.

In fact a more general result also holds but we will not give a proof.

Theorem 7.25

Let $G \leq H$ be a closed subgroup of a Lie group H . Then G is a Lie subgroup of H .

7.5 Some Useful Formulae in Matrix Groups

Let $G \leq \mathrm{GL}_n(\mathbb{R})$ be a closed matrix subgroup. Applying Proposition 2.4, we may choose $r \in \mathbb{R}$ so that $0 < r \leq 1/2$ and if $A, B \in N_{M_n(\mathbb{R})}(O; r)$ then $\exp(A)\exp(B) \in \exp(N_{M_n(\mathbb{R})}(O; 1/2))$. Since \exp is injective on $N_{M_n(\mathbb{R})}(O; r)$, there is a unique $C \in M_n(\mathbb{R})$ for which

$$\exp(A)\exp(B) = \exp(C). \quad (7.4)$$

We also set

$$S = C - A - B - \frac{1}{2}[A, B] \in M_n(\mathbb{R}). \quad (7.5)$$

For $X \in M_n(\mathbb{R})$ we have

$$\exp(X) = I + X + R_1(X),$$

where the remainder term $R_1(X)$ is given by

$$R_1(X) = \sum_{k \geq 2} \frac{1}{k!} X^k.$$

Hence,

$$\|R_1(X)\| \leq \|X\|^2 \sum_{k \geq 2} \frac{1}{k!} \|X\|^{k-2},$$

and therefore if $\|X\| \leq 1$,

$$\|R_1(X)\| \leq \|X\|^2 \left(\sum_{k \geq 2} \frac{1}{k!} \right) = \|X\|^2(e - 2) < \|X\|^2.$$

Since $\|C\| < \frac{1}{2}$ we obtain

$$\|R_1(C)\| < \|C\|^2. \quad (7.6)$$

Similar considerations lead to

$$\exp(C) = \exp(A)\exp(B) = I + A + B + R_1(A, B),$$

where

$$R_1(A, B) = \sum_{k \geq 2} \frac{1}{k!} \left(\sum_{r=0}^k \binom{k}{r} A^r B^{k-r} \right).$$

This gives

$$\begin{aligned} \|R_1(A, B)\| &\leq \sum_{k \geq 2} \frac{1}{k!} \left(\sum_{r=0}^k \binom{k}{r} \|A\|^r \|B\|^{k-r} \right) \\ &= \sum_{k \geq 2} \frac{(\|A\| + \|B\|)^k}{k!} \\ &= (\|A\| + \|B\|)^2 \sum_{k \geq 2} \frac{(\|A\| + \|B\|)^{k-2}}{k!} \\ &\leq (\|A\| + \|B\|)^2 \end{aligned}$$

since $\|A\| + \|B\| < 1$.

Combining the two ways of writing $\exp(C)$ from above, we have

$$C = A + B + R_1(A, B) - R_1(C) \quad (7.7)$$

and so

$$\begin{aligned} \|C\| &\leq \|A\| + \|B\| + \|R_1(A, B)\| + \|R_1(C)\| \\ &< \|A\| + \|B\| + (\|A\| + \|B\|)^2 + \|C\|^2 \\ &\leq 2(\|A\| + \|B\|) + \frac{1}{2}\|C\|^2, \end{aligned}$$

since $\|A\|, \|B\|, \|C\| \leq \frac{1}{2}$. Finally this gives

$$\|C\| \leq 4(\|A\| + \|B\|).$$

Equation (7.7) also gives

$$\begin{aligned} \|C - A - B\| &\leq \|R_1(A, B)\| + \|R_1(C)\| \\ &\leq (\|A\| + \|B\|)^2 + (4(\|A\| + \|B\|))^2, \end{aligned}$$

giving

$$\|C - A - B\| \leq 17(\|A\| + \|B\|)^2. \quad (7.8)$$

Now we will refine these estimates further. Write

$$\exp(C) = I + C + \frac{1}{2}C^2 + R_2(C)$$

where

$$R_2(C) = \sum_{k \geq 3} \frac{1}{k!} C^k$$

which satisfies the estimate

$$\|R_2(C)\| \leq \frac{1}{3} \|C\|^3$$

since $\|C\| \leq 1$. With the aid of Equation (7.5) we obtain

$$\begin{aligned} \exp(C) &= I + A + B + \frac{1}{2}[A, B] + S + \frac{1}{2}C^2 + R_2(C) \\ &= I + A + B + \frac{1}{2}[A, B] + \frac{1}{2}(A + B)^2 + T \\ &= I + A + B + \frac{1}{2}(A^2 + 2AB + B^2) + T, \end{aligned} \quad (7.9)$$

where

$$T = S + \frac{1}{2}(C^2 - (A + B)^2) + R_2(C). \quad (7.10)$$

Also,

$$\exp(A) \exp(B) = I + A + B + \frac{1}{2}(A^2 + 2AB + B^2) + R_2(A, B) \quad (7.11)$$

where

$$R_2(A, B) = \sum_{k \geq 3} \frac{1}{k!} \left(\sum_{r=0}^k \binom{k}{r} A^r B^{k-r} \right),$$

which satisfies

$$\|R_2(A, B)\| \leq \frac{1}{3} (\|A\| + \|B\|)^3$$

since $\|A\| + \|B\| \leq 1$.

Comparing Equations (7.9) and (7.11) and using (7.4) we see that

$$S = R_2(A, B) + \frac{1}{2}((A + B)^2 - C^2) - R_2(C).$$

Taking norms we have

$$\begin{aligned} \|S\| &\leq \|R_2(A, B)\| + \frac{1}{2}\|(A + B)(A + B - C) - (A + B - C)C\| + \|R_2(C)\| \\ &\leq \frac{1}{3}(\|A\| + \|B\|)^3 + \frac{1}{2}(\|A\| + \|B\| + \|C\|)\|A + B - C\| + \frac{1}{3}\|C\|^3 \\ &\leq \frac{1}{3}(\|A\| + \|B\|)^3 + \frac{5}{2}(\|A\| + \|B\|) \cdot 17(\|A\| + \|B\|)^2 + \frac{1}{3}(4\|A\| + \|B\|)^3 \\ &\leq 65(\|A\| + \|B\|)^3, \end{aligned}$$

which yields the estimate

$$\|S\| \leq 65(\|A\| + \|B\|)^3. \quad (7.12)$$

Theorem 7.26

For $U, V \in M_n(\mathbb{R})$ we have the following formulæ.

Trotter Product Formula:

$$\exp(U + V) = \lim_{r \rightarrow \infty} (\exp((1/r)U) \exp((1/r)V))^r.$$

Commutator Formula:

$$\exp([U, V]) = \lim_{r \rightarrow \infty} (\exp((1/r)U) \exp((1/r)V) \exp(-(1/r)U) \exp(-(1/r)V))^{r^2}.$$

Proof

For large r we may take $A = (1/r)U$ and $B = (1/r)V$ and apply Equation (7.5) to give

$$\exp((1/r)U) \exp((1/r)V) = \exp(C_r)$$

with

$$\|C_r - (1/r)(U + V)\| \leq \frac{17(\|U\| + \|V\|)^2}{r^2}.$$

As $r \rightarrow \infty$,

$$\|rC_r - (U + V)\| = \frac{17(\|U\| + \|V\|)^2}{r} \rightarrow 0,$$

Hence $rC_r \rightarrow (U + V)$. 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

$$\|S_r\| \leq 65 \frac{(\|U\| + \|V\|)^3}{r^3}.$$

Similarly, replacing U, V with $-U, -V$ we obtain

$$\exp((-1/r)U) \exp((-1/r)V) = \exp(C'_r),$$

where

$$C'_r = \frac{1}{r}(U + V) + \frac{1}{2r^2}[U, V] + S'_r$$

and

$$\|S'_r\| \leq 65 \frac{(\|U\| + \|V\|)^3}{r^3}.$$

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

$$\begin{aligned} E_r &= C_r + C'_r + \frac{1}{2}[C_r, C'_r] + T_r \\ &= \frac{1}{r^2} + \frac{1}{2}[C_r, C'_r] + S_r + S'_r + T_r. \end{aligned} \quad (7.13)$$

Here T_r is defined from Equation (7.5) by setting $C_r = A$, $C'_r = B$ and $T_r = S$.

Another tedious computation shows that

$$\begin{aligned} [C_r, C'_r] &= \left[\frac{1}{r}(U + V) + \frac{1}{2r^2}[U, V] + S_r, \frac{-1}{r}(U + V) + \frac{1}{2r^2}[U, V] + S'_r \right] \\ &= \frac{1}{r^3}[U + V, [U, V]] + \frac{1}{r}[U + V, S_r + S'_r] \\ &\quad + \frac{1}{2r^2}[[U, V], S'_r - S_r] + [S_r, S'_r]. \end{aligned}$$

By the estimate of (7.12), all four of these terms has norm bounded by an expression of the form $(\text{constant})/r^3$, so the same is true of $[C_r, C'_r]$. Estimate (7.12) also implies that S_r, S'_r, T_r have similarly bounded norms. Setting

$$Q_r = r^2 E_r - [U, V],$$

we obtain

$$\|Q_r\| = \left\| E_r - \frac{1}{r^2}[U, V] \right\| \leq \frac{(\text{constant})}{r^3} \rightarrow 0$$

as $r \rightarrow \infty$, so

$$\exp(E_r)^{r^2} = \exp([U, V] + Q_r) \rightarrow \exp([U, V]).$$

The commutator formula now follows using continuity of \exp . □

Proof (Another proof of Lemma 3.23)

As an application of the Trotter formula, we will reprove the formula of Lemma 3.23:

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

The case $n = 1$ is immediate, so assume that $n > 1$.

If $U, V \in M_n(\mathbb{C})$ then by the Trotter formula together with the fact that \det is continuous and multiplicative,

$$\begin{aligned}\det \exp(U + V) &= \det \left(\lim_{r \rightarrow \infty} (\exp((1/r)U) \exp((1/r)V))^r \right) \\ &= \lim_{r \rightarrow \infty} \det (\exp((1/r)U) \exp((1/r)V))^r \\ &= \lim_{r \rightarrow \infty} \det \exp((1/r)U)^r \det \exp((1/r)V)^r \\ &= \lim_{r \rightarrow \infty} \det (\exp((1/r)U)^r) \det (\exp((1/r)V)^r) \\ &= \lim_{r \rightarrow \infty} \det \exp(U) \det \exp(V) \\ &= \det \exp(U) \det \exp(V).\end{aligned}$$

More generally, given $U_1, \dots, U_k \in M_n(\mathbb{C})$ we have

$$\det \exp(U_1 + \cdots + U_k) = \det \exp(U_1) \cdots \det \exp(U_k). \quad (7.14)$$

So if $A = A_1 + \cdots + A_k$ where the A_j satisfy

$$\det \exp(A_j) = \exp(\operatorname{tr} A_j) \quad (j = 1, \dots, k),$$

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

$$A = \sum_{\substack{1 \leq r \leq n, \\ 1 \leq s \leq n}} a_{rs} E^{rs},$$

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 $z \in \mathbb{C}$,

$$\begin{aligned}\det \exp(zE^{rs}) &= \det \left(\sum_{k \geq 0} \frac{1}{k!} (z^k (E^{rs})^k) \right) \\ &= \begin{cases} \det((e^z - 1)E^{rr} + I_n) & \text{if } r = s, \\ \det I_n & \text{if } r \neq s \end{cases} \\ &= \begin{cases} e^z & \text{if } r = s, \\ 1 & \text{if } r \neq s. \end{cases}\end{aligned}$$

On the other hand,

$$\operatorname{tr} z E^{rs} = \begin{cases} z & \text{if } r = s, \\ 0 & \text{if } r \neq s. \end{cases}$$

Thus

$$\exp(\operatorname{tr} z E^{rs}) = \begin{cases} e^z & \text{if } r = s, \\ 1 & \text{if } r \neq s, \end{cases}$$

so we obtain

$$\det \exp(z E^{rs}) = \exp(\operatorname{tr} z E^{rs}),$$

which is the desired equation. \square

7.6 Matrix Groups are Lie Groups

Our aim in this section is to prove Theorem 7.24. Let $G \leq \mathrm{GL}_n(\mathbb{R})$ be a matrix subgroup. Recall that the Lie algebra $\mathfrak{g} = T_I G$ is an \mathbb{R} -Lie subalgebra of $\mathfrak{gl}_n(\mathbb{R}) = M_n(\mathbb{R})$. Let

$$\tilde{\mathfrak{g}} = \{A \in M_n(\mathbb{R}) : \forall t \in \mathbb{R}, \exp(tA) \in G\}.$$

Theorem 7.27

$\tilde{\mathfrak{g}}$ is an \mathbb{R} -Lie subalgebra of $M_n(\mathbb{R})$.

Proof

By definition $\tilde{\mathfrak{g}}$ is closed under multiplication by real scalars. If $U, V \in \tilde{\mathfrak{g}}$ and $r \geq 1$, 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 7.26, for $t \in \mathbb{R}$, we have

$$\exp(tU + tV) = \lim_{r \rightarrow \infty} (\exp((1/r)tU) \exp((1/r)tV))^r,$$

and

$$\exp(t[U, V]) = \exp([tU, V])$$

$$= \lim_{r \rightarrow \infty} (\exp((1/r)tU) \exp((1/r)V) \exp(-(1/r)tU) \exp(-(1/r)V))^{r^2}.$$

As these are both limits of elements of the closed subgroup $G \leq \mathrm{GL}_n(\mathbb{R})$ they are also in G . This shows that $\tilde{\mathfrak{g}}$ is a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{R}) = M_n(\mathbb{R})$. \square

Proposition 7.28

For a matrix subgroup $G \leq \mathrm{GL}_n(\mathbb{R})$, $\tilde{\mathfrak{g}}$ is an \mathbb{R} -Lie subalgebra of \mathfrak{g} .

Proof

Let $U \in \tilde{\mathfrak{g}}$. Then the curve

$$\gamma: \mathbb{R} \longrightarrow G; \quad \gamma(t) = \exp(tU),$$

has $\gamma(0) = I$ and $\gamma'(0) = U$, hence $U \in \mathfrak{g}$. \square

Remark 7.29

Eventually we will see that $\tilde{\mathfrak{g}} = \mathfrak{g}$.

We will require a technical result.

Lemma 7.30

Let $\{A_n \in \exp^{-1} G\}_{n \geq 1}$ and $\{s_n \in \mathbb{R}\}_{n \geq 1}$ be sequences for which $\|A_n\| \rightarrow 0$ and $s_n A_n \rightarrow A \in M_n(\mathbb{R})$ as $n \rightarrow \infty$. Then $A \in \tilde{\mathfrak{g}}$.

Proof

Let $t \in \mathbb{R}$. For each n , choose an integer $m_n \in \mathbb{Z}$ so that $|ts_n - m_n| \leq 1$. Then

$$\begin{aligned} \|m_n A_n - tA\| &\leq \|(m_n - ts_n)A_n\| + \|ts_n A_n - tA\| \\ &= |m_n - ts_n| \|A_n\| + \|ts_n A_n - tA\| \\ &\leq \|A_n\| + \|ts_n A_n - tA\| \rightarrow 0 \end{aligned}$$

as $n \rightarrow \infty$, showing that $m_n A_n \rightarrow tA$. 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

$$\exp(tA) = \lim_{n \rightarrow \infty} \exp(m_n A_n) \in G.$$

Thus every real scalar multiple tA is in $\exp^{-1} G$, showing that $A \in \tilde{\mathfrak{g}}$. \square

Choose a complementary \mathbb{R} -subspace \mathfrak{w} to $\tilde{\mathfrak{g}}$ in $\mathfrak{gl}_n(\mathbb{R}) = M_n(\mathbb{R})$, i.e., any vector subspace such that

$$\begin{aligned}\tilde{\mathfrak{g}} + \mathfrak{w} &= M_n(\mathbb{R}), \\ \dim_{\mathbb{R}} \tilde{\mathfrak{g}} + \dim_{\mathbb{R}} \mathfrak{w} &= \dim_{\mathbb{R}} M_n(\mathbb{R}) = n^2.\end{aligned}$$

The second of these conditions is equivalent to

$$\tilde{\mathfrak{g}} \cap \mathfrak{w} = 0.$$

This gives a direct sum decomposition of $M_n(\mathbb{R})$, so every element $X \in M_n(\mathbb{R})$ has a unique expression of the form

$$X = U + V \quad (U \in \tilde{\mathfrak{g}}, V \in \mathfrak{w}).$$

Consider the map

$$\Phi: M_n(\mathbb{R}) \longrightarrow GL_n(\mathbb{R}); \quad \Phi(U + V) = \exp(U) \exp(V) \quad (U \in \tilde{\mathfrak{g}}, V \in \mathfrak{w}).$$

Φ is a smooth function which maps 0 to I . Notice that the factor $\exp(U)$ is in G . Consider the derivative at O ,

$$d\Phi_O: M_n(\mathbb{R}) \longrightarrow \mathfrak{gl}_n(\mathbb{R}) = M_n(\mathbb{R}).$$

To determine $d\Phi_O(A + B)$, where $A \in \tilde{\mathfrak{g}}$ 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 using the notation of Equations (7.4) and (7.5) for small $t \in \mathbb{R}$, there is a unique $C(t)$ depending on t for which

$$\Phi(t(A + B)) = \exp(C(t)).$$

Estimate (7.12) gives

$$\|(C(t) - tA - tB) - \frac{t^2}{2} [A, B]\| \leq 65|t|^3(\|A\| + \|B\|)^3.$$

From this we obtain

$$\begin{aligned}\|(C(t) - tA - tB)\| &\leq \frac{t^2}{2} \|[A, B]\| + 65|t|^3(\|A\| + \|B\|)^3 \\ &= \frac{t^2}{2} (\|[A, B]\| + 130|t|(\|A\| + \|B\|)^3)\end{aligned}$$

and so

$$\frac{d}{dt} \Phi(t(A + B))|_{t=0} = \frac{d}{dt} \exp(C(t))|_{t=0} = A + B.$$

By linearity of the derivative, for small A, B ,

$$d\Phi_O(A + B) = A + B,$$

so $d\Phi_O$ is the identity function on $M_n(\mathbb{R})$. By the Inverse Function Theorem 7.12, 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 $N_{M_n(\mathbb{R})}(O; \delta)$ for some $\delta > 0$. Thus the restriction of Φ to

$$\Phi_1 : N_{M_n(\mathbb{R})}(O; \delta) \longrightarrow \Phi N_{M_n(\mathbb{R})}(O; \delta)$$

is a diffeomorphism.

Now we must show that Φ maps some open subset (which we could assume to be an open disc) of $N_{M_n(\mathbb{R})}(O; \delta) \cap \tilde{\mathfrak{g}}$ containing O onto an open neighbourhood of I in G . Suppose not; then there is a sequence of elements $U_n \in G$ with $U_n \rightarrow I$ as $n \rightarrow \infty$ but $U_n \notin \Phi \tilde{\mathfrak{g}}$. For large enough n , $U_n \in \Phi N_{M_n(\mathbb{R})}(O; \delta)$, hence there are unique elements $A_n \in \tilde{\mathfrak{g}}$ and $B_n \in \mathfrak{w}$ with $\Phi(A_n + B_n) = U_n$. Notice that $B_n \neq O$ since otherwise $U_n \in \Phi \tilde{\mathfrak{g}}$. As Φ_1 is a diffeomorphism, $A_n + B_n \rightarrow O$ and this implies that $A_n \rightarrow O$ and $B_n \rightarrow O$. 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 $\bar{B}_n = (1/\|B_n\|)B_n$ of unit norm. Each \bar{B}_n is in the unit sphere in $M_n(\mathbb{R})$, which is compact hence there is a convergent subsequence of $\{\bar{B}_n\}$. By renumbering this subsequence, we can assume that $\bar{B}_n \rightarrow B$ as $n \rightarrow \infty$, where $\|B\| = 1$. Applying Lemma 7.30 to the sequences $\{B_n\}$ and $\{1/\|B_n\|\}$ we find that $B \in \tilde{\mathfrak{g}}$. But each B_n (and hence \bar{B}_n) is in \mathfrak{w} , so B must be too. Thus $B \in \tilde{\mathfrak{g}} \cap \mathfrak{w}$, which contradicts the fact that $B \neq O$.

So there must be an open disc

$$N_{\tilde{\mathfrak{g}}}(O; \delta_1) = N_{M_n(\mathbb{R})}(O; \delta_1) \cap \tilde{\mathfrak{g}}$$

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 $GL_n(\mathbb{R})$ at I and moreover $N_{\tilde{\mathfrak{g}}}(O; \delta_1)$ is then a submanifold of $N_{M_n(\mathbb{R})}(O; \delta_1)$.

We can use left translation to move this chart to a new chart at any other point $U \in G$, by considering $L_U \circ \Phi$. The details are left as an exercise.

So we have shown that $G \leq GL_n(\mathbb{R})$ is a Lie subgroup, proving Theorem 7.24. This is a fundamental result that can be usefully reformulated as follows. The proof of the second part is similar to our proof of the first with minor adjustments required for the general case.

Theorem 7.31

A subgroup of $GL_n(\mathbb{R})$ 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.

Notice that the dimension of a matrix group G as a manifold is $\dim_{\mathbb{R}} \tilde{g}$. By Proposition 7.28, $\tilde{g} \subseteq g$ and $\dim_{\mathbb{R}} \tilde{g} \leq \dim_{\mathbb{R}} g$. By Theorem 7.23, these dimensions are in fact equal, giving

$$\tilde{g} = g. \quad (7.15)$$

Combined with Proposition 2.4, this gives the following theorem.

Theorem 7.32

For a matrix group $G \leq \mathrm{GL}_n(\mathbb{R})$, the exponential map $\exp: g \rightarrow M_n(\mathbb{R})$ has image in G , $\mathrm{im} \exp \leq G$. Moreover, \exp is a local diffeomorphism at O , mapping some open neighbourhood of O onto an open neighbourhood of I in G .

7.7 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 infinitely many related examples of such *Heisenberg groups* Heis_n , and the example Heis_3 that we will discuss in detail is particularly important in Quantum Physics.

For $n \geq 3$, 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

$$\begin{bmatrix} 1 & a_{12} & \cdots & \cdots & \cdots & a_{1n} \\ 0 & 1 & a_{23} & \ddots & \ddots & a_{2n} \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 & a_{n-2,n-1} & \vdots \\ \vdots & \vdots & \ddots & 0 & 1 & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix},$$

with $a_{ij} \in \mathbb{R}$. Then SUT_n is a matrix subgroup of $\mathrm{GL}_n(\mathbb{R})$ whose Lie algebra $\mathrm{sut}_n(\mathbb{R})$ consists of the matrices of the form

$$\begin{bmatrix} 0 & t_{12} & \cdots & \cdots & \cdots & t_{1n} \\ 0 & 0 & t_{23} & \ddots & \ddots & t_{2n} \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & t_{n-2,n-1} & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 & t_{n-1,n} \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

with $t_{ij} \in \mathbb{R}$, hence $\dim \text{SUT}_n = \binom{n}{2}$. It is a nice exercise to show that the following hold in general.

Proposition 7.33

For $n \geq 3$, the centre $Z(\text{SUT}_n)$ of SUT_n consists of all the matrices $[a_{ij}] \in \text{Heis}_n$ with $a_{ij} = 0$ except when $i = 1$ and $j = n$. Furthermore, $Z(\text{SUT}_n)$ is contained in the commutator subgroup of SUT_n .

Notice that there is an isomorphism of Lie groups $\mathbb{R} \cong Z(\text{SUT}_n)$ under which the subgroup of integers $\mathbb{Z} \subseteq \mathbb{R}$ corresponds to the matrices with $a_{1n} \in \mathbb{Z}$ and these form a discrete normal (in fact central) subgroup $\mathbb{Z}_n \triangleleft \text{SUT}_n$. We can form the quotient group

$$\text{Heis}_n = \text{SUT}_n / \mathbb{Z}_n.$$

Heis_n has the quotient space topology and since \mathbb{Z}_n is a discrete subgroup, the quotient map $q: \text{SUT}_n \rightarrow \text{Heis}_n$ 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 , $\mathfrak{heis}_n = \mathfrak{sut}_n$.

Proposition 7.34

For $n \geq 3$, the centre $Z(\text{Heis}_n)$ of Heis_n consists of the image under q of $Z(\text{SUT}_n)$. Furthermore, $Z(\text{Heis}_n)$ is contained in the commutator subgroup of Heis_n .

We note that $Z(\text{Heis}_n) = Z(\text{SUT}_n)/\mathbb{Z}_n$ is isomorphic to the circle group

$$\mathbf{T} = \{z \in \mathbb{C} : |z| = 1\},$$

with the correspondence coming from the map

$$\mathbb{R} \rightarrow \mathbf{T}; \quad t \mapsto e^{2\pi i t}.$$

When $n = 3$, there is a surjective Lie homomorphism

$$p: \text{SUT}_3 \rightarrow \mathbb{R}^2; \quad \begin{bmatrix} 1 & x & t \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \end{bmatrix}$$

whose kernel is $\ker p = Z(\text{SUT}_3)$. Since $\mathbb{Z}_3 \leq \ker p$, there is an induced surjective Lie homomorphism $\bar{p}: \text{Heis}_3 \rightarrow \mathbb{R}^2$ satisfying $\bar{p} \circ q = p$. In this case the

isomorphism $Z(\text{Heis}_n) \cong \mathbf{T}$ is given by

$$\begin{bmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} Z_3 \longleftrightarrow e^{2\pi i t}.$$

From now on we will denote the coset

$$\begin{bmatrix} 1 & x & t \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} Z_3 \in \text{Heis}_3$$

by $[x, y, e^{2\pi i t}]$. Thus a general element of Heis_3 has the form $[x, y, z]$ for $x, y \in \mathbb{R}$ and $z \in \mathbf{T}$; the identity element is $1 = [0, 0, 1]$. Similarly, the element

$$\begin{bmatrix} 1 & x & t \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \in \mathfrak{heis}_3$$

will be denoted (x, y, t) .

Proposition 7.35

Multiplication, inverses and commutators in Heis_3 are given by

$$[x_1, y_1, z_1][x_2, y_2, z_2] = [x_1 + x_2, y_1 + y_2, z_1 z_2 e^{2\pi i z_1 y_2}],$$

$$[x, y, z]^{-1} = [-x, -y, z^{-1} e^{2\pi i z y}]$$

$$[x_1, y_1, z_1][x_2, y_2, z_2][x_1, y_1, z_1]^{-1}[x_2, y_2, z_2]^{-1} = [0, 0, e^{2\pi i (z_1 y_2 - y_1 z_2)}].$$

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 $f: \mathbb{R} \rightarrow \mathbb{R}$ spanned by the three operators $1, p, q$ defined by

$$1f(x) = f(x), \quad pf(x) = \frac{df(x)}{dx}, \quad qf(x) = xf(x).$$

The non-trivial commutator involving these three operators is given by the *canonical commutation relation*

$$[p, q] = pq - qp = 1.$$

In \mathfrak{heis}_3 , the elements $(1, 0, 0), (1, 0, 0), (0, 0, 1)$ form a basis with the only non-trivial commutator

$$[(1, 0, 0), (1, 0, 0)] = (0, 0, 1).$$

Theorem 7.36

There are no continuous homomorphisms $\varphi: \text{Heis}_3 \rightarrow \text{GL}_n(\mathbb{C})$ with trivial kernel $\ker \varphi = 1$.

Proof

Suppose that $\varphi: \text{Heis}_3 \rightarrow \text{GL}_n(\mathbb{C})$ is a continuous homomorphism with trivial kernel and suppose that n is minimal with this property. For each $g \in \text{Heis}_3$, the matrix $\varphi(g)$ acts on vectors in \mathbb{C}^n .

We will identify $Z(\text{Heis}_3)$ with the circle T as above. Then T has a *topological generator* z_0 ; this is an element whose powers form a cyclic subgroup $\langle z_0 \rangle \leq T$ which has closure T . Proposition 10.7 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 $\epsilon > 0$, there are integers $p, q \in \mathbb{Z}$ such that

$$|s - pr - q| < \epsilon.$$

This implies that $e^{2\pi i r}$ is a topological generator of T since its powers are dense.

Let λ be an eigenvalue for the matrix $\varphi(z_0)$, with eigenvector v . If necessary replacing z_0 with z_0^{-1} , we may assume that $\lambda \geq 1$. If $\|\lambda\| > 1$, then

$$\varphi(z_0^k)v = \varphi(z_0)^k v = \lambda^k v$$

and so

$$\|\varphi(z_0^k)\| \geq \|\lambda\|^k.$$

Thus $\|\varphi(z_0^k)\| \rightarrow \infty$ as $k \rightarrow \infty$, which implies that φT is unbounded. But φ is continuous and T is compact hence φT is bounded. So in fact $\|\lambda\| = 1$.

Since φ is a homomorphism and $z_0 \in Z(\text{Heis}_3)$, for any $g \in \text{Heis}_3$ we have

$$\varphi(z_0)\varphi(g)v = \varphi(z_0g)v = \varphi(gz_0)v = \varphi(g)\varphi(z_0)v = \lambda\varphi(g)v,$$

which shows that $\varphi(g)$ is another eigenvector of $\varphi(z_0)$ for the eigenvalue λ . If we set

$$V_\lambda = \{v \in \mathbb{C}^n : \exists k \geq 1 \text{ s.t. } (\varphi(z_0) - \lambda I_n)^k v = 0\},$$

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 $g \in \text{Heis}_3$. Choose $k_0 \geq 1$ to be the largest number for which there is a vector $v_0 \in V_\lambda$ satisfying

$$(\varphi(z_0) - \lambda I_n)^{k_0} v_0 = 0, \quad (\varphi(z_0) - \lambda I_n)^{k_0-1} v_0 \neq 0.$$

If $k_0 > 1$, there are vectors $u, v \in V_\lambda$ for which

$$\varphi(z_0)u = \lambda u + v, \quad \varphi(z_0)v = \lambda v.$$

Then

$$\varphi(z_0^k)u = \varphi(z_0)^k u = \lambda^k u + k\lambda^{k-1}v$$

and since $|\lambda| = 1$,

$$\|\varphi(z_0^k)\| = \|\varphi(z_0)^k\| \geq |\lambda u + kv| \rightarrow \infty$$

as $k \rightarrow \infty$. This also contradicts the fact that φT 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 7.37

Let G be a compact Lie group and $\rho: G \rightarrow \mathrm{GL}_n(\mathbb{C})$ be a continuous homomorphism. Then for any $g \in G$, $\rho(g)$ is diagonalisable.

Having chosen a basis for V_λ , we obtain a continuous homomorphism $\theta: \mathrm{Heis}_3 \rightarrow \mathrm{GL}_d(\mathbb{C})$ for which $\theta(z_0) = \lambda I_d$. By continuity, every element of T also has the form (scalar) I_d . The minimality of n implies that we must have $d = n$ and so we can assume that $\varphi(z_0) = \lambda I_n$.

By the equation for commutators given in Proposition 7.35, each element $z \in T \leq \mathrm{Heis}_3$ is a commutator $z = ghg^{-1}h^{-1}$ in Heis_3 . Since \det and φ are homomorphisms,

$$\det \varphi(z) = \varphi(ghg^{-1}h^{-1}) = 1.$$

Thus there is a continuous function $\mu: T \rightarrow \mathbb{C}^\times$ such that for every $z \in T$,

$$\varphi(z) = \mu(z)I_d, \quad \mu(z)^d = 1.$$

As T is path connected, $\mu(z) = 1$ for every $z \in T$. Therefore for each $z \in T$, the only eigenvalue of $\varphi(z)$ is 1. This shows that $T \leq \ker \varphi$, contradicting the assumption that $\ker \varphi$ is trivial. \square

The argument used in this proof can be modified to show that none of the Heisenberg groups Heis_n with $n \geq 3$ can be realised as a matrix group.

EXERCISES

- 7.1. a) Show that the subset

$$M = \{(A, b) \in M_n(\mathbb{R}) \times \mathbb{R} : b \det A = 1\} \subseteq M_n(\mathbb{R}) \times \mathbb{R}$$

is a closed submanifold of $M_n(\mathbb{R}) \times \mathbb{R}$ and determine $T_{(A,b)} M$ for $(A,b) \in M$.

b) 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?

c) Repeat this with \mathbb{R} replaced by \mathbb{C} .

7.2. Work through the details of the calculation in Example 7.22 for the cases $n = 2, 3$.

7.3. a) Modify the details of Example 7.22 to show that $U(n) \leqslant \mathrm{GL}_n(\mathbb{C})$ is a Lie subgroup. It might be helpful to work through the cases $n = 1, 2, 3$ first.

b) Show that $SU(n) \leqslant U(n)$ is a Lie subgroup by using the determinant function

$$\det: U(n) \longrightarrow \mathbf{T} = \{z \in \mathbb{C}: |z| = 1\}$$

together with the Identity Check Trick 7.21.

7.4. Let G be a matrix group. Use Theorem 7.24 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 centraliser of g , $Z_G(g) = \{x \in G : xgx^{-1} = g\}$.

b) The centre of G , $Z(G) = \bigcap_{g \in G} Z_G(g)$.

c) For a closed subgroup $H \leqslant G$, the normaliser of H ,

$$N_G(H) = \{x \in G : xHx^{-1} = H\}.$$

d) The kernel of φ , $\ker \varphi$, where $\varphi: G \longrightarrow H$ is a continuous homomorphism into a matrix group H .

7.5. Let G be a matrix group and M be a smooth manifold. Suppose that $\mu: G \times M \longrightarrow M$ is a continuous group action as defined in Section 1.9 and investigated in the exercises for Chapter 1. Also suppose that μ is smooth, i.e., is a *smooth group action*.

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 its stabiliser

$$\mathrm{Stab}_G(X) = \{g \in G : gX = X\} \leqslant G$$

is a Lie subgroup.

- 7.6. For a Lie group G and a closed subgroup $H \leq G$, show that the cosets gH and Hg and the conjugate gHg^{-1} are submanifolds of G . In each case, identify the tangent space at a point in terms of a suitable tangent space to H .
- 7.7. Let G and H be Lie groups and $\varphi: G \rightarrow H$ be a Lie homomorphism. Show that $\ker \varphi \leq G$ is a Lie subgroup and identify the tangent space $T_g \ker \varphi$ at $g \in \ker \varphi$.
- 7.8. For $n \geq 1$, consider the subset

$$\mathcal{C}_{2n} = \{J \in \mathrm{GL}_{2n}(\mathbb{R}) : J^2 = -I_{2n}\} \subseteq \mathrm{GL}_{2n}(\mathbb{R}).$$

- a) Show that $\mathcal{C}_{2n} \subseteq \mathrm{GL}_{2n}(\mathbb{R})$ is a closed subset.
 b) Show that there is a continuous action of $\mathrm{GL}_{2n}(\mathbb{R})$ on \mathcal{C}_{2n} given by

$$A \cdot J = AJA^{-1} \quad (A \in \mathrm{GL}_{2n}(\mathbb{R})).$$

By considering the eigenvalues of elements of \mathcal{C}_{2n} , show that this action has only one orbit.

- c) Find the stabilisers of the matrices J_{2n} and J'_{2n} of Section 1.6.
 d) $\Delta\Delta$ Show that $\mathcal{C}_{2n} \subseteq \mathrm{GL}_{2n}(\mathbb{R})$ is a submanifold and find its tangent space at a point $J \in \mathcal{C}_{2n}$.
 [An element $J \in \mathrm{GL}_{2n}(\mathbb{R})$ satisfying $J^2 = -I_{2n}$ is often called a *complex structure* on \mathbb{R}^{2n} and \mathcal{C}_{2n} can be interpreted as the space of all complex structures on \mathbb{R}^{2n} .]

- 7.9. Modify the previous exercise by replacing $\mathrm{GL}_{2n}(\mathbb{R})$ with $\mathrm{O}(2n)$ and \mathcal{C}_{2n} with

$$\mathcal{C}'_{2n} = \{J \in \mathrm{O}(2n) : J^2 = -I_{2n}\} \subseteq \mathrm{O}(2n).$$

In particular show that $\mathcal{C}'_{2n} \subseteq \mathrm{O}(2n)$ is a submanifold and find its dimension. Also show that $\mathcal{C}'_{2n} \subseteq \mathrm{SO}(2n)$.

- 7.10. Determine the Lie bracket $[,]$ of the Lie algebra \mathfrak{heis}_4 of the Heisenberg group Heis_4 .

Homogeneous Spaces

Homogeneous spaces are as important in connection with Lie groups and their applications as sets of cosets are in ordinary group theory. Indeed, in the Kleinian view, a geometry consists of a homogeneous space with the group acting as its symmetry group. For detailed discussions of homogeneous spaces and the related notion of symmetric space see [6, 11]. We describe the basic ideas although we omit the proof that a homogeneous space is actually a smooth manifold since that would require somewhat more differential geometry than we have developed. Instead we focus on homogeneous spaces arising as orbits for smooth group actions and these can be studied as submanifolds.

We give a number of examples based on standard constructions in linear algebra, such as the Gram–Schmidt process and the transformation theory of real quadratic forms. In Chapter 9 we will use homogeneous spaces to investigate connectivity of various families of matrix groups.

8.1 Homogeneous Spaces as Manifolds

Let G be a Lie group of dimension $\dim G = n$ and $H \leqslant G$ be 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

$$\pi: G \longrightarrow G/H; \quad \pi(g) = gH.$$

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 .

Remark 8.1

This idea can be generalised to the case of an arbitrary surjection $q: X \rightarrow Y$ where X is a topological space and Y is a set. The resulting topology on Y is also called the *quotient topology with respect to q*. A particular case of this occurs when Y is the set of equivalence classes of an equivalence relation \sim on X and $q(x)$ is the equivalence class of $x \in X$.

Lemma 8.2

The projection map $\pi: G \rightarrow G/H$ 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

$$\theta: G \times G \rightarrow G; \quad \theta(x, y) = x^{-1}y.$$

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 $x, y \in G$ satisfy $xH \neq yH$, 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 $\pi: G \rightarrow G/H$ is characterised by a universal property which is similar in spirit to that of Theorem 5.4.

Proposition 8.3 (Universal property of the quotient topology)

Let X be a topological space and $f: G \rightarrow X$ be a continuous function for which $f(gh) = f(g)$ whenever $g \in G$ and $h \in H$. Then there is a unique continuous function $\bar{f}: G/H \rightarrow X$ for which $\bar{f} \circ \pi = f$.

$$\begin{array}{ccc} & G & \\ \pi \swarrow & & \searrow f \circ \pi \\ G/H & \xrightarrow{\exists! f \text{ continuous}} & X \end{array}$$

The following consequence characterises the topological space G/H up to homeomorphism and is proved in a similar fashion to Corollary 5.5.

Corollary 8.4

Let Q be a topological space and $q: G \rightarrow Q$ be a continuous function for which $q(gh) = q(g)$ whenever $g \in G$ and $h \in H$, and suppose that it has the following universal property:

Let X be a topological space and $f: G \rightarrow X$ be a continuous function for which $f(gh) = f(g)$ whenever $g \in G$ and $h \in H$. Then there is a unique continuous function $\tilde{f}: Q \rightarrow X$ for which $\tilde{f} \circ q = f$.

Then there is a unique homeomorphism $\bar{q}: G/H \rightarrow Q$ satisfying $\bar{q} \circ \pi = q$.

We would like to make G/H into a smooth manifold so that $\pi: G \rightarrow G/H$ is smooth. The explicit construction of an atlas is rather complicated so we merely state a result on this and then consider some examples where the smooth structure comes from an existing manifold which is diffeomorphic to a quotient. The reader is referred to [6, 29] for further details.

Theorem 8.5

G/H can be given the unique structure of a smooth manifold of dimension

$$\dim G/H = \dim G - \dim H$$

so that the projection map $\pi: G \rightarrow G/H$ is smooth and at each $g \in G$,

$$\ker(d\pi: T_g G \rightarrow T_{gH} G/H) = dL_g \mathfrak{h},$$

and there is an atlas for G/H consisting of charts of the form $\theta: W \rightarrow \theta W \subseteq \mathbb{R}^{n-k}$ for which there is a diffeomorphism $\Theta: W \times H \rightarrow \pi^{-1}W$ satisfying the conditions

$$\Theta(w, h_1 h_2) = \Theta(w, h_1) h_2, \quad \pi(\Theta(w, h)) = w \quad (w \in W, h, h_1, h_2 \in H).$$

$$\begin{array}{ccc} W \times H & \xrightarrow{\Theta} & \pi^{-1}W \\ & \searrow \text{proj}_1 & \swarrow \pi \\ & W & \end{array}$$

The projection π looks like $\text{proj}_1: \pi^{-1}W \rightarrow W$, 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 $(\theta: W \rightarrow \theta W, \Theta)$ provides a *local trivialisation 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 $(\theta: W \rightarrow \theta W, \Theta)$ and combining the map θ with a chart $\psi: U \rightarrow \psi U \subseteq \mathbb{R}^k$ for H to obtain a chart

$$(\theta \times \psi) \circ \Theta^{-1}: \Theta(W \times U) \rightarrow \theta W \times \psi U \subseteq \mathbb{R}^{n-k} \times \mathbb{R}^k = \mathbb{R}^n.$$

Such a manifold G/H is called a *homogeneous space* since each left translation map L_g on G gives rise to a diffeomorphism

$$\bar{L}_g: G/H \rightarrow G/H; \quad \bar{L}_g(xH) = gxH,$$

for which $\pi \circ L_g = \bar{L}_g \circ \pi$.

$$\begin{array}{ccc} G & \xrightarrow{L_g} & G \\ \pi \downarrow & & \pi \downarrow \\ G/H & \xrightarrow{\bar{L}_g} & G/H \end{array}$$

So each point gH has a neighbourhood diffeomorphic under \bar{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 $H \leqslant G$.

8.2 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 group actions of G on a manifolds.

Theorem 8.6

Suppose that a Lie group G acts smoothly on a manifold M . If the element $x \in M$ has stabiliser $\text{Stab}_G(x) \leq G$ and the orbit $\text{Orb}_G(x) \subseteq M$ is a closed submanifold, then the function

$$f: G/\text{Stab}_G(x) \longrightarrow \text{Orb}_G(x); \quad f(g\text{Stab}_G(x)) = gx$$

is a diffeomorphism.

Example 8.7

For $n \geq 1$, $O(n)$ acts smoothly on \mathbb{R}^n by matrix multiplication. For any non-zero vector $v \in \mathbb{R}^n$, the orbit $\text{Orb}_{O(n)}(v) \subseteq \mathbb{R}^n$ is diffeomorphic to $O(n)/O(n-1)$.

Proof

First observe that when v is the standard basis vector e_n , for $A \in O(n)$, $Ae_n = e_n$ if and only if e_n is the last column of A , while all the other columns of A are orthogonal to 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

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} \\ a_{21} & a_{22} & \cdots & a_{2n-1} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n-1} \end{bmatrix}$$

is orthogonal and hence in $O(n - 1)$. We identify $O(n - 1)$ with the subset of $O(n)$ consisting of matrices of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n-1} & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n-1} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

and then have $\text{Stab}_{O(n)}(e_n) = O(n - 1)$. The orbit of e_n is the whole unit sphere $S^{n-1} \subseteq \mathbb{R}^n$ since given a unit vector u we can extend it to an orthonormal basis $u_1, \dots, u_{n-1}, u_n = u$ whose vectors form the columns of an orthogonal matrix $U \in O(n)$ for which $Ue_n = u$. Then there is a diffeomorphism

$$O(n)/\text{Stab}_{O(n)}(e_n) = O(n)/O(n - 1) \longrightarrow \text{Orb}_{O(n)}(e_n) = S^{n-1}.$$

Now for a general non-zero vector v notice that $\text{Stab}_{O(n)}(v) = \text{Stab}_{O(n)}(\hat{v})$ where $\hat{v} = (1/|v|)v$ and

$$\text{Orb}_{O(n)}(v) = S^{n-1}(|v|),$$

the sphere of radius $|v|$. If we choose any $P \in O(n)$ with $\hat{v} = Pe_n$, we have

$$\text{Stab}_{O(n)}(v) = P \text{Stab}_{O(n)}(e_n) P^{-1}$$

and so the map

$$\text{Orb}_{O(n)}(v) \longrightarrow O(n)/P O(n - 1) P^{-1} \xrightarrow{\chi_{P^{-1}}} O(n)/O(n - 1)$$

is a diffeomorphism. □

A similar result holds for $SO(n)/SO(n - 1)$ as a homogeneous space of $SO(n)$. We can also obtain the homogeneous spaces $U(n)/U(n - 1)$ and $SU(n)/SU(n - 1)$ of the unitary and special unitary groups as orbits of non-zero vectors in \mathbb{C}^n on which these groups act by matrix multiplication; these are diffeomorphic to S^{2n-1} . The action of the quaternionic symplectic group $Sp(n)$ on \mathbb{H}^n leads to orbits of non-zero vectors diffeomorphic to $Sp(n)/Sp(n - 1)$ and S^{4n-1} .

8.3 Projective Spaces

More exotic orbit spaces are obtained as follows. Let $\mathbf{k} = \mathbb{R}$, \mathbb{C} or \mathbb{H} and $d = \dim_{\mathbb{R}} \mathbf{k}$. Consider \mathbf{k}^{n+1} as a right \mathbf{k} -vector space. There is an action of the group of units \mathbf{k}^\times on the subset of non-zero vectors $\mathbf{k}_0^{n+1} = \mathbf{k}^{n+1} - \{\mathbf{0}\}$, given by

$$z \cdot \mathbf{x} = \mathbf{x} z^{-1}.$$

The set of orbits is denoted $\mathbf{k}\mathbf{P}^n$ and is called n -dimensional \mathbf{k} -projective space. An element of $\mathbf{k}\mathbf{P}^n$ is a subset of \mathbf{k}_0^{n+1} of the form

$$[\mathbf{x}] = \{\mathbf{x} z^{-1} : z \in \mathbf{k}^\times\}.$$

Clearly $[\mathbf{x}] = [\mathbf{y}]$ if and only if there is a $z \in \mathbf{k}^\times$ for which $\mathbf{y} = \mathbf{x} z^{-1}$.

Remark 8.8

Because of this we can identify elements $\mathbf{k}\mathbf{P}^n$ with \mathbf{k} -lines in \mathbf{k}^{n+1} , i.e., 1-dimensional \mathbf{k} -vector subspaces. $\mathbf{k}\mathbf{P}^n$ is often taken to be the set of all such lines, particularly in the subject of Projective Geometry.

Associated to the map

$$q_n: \mathbf{k}_0^{n+1} \longrightarrow \mathbf{k}\mathbf{P}^n; \quad q_n(\mathbf{x}) = [\mathbf{x}],$$

is a quotient topology on $\mathbf{k}\mathbf{P}^n$ (see Remark 8.1) which is Hausdorff and separable.

Proposition 8.9

$\mathbf{k}\mathbf{P}^n$ is a smooth manifold of dimension $\dim \mathbf{k}\mathbf{P}^n = n \dim_{\mathbb{R}} \mathbf{k}$. Moreover, the quotient map $q_n: \mathbf{k}_0^{n+1} \longrightarrow \mathbf{k}\mathbf{P}^n$ is smooth with surjective derivative at every point in \mathbf{k}_0^{n+1} .

Proof

As usual, for $\mathbf{x} \in \mathbf{k}^{n+1}$ write $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_{n+1} \end{bmatrix}$. Now for each $r = 1, 2, \dots, n$, set

$$\mathbf{k}\mathbf{P}_r^n = \{[\mathbf{x}] : x_r \neq 0\} \subseteq \mathbf{k}\mathbf{P}^n.$$

Then $\mathbf{k}\mathbf{P}_r^n \subseteq \mathbf{k}\mathbf{P}^n$ is open and there is a function

$$\sigma_r: \mathbf{k}\mathbf{P}_r^n \rightarrow \mathbf{k}^n; \quad \sigma_r([\mathbf{x}]) = \begin{bmatrix} x_1 x_r^{-1} \\ \vdots \\ x_{r-1} x_r^{-1} \\ x_{r+1} x_r^{-1} \\ \vdots \\ x_{n+1} x_r^{-1} \end{bmatrix}$$

which is a continuous bijection that is actually a homeomorphism. Whenever $r \neq s$, the induced map

$$\sigma_s^{-1} \circ \sigma_r: \sigma_r^{-1} \mathbf{k}\mathbf{P}_r^n \cap \mathbf{k}\mathbf{P}_s^n \rightarrow \sigma_s^{-1} \mathbf{k}\mathbf{P}_r^n \cap \mathbf{k}\mathbf{P}_s^n$$

is given by

$$\sigma_s^{-1} \circ \sigma_r(\mathbf{x}) = \begin{bmatrix} y_1 \\ \vdots \\ y_{s-1} \\ y_{s+1} \\ \vdots \\ y_{n+1} \end{bmatrix}$$

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 \mathbf{k} . \square

An alternative description of $\mathbf{k}\mathbf{P}^n$ is obtained by considering the action of the subgroup

$$\mathbf{k}_1^\times = \{z \in \mathbf{k}^\times : |z| = 1\} \leq \mathbf{k}^\times$$

on the unit sphere $\mathbf{S}^{(n+1)d-1} \subseteq \mathbf{k}_0^{n+1}$. Notice that every element $[\mathbf{x}] \in \mathbf{k}\mathbf{P}^n$ contains elements of $\mathbf{S}^{(n+1)d-1}$. Also, if $\mathbf{x}, \mathbf{y} \in \mathbf{k}_0^{n+1}$ 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 $z \in \mathbf{k}_1^\times$. This means we can also view $\mathbf{k}\mathbf{P}^n$ as the orbit space of this action of \mathbf{k}_1^\times on $\mathbf{S}^{(n+1)d-1}$, and we also write the quotient map as $q_n: \mathbf{S}^{(n+1)d-1} \rightarrow \mathbf{k}\mathbf{P}^n$; this map is also smooth.

Proposition 8.10

The quotient space associated to the map $q_n: \mathbf{S}^{(n+1)d-1} \rightarrow \mathbf{k}\mathbf{P}^n$ is compact.

Proof

This follows from the standard fact that the image of a compact space under a continuous mapping is compact. Of course \mathbf{kP}^n is also Hausdorff. \square

Consider the action of $O(n+1)$ on the unit sphere $S^n \subseteq \mathbb{R}^{n+1}$. Then for $A \in O(n+1)$, $z = \pm 1$ and $\mathbf{x} \in S^n$, we have

$$A(\mathbf{x}z^{-1}) = (Ax)z^{-1}.$$

Hence there is an induced action of $O(n+1)$ on \mathbf{RP}^n given by

$$A \cdot [\mathbf{x}] = [Ax].$$

This action is transitive and also the matrices $\pm I_{n+1}$ fix every point of \mathbf{RP}^n . There is also an action of $SO(n+1)$ on \mathbf{RP}^n ; notice that $-I_{n+1} \in SO(n+1)$ only if n is odd.

Similarly, $U(n+1)$ and $SU(n+1)$ act on \mathbf{CP}^n with scalar matrices wI_{n+1} ($w \in \mathbf{C}_1^\times$) fixing every element. Notice that if $wI_{n+1} \in SU(n+1)$ then $w^{n+1} = 1$, so there are exactly $(n+1)$ such values.

Finally, $Sp(n+1)$ acts on \mathbf{HP}^n and the matrices $\pm I_{n+1}$ fix every element.

There are some important new quotient Lie groups associated to these actions, the *projective unitary*, *special unitary* and *quaternionic symplectic groups*

$$PU(n+1) = U(n+1)/\{wI_{n+1} : w \in \mathbf{C}_1^\times\},$$

$$PSU(n+1) = SU(n+1)/\{wI_{n+1} : w^{n+1} = 1\},$$

$$PSp(n+1) = Sp(n+1)/\{\pm I_{n+1}\}.$$

Each of these is a subgroup of one of the three *projective linear groups* $PGL_n(\mathbb{R})$, $PGL_n(\mathbb{C})$ and $PGL_n(\mathbb{H})$ where the first two were mentioned in Section 4.6, while the last is defined by

$$PGL_n(\mathbb{H}) = PGL_n(\mathbb{H})/\{tI_n : t \in \mathbb{R}^\times\}.$$

Projective spaces are themselves homogeneous spaces. Consider the subgroup of $O(n+1)$ 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 $O(n+1)$ by $O(n) \times O(1)$. There is a subgroup $\widetilde{O}(n) \leqslant SO(n+1)$ whose elements have the form

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

where

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} \\ a_{21} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ a_{n-11} & \ddots & \ddots & a_{n-1n-1} \end{bmatrix} \in O(n),$$

$$\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} \\ a_{21} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ a_{n-11} & \ddots & \ddots & a_{n-1n-1} \end{bmatrix} = w.$$

Similarly, there is a subgroup $U(n) \times U(1) \leqslant U(n+1)$ whose elements have the form

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

and $\widetilde{U}(n) \leqslant SU(n+1)$ with elements

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

where

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} \\ a_{21} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ a_{n-11} & \ddots & \ddots & a_{n-1n-1} \end{bmatrix} \in U(n),$$

$$\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} \\ a_{21} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ a_{n-11} & \ddots & \ddots & a_{n-1n-1} \end{bmatrix}^{-1} = w.$$

Finally we have $Sp(n) \times Sp(1) \in Sp(n+1)$ consisting of matrices of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & 0 \\ a_{21} & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-11} & \ddots & \ddots & a_{n-1n-1} & 0 \\ 0 & 0 & \cdots & 0 & w \end{bmatrix}.$$

Proposition 8.11

There are diffeomorphisms

$$\mathbb{R}P^n \longrightarrow O(n+1)/O(n) \times O(1), \quad \mathbb{R}P^n \longrightarrow SO(n+1)/\widetilde{O(n)};$$

$$\mathbb{C}P^n \longrightarrow U(n+1)/U(n) \times U(1), \quad \mathbb{C}P^n \longrightarrow SU(n+1)/\widetilde{U(n)};$$

$$\mathbb{H}P^n \longrightarrow Sp(n+1)/Sp(n) \times Sp(1).$$

There are similar homogeneous spaces of the general and special linear groups also diffeomorphic to these projective spaces. We illustrate this with one example.

The special linear group $SL_2(\mathbb{C})$ contains the matrix subgroup P consisting of its lower triangular matrices

$$\begin{bmatrix} u & 0 \\ w & v \end{bmatrix} \in SL_2(\mathbb{C}).$$

P is often referred to as a *parabolic subgroup* of $SL_2(\mathbb{C})$.

Proposition 8.12

$\mathrm{SL}_2(\mathbb{C})/\mathrm{P}$ is diffeomorphic to \mathbb{CP}^1 .

Proof

There is smooth map

$$\psi: \mathrm{SL}_2(\mathbb{C}) \longrightarrow \mathbb{CP}^1; \quad \psi(A) = [Ae_2].$$

Notice that for $B = \begin{bmatrix} u & 0 \\ w & v \end{bmatrix} \in \mathrm{P}$,

$$\begin{bmatrix} u & 0 \\ w & v \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ v \end{bmatrix},$$

so $[(AB)e_2] = [Ae_2]$ for any $A \in \mathrm{SL}_2(\mathbb{C})$. This means that $\psi(A)$ only depends on the coset $AP \in \mathrm{SL}_2(\mathbb{C})/\mathrm{P}$. 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. \square

8.4 Grassmannians

There are some important families of homogeneous spaces directly generalising projective spaces. These are the real, complex and quaternionic *Grassmannians*, which we now define.

Let $O(k) \times O(n-k) \leqslant O(n)$ be the closed subgroup whose elements have the form

$$\begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix} \quad (A \in O(k), B \in O(n-k)).$$

Similarly there are closed subgroups $U(k) \times U(n-k) \leqslant U(n)$ and $Sp(k) \times Sp(n-k) \leqslant Sp(n)$ with elements

$$U(k) \times U(n-k) : \quad \begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix} \quad (A \in U(k), B \in U(n-k));$$

$$Sp(k) \times Sp(n-k) : \quad \begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix} \quad (A \in Sp(k), B \in Sp(n-k)).$$

The associated homogeneous spaces are the *Grassmannians*

$$Gr_{k,n}(\mathbb{R}) = O(n)/O(k) \times O(n-k);$$

$$Gr_{k,n}(\mathbb{C}) = U(n)/U(k) \times U(n-k);$$

$$Gr_{k,n}(\mathbb{H}) = Sp(n)/Sp(k) \times Sp(n-k).$$

Proposition 8.13

For $\mathbf{k} = \mathbb{R}, \mathbb{C}, \mathbb{H}$, the Grassmannian $\text{Gr}_{k,n}(\mathbf{k})$ can be viewed as the set of all k -dimensional \mathbf{k} -vector subspaces in \mathbf{k}^n .

Proof

We describe the real case $\mathbf{k} = \mathbb{R}$, the others being analogous.

Associated to element $W \in O(n)$ is the subspace spanned by the first k columns of W , say w_1, \dots, w_k ; we will denote this subspace by $\langle w_1, \dots, w_k \rangle$. As the columns of W are an orthonormal set, they are linearly independent, hence $\dim_{\mathbb{R}} \langle w_1, \dots, w_k \rangle = k$. Notice that the remaining $(n - k)$ columns give rise to another subspace $\langle w_{k+1}, \dots, w_n \rangle$ of dimension $\dim_{\mathbb{R}} \langle w_{k+1}, \dots, w_n \rangle = n - k$. In fact these are mutually orthogonal in the sense that

$$\begin{aligned}\langle w_{k+1}, \dots, w_n \rangle &= \langle w_1, \dots, w_k \rangle^\perp \\ &= \{x \in \mathbb{R}^n : x \cdot w_r = 0, r = 1, \dots, k\}, \\ \langle w_1, \dots, w_k \rangle &= \langle w_{k+1}, \dots, w_n \rangle^\perp \\ &= \{x \in \mathbb{R}^n : x \cdot w_r = 0, r = k + 1, \dots, n\}.\end{aligned}$$

For a matrix

$$\begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix} \in O(k) \times O(n - k),$$

the columns in the product

$$W' = W \begin{bmatrix} A & O_{k,n-k} \\ O_{n-k,k} & B \end{bmatrix}$$

span subspaces $\langle w'_1, \dots, w'_k \rangle$ and $\langle w'_{k+1}, \dots, w'_n \rangle$. Note that w'_1, \dots, w'_k are orthonormal and also linear combinations of w_1, \dots, w_k ; similarly, w'_{k+1}, \dots, w'_n are linear combinations of w_{k+1}, \dots, w_n . Hence

$$\langle w'_1, \dots, w'_k \rangle = \langle w_1, \dots, w_k \rangle, \quad \langle w'_{k+1}, \dots, w'_n \rangle = \langle w_{k+1}, \dots, w_n \rangle.$$

So there is a well-defined function

$$O(n)/O(k) \times O(n - k) \longrightarrow k\text{-dimensional vector subspaces of } \mathbb{R}^n$$

which sends the coset of W to the subspace $\langle w_1, \dots, w_k \rangle$. This is actually a bijection.

Notice also that there is another bijection

$$O(n)/O(k) \times O(n - k) \longrightarrow (n - k)\text{-dimensional vector subspaces of } \mathbb{R}^n$$

which sends the coset of W to the subspace $\langle w_{k+1}, \dots, w_n \rangle$. This corresponds to a diffeomorphism $\text{Gr}_{k,n}(\mathbb{R}) \rightarrow \text{Gr}_{n-k,n}(\mathbb{R})$ which in turn corresponds to the obvious isomorphism $O(k) \times O(n-k) \rightarrow O(n-k) \times O(k)$ induced by conjugation by a suitable element $P \in O(n)$. \square

8.5 The Gram–Schmidt Process

The Gram–Schmidt process provides a useful algorithm which allows an arbitrary basis of \mathbb{R}^n to be replaced by an orthonormal basis. First we recall this and then explain how it gives rise to a homogeneous space of $GL_n(\mathbb{R})$.

Let $\{u_1, \dots, u_n\}$ be a basis of \mathbb{R}^n . Writing $u_i^{(0)} = u_i$, we construct a new sequence of vectors $u_1^{(1)}, \dots, u_n^{(1)}$ by

$$u_i^{(1)} = \frac{1}{|u_i^{(0)}|} u_i^{(0)}, \quad u_i^{(1)} = u_i^{(0)} - (u_1^{(1)} \cdot u_i^{(0)}) u_1^{(0)} \quad (i = 2, \dots, n).$$

It is clear that these form a basis of \mathbb{R}^n for which

$$|u_i^{(1)}| = 1, \quad u_1^{(1)} \cdot u_i^{(1)} = 0 \quad (i = 2, \dots, n).$$

If $U^{(0)}$ is the matrix whose columns are the $u_i^{(0)}$, then the matrix $U^{(1)}$ whose columns are the $u_i^{(1)}$ is obtained as

$$U^{(1)} = U^{(0)} \text{ diag} \left(\frac{1}{|u_1^{(0)}|}, 1, \dots, 1 \right) \begin{bmatrix} 1 & -u_1^{(1)} \cdot u_2^{(0)} & \dots & \dots & -u_1^{(1)} \cdot u_n^{(0)} \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \\ = U^{(0)} T^{(1)},$$

where $T^{(1)}$ is upper triangular and has positive diagonal entries.

This construction can be iterated by defining for each $k = 1, \dots, n$, a sequence $u_1^{(k)}, \dots, u_n^{(k)}$ with

$$u_i^{(k)} = \begin{cases} u_i^{(k-1)} & \text{if } i = 1, \dots, k-1, \\ \frac{1}{|u_k^{(k-1)}|} u_k^{(k-1)} & \text{if } i = k, \\ u_i^{(k-1)} - \sum_{r=1}^k (u_r^{(k)} \cdot u_i^{(k-1)}) u_r^{(k)} & \text{if } i = k+1, \dots, n. \end{cases}$$

It is straightforward to check that these vectors form a basis satisfying

$$\begin{aligned} |\mathbf{u}_i^{(k)}| &= 1 \quad (i = 1, \dots, k), \\ \mathbf{u}_i^{(k)} \cdot \mathbf{u}_j^{(k)} &= 0 \quad (i = 1, \dots, k, j = 1, \dots, n). \end{aligned}$$

In particular, the basis $\{\mathbf{u}_1^{(n)}, \dots, \mathbf{u}_n^{(n)}\}$ is orthonormal,

$$\begin{aligned} |\mathbf{u}_i^{(n)}| &= 1 \quad (i = 1, \dots, n), \\ \mathbf{u}_i^{(n)} \cdot \mathbf{u}_j^{(n)} &= 0 \quad (i, j = 1, \dots, n). \end{aligned}$$

At each stage of the iteration the matrix $U^{(k)}$ whose i th column is the vector $\mathbf{u}_i^{(k)}$ given by

$$U^{(k)} = U^{(k-1)}T^{(k)},$$

where $T^{(k)}$ is upper triangular with positive diagonal entries.

From all of this we can deduce that

$$U^{(n)} = U^{(0)}T$$

where $T = T^{(1)} \dots T^{(n)}$ is upper triangular. Since $U^{(n)}$ is orthogonal and T^{-1} is upper triangular with positive diagonal entries, we obtain

$$U^{(0)} = U^{(n)}T^{-1} \in O(n) \cap UT_n(\mathbb{R}).$$

If $PDUT_n(\mathbb{R}) \leqslant UT_n(\mathbb{R})$ is the matrix subgroup consisting of all upper triangular matrices with positive diagonal entries, then

$$PDUT_n(\mathbb{R}) \cap O(n) = \{I\},$$

hence we have shown the following.

Proposition 8.14

Every matrix $A \in \mathbb{M}_n(\mathbb{R})$ has a unique factorisation $A = UT$ with $U \in O(n)$ and $T \in PDUT_n(\mathbb{R})$.

Corollary 8.15

There is a diffeomorphism $GL_n(\mathbb{R}) / PDUT_n(\mathbb{R}) \rightarrow O(n)$.

By conjugating and considering the matrix subgroup $PDLT_n(\mathbb{R}) \leqslant GL_n(\mathbb{R})$ of lower triangular matrices with positive diagonal entries, we also obtain the following.

Corollary 8.16

There is a diffeomorphism $\mathrm{GL}_n(\mathbb{R})/\mathrm{O}(n) \longrightarrow \mathrm{PDLT}_n(\mathbb{R})$.

Here the space $\mathrm{PDLT}_n(\mathbb{R})$ is contractible, hence so is the homogeneous space $\mathrm{GL}_n(\mathbb{R})/\mathrm{O}(n)$.

A similar construction works for a \mathbb{C} -basis of \mathbb{C}^n , using the standard hermitian inner product.

8.6 Reduced Echelon Form

The basic theory of systems of linear equations makes use of *Gaussian elimination* to produce the *reduced echelon form* of a matrix. For example, starting with an invertible matrix $A \in M_n(\mathbf{k})$ where $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , a sequence of elementary column operations will convert A into a lower triangular matrix. If A is invertible, $A \in \mathrm{GL}_n(\mathbf{k})$, this lower triangular matrix is the identity matrix I_n . If we only use elementary column operations which involve adding a multiple of column j to column i where $j < i$, then the best result will be a lower triangular matrix with non-zero diagonal terms; but even this may not be possible. To ensure that we obtain a lower triangular matrix we may also need to permute columns. This leads to the *LPU-decomposition* of A , see [28].

Proposition 8.17

Let $A \in \mathrm{GL}_n(\mathbf{k})$. Then there are matrices $L, P, U \in \mathrm{GL}_n(\mathbf{k})$ where L is lower triangular, P is a permutation matrix, U is upper triangular and $A = LPU$. Moreover, P is unique and if we require that the diagonal entries of L are all 1, then so are L and U . Therefore the homogeneous space $\mathrm{GL}_n(\mathbf{k})/\mathrm{UT}_n(\mathbf{k})$ is diffeomorphic to the manifold

$$\mathbf{k}^{\binom{n+1}{2}} \times \mathrm{Sym}_n,$$

where Sym_n is the symmetric group on n elements.

This is a closely related to the important concept of a *Bruhat decomposition*.

8.77 Real Inner Products

Recall that an *inner product* on the real vector space \mathbb{R}^n is an \mathbb{R} -bilinear map $\beta: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ which satisfies

$$\beta(\mathbf{x}, \mathbf{y}) = \beta(\mathbf{y}, \mathbf{x})$$

and so is *symmetric*. β is *non-degenerate* if for every non-zero $\mathbf{x} \in \mathbb{R}^n$ there is a $\mathbf{y} \in \mathbb{R}^n$ for which $\beta(\mathbf{x}, \mathbf{y}) \neq 0$.

(Given the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ of \mathbb{R}^n , β is determined by the $n \times n$ matrix $B_\beta = [b_{ij}]$ for which

$$b_{ij} = \beta(\mathbf{e}_i, \mathbf{e}_j) = b_{ji}.$$

Hence B_β is a symmetric matrix and

$$\beta(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T B_\beta \mathbf{y}.$$

If we express an arbitrary vector $\mathbf{x} \in \mathbb{R}^n$ as $\mathbf{x} = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n$ for $x_i \in \mathbb{R}$, then

$$\begin{aligned}\beta(\mathbf{x}, \mathbf{x}) &= \sum_{i,j=1}^n b_{ij} x_i x_j \\ &= \sum_{i=1}^n b_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} b_{ij} x_i x_j \\ &= Q_\beta(\mathbf{x}),\end{aligned}$$

which is a real symmetric *quadratic form*. The following is a well-known result on real inner products and quadratic forms.

Theorem 8.18

Given an inner product β on \mathbb{R}^n , there is a matrix $P \in \mathrm{GL}_n(\mathbb{R})$ for which

$$(P^T)^{-1} B_\beta P^{-1} = \mathrm{diag}(\underbrace{1, \dots, 1}_{r_+}, \underbrace{0, \dots, 0}_{r_0}, \underbrace{-1, \dots, -1}_{r_-})$$

with r_+ 1's, r_0 0's and r_- -1's. Moreover, the numbers r_+ , r_0 and r_- are independent of the matrix P .

We will say that two inner products β_1 and β_2 on \mathbb{R}^n (or equivalently the symmetric matrices B_{β_1} and B_{β_2}) are *related* and write $\beta_1 \approx \beta_2$ (or $B_{\beta_1} \approx B_{\beta_2}$) if

$$B_{\beta_2} = (P^T)^{-1} B_{\beta_1} P^{-1}$$

for some $P \in \mathrm{GL}_n(\mathbb{R})$.

Corollary 8.19

β is non-degenerate if and only if $\det B \neq 0$, or equivalently if $r_0 \neq 0$.

Certain combinations of the numbers occurring here are given special names:

- the *rank* of $\beta = \text{rank } \beta = n = r_+ + r_0 + r_-$,
- the *index* or *signature* of $\beta = \text{sign } \beta = r_+ - r_-$,
- the *nullity* of $\beta = \text{null } \beta = r_0$.

These three numbers determine the diagonal matrix $(P^T)^{-1}B\beta P$. We can also sensibly refer to $\text{rank } B$, $\text{sign } B$ and $\text{null } B$ for a real symmetric matrix.

The $n \times n$ real symmetric matrices form a subspace $\text{Sym}_n(\mathbb{R}) \subseteq M_n(\mathbb{R})$ of dimension $\binom{n}{2}$. We introduce a continuous action of $\text{GL}_n(\mathbb{R})$ on $\text{Sym}_n(\mathbb{R})$ by

$$\text{GL}_n(\mathbb{R}) \times \text{Sym}_n(\mathbb{R}) \longrightarrow \text{Sym}_n(\mathbb{R}); \quad P \cdot B = (P^T)^{-1}BP^{-1}.$$

Notice that for each $P \in \text{GL}_n(\mathbb{R})$, P acts as a linear transformation on $\text{Sym}_n(\mathbb{R})$.

We can of course consider the stabiliser and orbit of an element $B \in \text{Sym}_n(\mathbb{R})$. For example, if $B = I_n$, then

$$\begin{aligned} \text{Stab}_{\text{GL}_n(\mathbb{R})}(I_n) &= \{P \in \text{GL}_n(\mathbb{R}) : (P^T)^{-1}I_nP^{-1} = I_n\} = O(n), \\ \text{Orb}_{\text{GL}_n(\mathbb{R})}(I_n) &= \{B \in \text{Sym}_n(\mathbb{R}) : \text{sign } B = n\}. \end{aligned}$$

Hence we see that the homogenous space of $O(n)$ is

$$\text{GL}_n(\mathbb{R})/O(n) = \{B \in \text{Sym}_n(\mathbb{R}) : \text{sign } B = n\},$$

which can also be interpreted as the set of all *positive definite* inner products.

More generally, following the ideas of Section 1.5, considering any $n \times n$ real symmetric matrix Q , we find that its stabiliser is

$$\text{Stab}_{\text{GL}_n(\mathbb{R})}(Q) = O_Q = \{P \in \text{GL}_n(\mathbb{R}) : P^TQP = Q\},$$

the generalised orthogonal group associated with Q . In particular, when

$$Q = \text{diag}(\underbrace{1, \dots, 1}_{r_+}, \underbrace{0, \dots, 0}_{r_0}, \underbrace{-1, \dots, -1}_{r_-}),$$

the homogeneous space $\text{GL}_n(\mathbb{R})/O_Q$ can be interpreted as the space of all inner products with rank n , signature $(r_+ - r_-)$ and nullity r_0 .

8.8 Symplectic Forms

Definition 8.20

A *symplectic form* on \mathbb{R}^n is a bilinear form

$$\omega: \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

which is *skew symmetric*, i.e., for all $x, y \in \mathbb{R}^n$,

$$\omega(y, x) = -\omega(x, y),$$

and non-degenerate, i.e., for every non-zero vector $u \in \mathbb{R}^n$, there is a vector $v \in \mathbb{R}^n$ for which $\omega(u, v) \neq 0$.

Suppose that $v = \{v_1, \dots, v_n\}$ is a basis for \mathbb{R}^n . Then associated with a skew symmetric bilinear form ω on \mathbb{R}^n is a matrix $S_{\omega, v} = [s_{ij}]$ given by

$$s_{ij} = \omega(v_i, v_j).$$

It is easy to see that $S_{\omega, v}$ is skew symmetric. Conversely, a skew symmetric matrix $S = [s_{ij}]$ gives rise to a skew symmetric bilinear form ω for which

$$s_{ij} = \omega(v_i, v_j).$$

An alternative way to see this connection is using the formula

$$\omega(x, y) = x^T S_{\omega, v} y \quad (x, y \in \mathbb{R}^n). \quad (8.1)$$

Notice that for any vector $x \in \mathbb{R}^n$,

$$\omega(x, x) = -\omega(x, x),$$

hence $\omega(x, x) = 0$.

Lemma 8.21

- i) A skew symmetric bilinear form ω on \mathbb{R}^n is a symplectic form if and only if the associated matrix $S_{\omega, v}$ is non-singular, i.e., $\det S_{\omega, v} \neq 0$.
- ii) If ω is a symplectic form on \mathbb{R}^n , then n must be even.

Proof

- (i) Suppose that ω is a symplectic form. If $S_{\omega, v}$ is singular, there is a non-zero vector u for which $S_{\omega, v} u = 0$. If $v \in \mathbb{R}^n$ is any vector for which $\omega(u, v) \neq 0$, then

$$\omega(u, v) = -\omega(v, u) = v^T S_{\omega, v} u = 0,$$

which contradicts the assumption on \mathbf{v} .

If $S_{\omega, \mathbf{v}}$ is non-singular, then for any non-zero vector $\mathbf{u} \in \mathbb{R}^n$, $S_{\omega, \mathbf{v}}\mathbf{u} \neq 0$, hence there must be some vector $\mathbf{v} \in \mathbb{R}^n$ for which

$$\mathbf{v}^T(S_{\omega, \mathbf{v}}\mathbf{u}) \neq 0.$$

But then we have $\omega(\mathbf{v}, \mathbf{u}) \neq 0$. So ω is non-degenerate.

(ii) If $\det S_{\omega, \mathbf{v}} \neq 0$ then n must be even by Equation (1.5). \square

Theorem 8.22 (Darboux's Theorem)

Let ω be a symplectic form on \mathbb{R}^{2m} . Then there is a basis $\mathbf{w}_1, \dots, \mathbf{w}_{2m}$ of \mathbb{R}^{2m} for which

$$\begin{aligned}\omega(\mathbf{w}_i, \mathbf{w}_j) &= 0 \quad \text{if } (i - j) \neq \pm 1, \\ \omega(\mathbf{w}_{2k}, \mathbf{w}_{2k+1}) &= 0, \\ \omega(\mathbf{w}_{2k-1}, \mathbf{w}_{2k}) &= 1.\end{aligned}$$

Proof

Start with any basis $\mathbf{v} = \{\mathbf{v}_1, \dots, \mathbf{v}_{2m}\}$ of \mathbb{R}^{2m} . Taking $\mathbf{w}_1 = \mathbf{v}_1$, notice that since ω is non-degenerate, at least one of the \mathbf{v}_j with $j \neq 1$ must satisfy

$$\omega(\mathbf{w}_1, \mathbf{v}_j) \neq 0.$$

By reordering the vectors \mathbf{v}_j if necessary, we can assume that $j = 2$. Setting

$$\mathbf{w}_2 = \frac{1}{\omega(\mathbf{v}_2, \mathbf{w}_1)} \mathbf{v}_2,$$

we have $\omega(\mathbf{w}_1, \mathbf{w}_2) = 1$. Notice that

$$\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{v}_3, \dots, \mathbf{v}_{2m}\}$$

is a basis. Replace each of the vectors $\mathbf{v}_3, \dots, \mathbf{v}_{2m}$ by

$$\mathbf{v}'_j = \mathbf{v}_j - \omega(\mathbf{v}_j, \mathbf{w}_2)\mathbf{w}_1 + \omega(\mathbf{v}_j, \mathbf{w}_1)\mathbf{w}_2.$$

It is easy to see that

$$\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{v}'_3, \dots, \mathbf{v}'_{2m}\}$$

is a basis for which

$$\omega(\mathbf{w}_i, \mathbf{v}'_j) = 0.$$

We can keep repeating this process to replace this basis by another of the form

$$\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \dots, \mathbf{w}_{2k-1}, \mathbf{w}_{2k}, \mathbf{v}''_{2k}, \dots, \mathbf{v}''_{2m}\}$$

for which

$$\begin{aligned}\omega(\mathbf{w}_i, \mathbf{w}_j) &= 0 \quad \text{if } (i - j) \neq \pm 1, \\ \omega(\mathbf{w}_{2r}, \mathbf{w}_{2r+1}) &= 0, \\ \omega(\mathbf{w}_{2r-1}, \mathbf{w}_{2r}) &= 1, \\ \omega(\mathbf{w}_i, \mathbf{v}_j'') &= 0.\end{aligned}$$

Continuing with this we obtain a basis

$$\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \dots, \mathbf{w}_{2m-1}, \mathbf{w}_{2m}\}$$

with $\omega(\mathbf{w}_i, \mathbf{w}_j)$ with properties as above. \square

Corollary 8.23

If ω is a symplectic form on \mathbb{R}^{2m} , there is a basis

$$\mathbf{w} = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \dots, \mathbf{w}_{2m-1}, \mathbf{w}_{2m}\}$$

for which $S_{\omega, \mathbf{w}} = J_{2m}$.

Corollary 8.24

Let S be a non-singular $2m \times 2m$ skew symmetric matrix. Then there is a matrix $P \in \mathrm{GL}_{2m}(\mathbb{R})$ for which $S = P^T J_{2m} P$.

It is sometimes useful to reorder such a basis \mathbf{w} with $S_{\omega, \mathbf{w}} = J_{2m}$ to give another basis

$$\tilde{\mathbf{w}} = \{\mathbf{w}_1, \mathbf{w}_3, \dots, \mathbf{w}_{2m-1}, \mathbf{w}_2, \mathbf{w}_4, \dots, \mathbf{w}_{2m}\}$$

for which

$$S_{\omega, \tilde{\mathbf{w}}} = \begin{bmatrix} O_m & I_n \\ -I_n & O_m \end{bmatrix} = J'_{2m}.$$

This leads to another canonical form for a skew symmetric matrix.

Corollary 8.25

Let S be a non-singular $2m \times 2m$ skew symmetric matrix. Then there is a matrix $Q \in \mathrm{GL}_{2m}(\mathbb{R})$ for which $S = Q^T J'_{2m} Q$.

Let $\Sigma_{2m} \subseteq M_{2m}(\mathbb{R})$ be the subspace of all $2m \times 2m$ non-singular real skew symmetric matrices. There is a continuous action of $GL_{2m}(\mathbb{R})$ on Σ_{2m} given by

$$P \cdot S = (P^{-1})^T S P^{-1}.$$

Then using Corollary 8.24, the homogeneous space $GL_{2m}(\mathbb{R}) / Symp_{2m}(\mathbb{R})$ can be identified with Σ_{2m} . A similar result is true for $GL_{2m}(\mathbb{R}) / Symp'_{2m}(\mathbb{R})$ and Proposition 1.42 can be viewed as a special case of part (iii) of Theorem 1.60. The details are left as an exercise.

We end by giving another interpretation of the *Pfaffian function* introduced in the exercises of Chapter 5.

Proposition 8.26

There is a continuous function

$$\tilde{\text{pf}}: GL_{2m}(\mathbb{R}) / Symp_{2m}(\mathbb{R}) \longrightarrow \mathbb{R}^\times; \quad \tilde{\text{pf}}(P Symp_{2m}(\mathbb{R})) = \det P.$$

Proof

The point is that the coset $P Symp_{2m}(\mathbb{R}) \in GL_{2m}(\mathbb{R}) / Symp_{2m}(\mathbb{R})$ corresponds to a skew symmetric matrix $P^T J_{2m} P$ and then

$$\tilde{\text{pf}}(P Symp_{2m}(\mathbb{R})) = \text{pf}(P^T J_{2m} P).$$

We leave the reader to work out the details. □

EXERCISES

- 8.1. Let $\mathbf{k} = \mathbb{R}$ or \mathbb{C} and $n \geq 1$. From the proof of Proposition 8.9, recall that the projective space $\mathbf{k}\mathbf{P}^n$ contains an open subset

$$\mathbf{k}\mathbf{P}_{n+1}^n = \{[x] : x \neq 0, x_{n+1} \neq 0\}$$

with a diffeomorphism $\sigma = \sigma_{n+1}: \mathbf{k}\mathbf{P}_{n+1}^n \longrightarrow \mathbf{k}^n$, where as usual we write

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_{n+1} \end{bmatrix}.$$

Also recall the affine group $\text{Aff}_n(\mathbf{k}) \leqslant GL_{n+1}(\mathbf{k})$.

- a) Show that the action of $\text{Aff}_n(\mathbf{k})$ on \mathbf{k}_0^{n+1} gives rise to a continuous

action of $\text{Aff}_n(\mathbf{k})$ on $\mathbf{k}\mathbf{P}^n$.

b) Show that the action of (a) restricts to an action of $\text{Aff}_n(\mathbf{k})$ on $\mathbf{k}\mathbf{P}_{n+1}^n$.

c) Using σ , obtain an action of $\text{Aff}_n(\mathbf{k})$ on \mathbf{k}^n . Explain how this action is related to that of Section 1.5.

d) How much of this would make sense with $\mathbf{k} = \mathbf{H}$, the skew field of quaternions?

- 8.2. For $n \geq 1$, consider the matrix group $\text{Sp}(n) \times \text{Sp}(1)$ and its quotient, the *quasi-symplectic group*

$$\text{Sp}(n) \tilde{\times} \text{Sp}(1) = \text{Sp}(n) \times \text{Sp}(1) / \{(I_n, 1), (-I_n, -1)\}.$$

a) Show that the Lie algebra of $\text{Sp}(n) \tilde{\times} \text{Sp}(1)$ is the direct product of two Lie ideals of dimensions $(2n^2 + n)$ and 3. Is the Lie group $\text{Sp}(n) \tilde{\times} \text{Sp}(1)$ a semi-direct product of groups of those dimensions?

b) Show that \mathbf{HP}^{n-1} is a homogeneous space of $\text{Sp}(n) \tilde{\times} \text{Sp}(1)$.

- 8.3. In the exercises for Chapter 7 we saw that the set of complex structures on \mathbb{R}^{2n} is a submanifold

$$\mathcal{C}_{2n} = \{J \in \text{GL}_{2n}(\mathbb{R}) : J^2 = -I_{2n}\} \subseteq \text{GL}_{2n}(\mathbb{R})$$

and that $\text{GL}_{2n}(\mathbb{R})$ acts smoothly on \mathcal{C}_{2n} with a single orbit. Use this result to identify the homogeneous spaces $\text{GL}_{2n}(\mathbb{R})/\rho_n \text{GL}_{2n}(\mathbb{C})$ and $\text{GL}_{2n}(\mathbb{R})/\rho'_n \text{GL}_{2n}(\mathbb{C})$ of Section 1.6.

Connectivity of Matrix Groups

9.1 Connectivity of Manifolds

Definition 9.1

Let X be a topological space.

- X is *connected* if whenever $X = U \cup V$ with $U, V \neq \emptyset$, then $U \cap V \neq \emptyset$.
- X is *path connected* if whenever $x, y \in X$, there is a continuous path $p: [0, 1] \rightarrow X$ 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 is a fundamental result of Real Analysis.

Proposition 9.2

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 9.3

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 X , there is a continuous map $p: [0, 1] \rightarrow X$ 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 9.4

Let X be a connected topological space which is locally path connected. Then X is path connected.

Proof

Let $x \in X$, and set

$$X_x = \{y \in X : \exists p: [0, 1] \rightarrow X \text{ continuous with } p(0) = x \text{ and } p(1) = y\}.$$

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 x to z via y , hence $U_y \subseteq X_z$. 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 $X - X_x = \emptyset$. Hence X is path connected. \square

Proposition 9.5

If the topological spaces X and Y are path connected then their product $X \times Y$ is path connected.

Corollary 9.6

For $n \geq 1$, \mathbb{R}^n is path connected and connected.

We also record the following standard results.

Proposition 9.7

- i) Let $n \geq 2$. The unit sphere $S^{n-1} \subseteq \mathbb{R}^n$ is path connected. In $S^0 = \{\pm 1\} \subseteq \mathbb{R}$, the subsets $\{1\}$ and $\{-1\}$ are path connected. The set of non-zero vectors $\mathbb{R}_0^n \subseteq \mathbb{R}^n$ is path connected.
- ii) For $n \geq 1$, the sets of non-zero complex and quaternionic vectors $\mathbb{C}_0^n \subseteq \mathbb{C}^n$ and $\mathbb{H}_0^n \subseteq \mathbb{H}^n$ are path connected.

Proposition 9.8

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 9.4. \square

Theorem 9.9

Let M be a connected manifold and $N \subseteq M$ be 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 9.10

Let G be a Lie group and $H \leq G$ be a closed subgroup. If H and G/H are connected, then so is G .

Proof

First we remark on the following: for any $g \in G$, the left translation map $\tau_g : H \rightarrow gH$ 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 non-empty open subsets for which $U \cap V = \emptyset$ and $U \cup V = G$. By Lemma 8.2, the projection $\pi: G \rightarrow G/H$ 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

Propositions 9.8 and 9.10 provide 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 a closed subgroup of a Lie group is a submanifold by Theorem 7.31.

Proposition 9.11

Let G be a Lie group and $H \leq G$ be a closed subgroup. If H and G/H are connected, then G is path connected.

9.2 Examples of Path Connected Matrix Groups

In this section we apply Proposition 9.11 to show that many familiar matrix groups are path connected.

Example 9.12

For $n \geq 1$, $\mathrm{SL}_n(\mathbb{R})$ 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 $n \geq 2$.

Recall that $\mathrm{SL}_n(\mathbb{R})$ acts continuously on \mathbb{R}^n by matrix multiplication. Consider the continuous function

$$f: \mathrm{SL}_n(\mathbb{R}) \longrightarrow \mathbb{R}^n; \quad f(A) = A\mathbf{e}_n.$$

The image of f is $\mathrm{im} f = \mathbb{R}_0^n = \mathbb{R}^n - \{0\}$ since every vector $\mathbf{v} \in \mathbb{R}_0^n$ can be extended to a basis

$$\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, \mathbf{v}_n = \mathbf{v}$$

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 stabiliser 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 $A = A_{\mathbf{v}}P$ 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}_0^n .

Since $n \geq 2$, it is well known that \mathbb{R}_0^n is path connected, hence it is connected. This implies that $\mathrm{SL}_n(\mathbb{R})/\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$ is connected.

The subgroup $\mathrm{SL}_{n-1}(\mathbb{R}) \leq \mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n)$ 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 9.6 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}_0^n to deduce that $\mathrm{SL}_n(\mathbb{R})$ is path connected, thus demonstrating the inductive step. \square

Example 9.13

For $n \geq 1$, $\mathrm{GL}_n^+(\mathbb{R})$ is path connected.

Proof

Since $SL_n(\mathbb{R}) \leq GL_n^+(\mathbb{R})$, it suffices to show that $GL_n^+(\mathbb{R})/SL_n(\mathbb{R})$ is path connected. But for this we can use the determinant to define a continuous map

$$\det: GL_n^+(\mathbb{R}) \longrightarrow \mathbb{R}^+ = (0, \infty),$$

which is surjective onto a path connected space. Then the homogeneous space $GL_n^+(\mathbb{R})/SL_n(\mathbb{R})$ is diffeomorphic to \mathbb{R}^+ and hence is path connected. So $GL_n^+(\mathbb{R})$ is path connected. \square

This result implies that

$$GL_n(\mathbb{R}) = GL_n^+(\mathbb{R}) \cup GL_n^-(\mathbb{R})$$

is the decomposition of $GL_n(\mathbb{R})$ into two path connected components.

Example 9.14

For $n \geq 1$, $SO(n)$ is path connected. Hence

$$O(n) = SO(n) \cup O(n)^-$$

is the decomposition of $O(n)$ into two path connected components.

Proof

For $n = 1$, $SO(1) = \{1\}$. So we will assume that $n \geq 2$ and proceed by induction on n . So assume that $SO(n - 1)$ is path connected.

Consider the continuous action of $SO(n)$ on \mathbb{R}^n by left multiplication. The stabiliser of e_n is $SO(n - 1) \leq SO(n)$ thought of as the closed subgroup of matrices of the form

$$\begin{bmatrix} P & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

with $P \in SO(n - 1)$ and $\mathbf{0}$ the $(n - 1) \times 1$ zero matrix. The orbit of e_n is the unit sphere S^{n-1} which is path connected. Since the orbit space is also diffeomorphic to $SO(n)/SO(n - 1)$ we have the inductive step. \square

Example 9.15

For $n \geq 1$, $U(n)$ and $SU(n)$ are path connected.

Proof

For $n = 1$, $U(1)$ is the unit circle in \mathbb{C} while $SU(1) = \{1\}$, so both of these are path connected. Assume that $U(n - 1)$ and $SU(n - 1)$ are path connected for some $n \geq 2$.

Then $U(n)$ and $SU(n)$ act on \mathbb{C}^n by matrix multiplication and by arguments of Chapter 8,

$$\text{Stab}_{U(n)}(\mathbf{e}_n) = U(n - 1), \quad \text{Stab}_{SU(n)}(\mathbf{e}_n) = SU(n - 1).$$

We also have

$$\text{Orb}_{U(n)}(\mathbf{e}_n) = \text{Orb}_{SU(n)}(\mathbf{e}_n) = S^{2n-1},$$

where $S^{2n-1} \subseteq \mathbb{C}^n \cong \mathbb{R}^{2n}$ denotes the unit sphere consisting of unit vectors. Since S^{2n-1} is path connected, we can deduce that $U(n)$ and $SU(n)$ are too, which gives the inductive step. \square

9.3 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 $p: [0, 1] \rightarrow G$ with $p(0) = x$ and $p(1) = y$; we will then write $x \sim_G y$.

Lemma 9.16

\sim_G 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* ,

$$G_g = \{x \in G : x \sim_G g\}.$$

Proposition 9.17

- i) The path component of the identity is a clopen normal subgroup of G , $G_1 \triangleleft G$; hence it is a closed Lie subgroup of dimension $\dim G$.
- ii) The path component G_g agrees with the coset of g with respect to G_1 , $G_g = gG_1 = G_1g$ and is a closed submanifold of G .

Proof

By Proposition 9.8, 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 9.4 shows that each G_g is actually clopen in G .

Let $x, y \in G_1$. Then there are continuous paths $p, q: [0, 1] \rightarrow G$ with $p(0) = 1 = q(0)$, $p(0) = x$ and $q(0) = y$. The product path

$$r: [0, 1] \rightarrow G; \quad r(t) = p(t)q(t)$$

has $r(0) = 1$ and $r(1) = xy$. So $G_1 \leq G$. For $g \in G$, the path

$$s: [0, 1] \rightarrow G; \quad s(t) = gp(t)g^{-1}$$

has $s(0) = 1$ and $s(1) = gxg^{-1}$; hence $G_1 \leq G$. If $z \in gG_1 = G_1g$, then $g^{-1}z \in G_1$ and so there is a continuous path $h: [0, 1] \rightarrow G$ with $h(0) = 1$ and $h(1) = g^{-1}z$. Then the path

$$gh: [0, 1] \rightarrow G; \quad gh(t) = g(h(t))$$

has $gh(0) = g$ and $gh(1) = z$. So each coset gG_1 is path connected, and $gG_1 \subseteq G_g$. To show equality, suppose that g is connected by a path $k: [0, 1] \rightarrow G$ 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 $G_g \subseteq gG_1$.

The remaining details are left to the reader. \square

The quotient group G/G_1 is the *group of path components* of G , which is denoted by $\pi_0 G$.

Example 9.18

We have the following groups of path components:

$$\begin{aligned} \pi_0 \mathrm{SO}(n) &= \pi_0 \mathrm{SL}_n(\mathbb{R}) = \pi_0 \mathrm{SU}(n) = \pi_0 \mathrm{U}(n) = \pi_0 \mathrm{SL}_n(\mathbb{C}) = \pi_0 \mathrm{GL}_n(\mathbb{C}) = \{1\} \\ \pi_0 \mathrm{O}(n) &\cong \pi_0 \mathrm{GL}_n(\mathbb{R}) \cong \{\pm 1\}. \end{aligned}$$

Example 9.19

Let

$$T = \left\{ \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} : \theta \in \mathbb{R} \right\} \leq \mathrm{SO}(3)$$

and let $G = \mathrm{N}_{\mathrm{SO}(3)}(T) \leq \mathrm{SO}(3)$ be the normaliser of T . Then T and G are Lie subgroups of $\mathrm{SO}(3)$ and $\pi_0 G \cong \{\pm 1\}$.

Proof

A straightforward computation shows that

$$N_{SO(3)}(T) = T \cup \left\{ \begin{bmatrix} -\cos \theta & \sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & -1 \end{bmatrix} : \theta \in \mathbb{R} \right\} = T \cup ZT,$$

where

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

Notice that T is isomorphic to the unit circle,

$$T \cong \mathbf{T}; \quad \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \longleftrightarrow e^{\theta i}.$$

This implies that T is path connected and abelian since \mathbf{T} is. The function

$$\varphi: G \longrightarrow \mathbb{R}^\times; \quad \varphi([a_{ij}]) = a_{33}$$

is continuous and

$$\varphi^{-1}\mathbb{R}^+ = T, \quad \varphi^{-1}\mathbb{R}^- = ZT,$$

hence these are clopen subsets. This shows that the path components of G are $G_I = T$ and ZT , so $\pi_0 G \cong \{\pm 1\}$.

Notice that $N_{SO(3)}(T)$ acts by conjugation on T and every element of $T \triangleleft N_{SO(3)}(T)$ acts trivially since T is abelian. Then $\pi_0 G$ acts on T with the action of the non-trivial coset arising from conjugation by the matrix Z , i.e.,

$$Z \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} Z^{-1} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}$$

which agrees with the inverse homomorphism on the unit circle $\mathbf{T} \cong T$. \square

Example 9.20

Let

$$T = \{x1 + yi : x, y \in \mathbb{R}, x^2 + y^2 = 1\} \leqslant \mathrm{Sp}(1),$$

the group of unit quaternions and let $G = N_{\mathrm{Sp}(1)}(T) \leqslant \mathrm{Sp}(1)$ be its normaliser. 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

$$\theta : G \rightarrow \mathbb{R}; \quad \theta(t1 + xi + yj + zk) = y^2 + z^2,$$

is continuous and

$$\theta^{-1}0 = T, \quad \theta^{-1}1 = jT.$$

Hence the path components of G are T, jT , while the group of path components is

$$\pi_0 G = G/T \cong \{\pm 1\}.$$

Under the conjugation action of G on T , every element of T acts 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 $T \cong \mathbb{T}$. □

The significance of such examples will become clearer when we discuss maximal tori and their normalisers in Chapter 10.

9.4 Another Connectivity Result

The following result will be used in Chapter 10.

Proposition 9.21

Let G be a connected Lie group and $H \leq G$ be 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 $\text{inv} : G \rightarrow G$ is a homeomorphism and maps H into itself, by replacing U with $U \cap \text{inv } U$ if necessary, we may assume that inv maps the open neighbourhood into itself, i.e., $\text{inv } U = U$.

For $k \geq 1$, consider

$$U^k = \{u_1 \cdots u_k \in G : u_j \in U\} \subseteq H.$$

Notice that $\text{inv } U^k = U^k$. Also, $U^k \subseteq G$ is open since for $u_1, \dots, u_k \in U$,

$$u_1 \cdots u_k \in L_{u_1 \cdots u_{k-1}} U \subseteq U^k$$

where

$$L_{u_1 \cdots u_{k-1}} U = L_{(u_1 \cdots u_{k-1})^{-1}}^{-1} U$$

is an open subset of G . Then

$$V = \bigcup_{k \geq 1} U^k \subseteq H$$

satisfies $\text{inv } V = V$.

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 that $g = v_1 \cdots v_s u_1^{-1} \cdots u_r^{-1} \in V$, contradicting the assumption on g .

So V is a non-empty clopen subset of G , which is connected. Hence $G - V = \emptyset$, and therefore $V = G$, which also implies that $H = G$. \square

EXERCISES

9.1. Let G be a Lie group.

- a) Using the fact that a manifold is separable, show that $\pi_0 G$ is a countable group.
- b) If G is compact, show that $\pi_0 G$ is a finite group.

9.2. Let $\text{SL}_n(\mathbb{R})$ act smoothly on \mathbb{R}^n by matrix multiplication.

- a) Find $\text{Stab}_{\text{SL}_n(\mathbb{R})}(e_n)$ and $\text{Orb}_{\text{SL}_n(\mathbb{R})}(e_n)$.
- b) Identify the homogeneous space $\text{SL}_n(\mathbb{R}) / \text{Stab}_{\text{SL}_n(\mathbb{R})}(e_n)$ and show that it is path connected if $n \geq 2$. Use this to give another proof that $\text{GL}_n(\mathbb{R})$ has two path components.

9.3. Let $\text{SL}_n(\mathbb{C})$ act smoothly on \mathbb{C}^n by matrix multiplication.

- a) Find $\text{Stab}_{\text{SL}_n(\mathbb{C})}(e_n)$ and $\text{Orb}_{\text{SL}_n(\mathbb{C})}(e_n)$.
- b) Identify the homogeneous space $\text{SL}_n(\mathbb{C}) / \text{Stab}_{\text{SL}_n(\mathbb{C})}(e_n)$ and show that it is path connected. Use this to prove that $\text{SL}_n(\mathbb{C})$ is path connected.
- c) By making use of the determinant $\det: \text{GL}_n(\mathbb{C}) \rightarrow \mathbb{C}^\times$, deduce that $\text{GL}_n(\mathbb{C})$ is path connected.

9.4. Let $A \in \mathrm{GL}_n(\mathbb{R})$.

- 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 $S_1^2 = S$.
- Show that $S_1^{-1}A$ is orthogonal.
- If $PR = QS$ where P, Q are positive definite real symmetric and $R, S \in \mathrm{O}(n)$, show that $P^2 = Q^2$.
- Let S_2 be a positive definite real symmetric matrix for which $S_2^2 = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. Show that $S_2 = \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$.
- 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 $R \in \mathrm{SO}(n)$.
- Show that the homogeneous space $\mathrm{GL}_n^+(\mathbb{R})/\mathrm{SO}(n)$ is path connected. Using Example 9.14, deduce that $\mathrm{GL}_n^+(\mathbb{R})$ is path connected.
- 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 9.15, deduce that $\mathrm{GL}_n(\mathbb{C})$ is path connected.

[Such factorisations are called polar decompositions of A and B .]

9.5. For $\mathbf{k} = \mathbb{R}, \mathbb{C}$ and $n \geq 1$, the affine group $\mathrm{Aff}_n(\mathbf{k})$ acts on \mathbf{k}^n as explained in Chapter 1.

- Find $\mathrm{Stab}_{\mathrm{Aff}_n(\mathbf{k})}(\mathbf{0})$ and $\mathrm{Orb}_{\mathrm{Aff}_n(\mathbf{k})}(\mathbf{0})$.
- Show that the affine group $\mathrm{Aff}_n(\mathbb{R})$ has two path components, while $\mathrm{Aff}_n(\mathbb{C})$ is path connected.

9.6. a) If $n \geq 1$, use the Gram–Schmidt orthogonalisation process to show that every matrix $A \in \mathrm{GL}_n(\mathbb{R})$ can be expressed in the form $A = QR$, where $Q \in \mathrm{O}(n)$ and $R \in \mathrm{UT}_n(\mathbb{R}) \leqslant \mathrm{GL}_n(\mathbb{R})$ the subgroup of upper triangular matrices. Deduce that the homogeneous space $\mathrm{GL}_n(\mathbb{R})/\mathrm{UT}_n(\mathbb{R})$ is compact. Is it connected?

- b) If $A \in \mathrm{SL}_n(\mathbb{R})$, show that we can take $Q \in \mathrm{SO}(n)$ and $\det R = 1$. Deduce that the homogeneous space $\mathrm{SL}_n(\mathbb{R})/(\mathrm{UT}_n(\mathbb{R}) \cap \mathrm{SL}_n(\mathbb{R}))$ is compact. Is it connected?

c) What is the relationship between these two homogeneous spaces?

9.7. Work through the analogue of the previous question for $\mathrm{GL}_n(\mathbb{C})$ after first showing that a matrix $A \in \mathrm{GL}_n(\mathbb{C})$ can be expressed in the form $A = QR$, where $Q \in \mathrm{U}(n)$ and $R \in \mathrm{UT}_n(\mathbb{C})$. Discuss the compactness and connectedness of the homogeneous spaces $\mathrm{GL}_n(\mathbb{C})/\mathrm{UT}_n(\mathbb{C})$ and $\mathrm{SL}_n(\mathbb{C})/(\mathrm{UT}_n(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C}))$.

9.8. a) Let $m \geq 1$. Using Corollary 8.24, show how to identify the hom-

ogeneous spaces $\mathrm{GL}_{2m}(\mathbb{R})/\mathrm{Symp}_{2m}(\mathbb{R})$ and $\mathrm{GL}_{2m}(\mathbb{R})/\mathrm{Symp}'_{2m}(\mathbb{R})$ with the subspace $\Sigma_{2m} \subseteq \mathrm{M}_{2m}(\mathbb{R})$ of all $2m \times 2m$ non-singular real skew symmetric matrices. Use these identifications to prove Proposition 1.42.

- b) Use the Pfaffian function $\mathrm{pf}: \Sigma_{2m} \rightarrow \mathbb{R}^\times$ to prove that Σ_{2m} is not connected.
c) Fill in the details of Remark 1.43, in particular verify the inclusions $\mathrm{Symp}_{2m}(\mathbb{R}) \leq \mathrm{SL}_{2m}(\mathbb{R})$ and $\mathrm{Symp}'_{2m}(\mathbb{R}) \leq \mathrm{SL}_{2m}(\mathbb{R})$.

9.9. Δ Since $\mathrm{Symp}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$, this is a connected group. The aim of this exercise is to prove that $\mathrm{Symp}_{2n}(\mathbb{R})$ is connected for all $n \geq 1$. Let ω be the standard symplectic form on \mathbb{R}^{2n} , hence $\omega(x, y) = x^T J_{2n} y$ where J_{2n} is defined in Section 1.5.

- a) For $n \geq 1$, show that

$$\Psi_{2n} = \{(u, v) \in \mathbb{R}^{2n} \times \mathbb{R}^{2n} : \omega(u, v) = 1\} \subseteq \mathbb{R}^{2n} \times \mathbb{R}^{2n}$$

is a closed subset.

- b) Show that the function

$$\varphi: \mathrm{Symp}_{2n}(\mathbb{R}) \rightarrow \Psi_{2n}; \quad \varphi([a_1, \dots, a_{2n}]) = (a_{2n-1}, a_{2n})$$

is continuous and surjective, hence identify the homogeneous space $\mathrm{Symp}_{2n}(\mathbb{R})/\mathrm{Symp}_{2n-2}(\mathbb{R})$ with Ψ_{2n} .

- c) Show that $\mathrm{Symp}_{2n}(\mathbb{R})$ acts continuously on Ψ_{2n} by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot (u, v) = (au + bv, cu + dv).$$

Show that this action has only one orbit. Deduce that Ψ_{2n} is connected.

- d) Prove by induction on n that $\mathrm{Symp}_{2n}(\mathbb{R})$ is connected.

9.10. Let

$$T = \{\mathrm{diag}(u, v) : |u| = |v| = 1\} \leq \mathrm{U}(2).$$

Find the normaliser $N_{\mathrm{U}(2)}(T) \leq \mathrm{U}(2)$ and the group of path components $\pi_0 N_{\mathrm{U}(2)}(T)$.

9.11. Let

$$G = \left\{ \begin{bmatrix} A & O_{2,2} \\ O_{2,2} & B \end{bmatrix} : A, B \in \mathrm{U}(2), \det A = \det B \right\} \leq \mathrm{SU}(4),$$

$$T' = \left\{ \begin{bmatrix} \mathrm{diag}(u, v) & O_{2,2} \\ O_{2,2} & \mathrm{diag}(w, uv\bar{w}) \end{bmatrix} : |u| = |v| = |w| = 1 \right\} \leq G.$$

Find the normaliser $N_G(T') \leq G$ and the group of path components $\pi_0 N_G(T')$.

Part III

Compact Connected Lie Groups and their Classification

Maximal Tori in Compact Connected Lie Groups

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 turns out to be central to the classification of simple compact connected Lie groups. From Chapter 9 we know that many familiar examples of compact matrix groups are path connected.

Although we state results for arbitrary Lie groups we will often give proofs for matrix groups. However, there is no loss of generality in assuming this because of the following important result whose proof requires *Haar measure* and integration on compact Lie groups.

Theorem 10.1

Let G be a compact Lie group. Then for some $m, n \geq 1$, there are injective Lie homomorphisms $G \rightarrow O(m)$ and $G \rightarrow U(n)$, hence G is a matrix group.

10.1 Tori

The familiar *circle group*

$$\mathbf{T} = \{z \in \mathbf{C} : |z| = 1\} \leq \mathbf{C}^\times$$

is a matrix group since $\mathbf{C}^\times = \mathrm{GL}_1(\mathbf{C})$; its Lie algebra is the 1-dimensional real abelian subalgebra

$$\{ti \in \mathbf{C} : t \in \mathbb{R}\} \leqslant \mathbf{C}.$$

Notice that \mathbf{T} is a closed and bounded subset of \mathbf{C} , hence it is compact; it is also path connected and abelian.

Definition 10.2

For each $r \geqslant 1$,

$$\mathbf{T}^r = \{\mathrm{diag}(z_1, \dots, z_r) : |z_1| = \dots = |z_r| = 1\} \leqslant \mathrm{GL}_r(\mathbf{C})$$

is the *standard torus of rank r*. More generally, a *torus of rank r* is any Lie group isomorphic to \mathbf{T}^r .

We will often view elements of \mathbf{T}^r as sequences of complex numbers (z_1, \dots, z_r) with $|z_1| = \dots = |z_r| = 1$, this corresponds to an identification

$$\mathbf{T}^r \cong \underbrace{\mathbf{T} \times \dots \times \mathbf{T}}_{r \text{ factors}} \leqslant (\mathbf{C}^\times)^r.$$

\mathbf{T}^r is an abelian Lie group of dimension r whose Lie algebra is the abelian Lie algebra

$$\{\mathrm{diag}(t_1 i, \dots, t_r i) : t_1, \dots, t_r \in \mathbb{R}\}$$

which can also be identified with

$$\{(t_1 i, \dots, t_r i) : \forall k, t_k \in \mathbb{R}\}.$$

Proposition 10.3

Let T be a torus. Then T is a compact, path connected and abelian Lie group.

Proof

Since the circle \mathbf{T} is compact and abelian the same is true for \mathbf{T}^r and hence for any torus.

If $(z_1, \dots, z_r) \in \mathbf{T}^r$, set $z_k = e^{\theta_k i}$. Then there is a continuous path

$$p: [0, 1] \longrightarrow \mathbf{T}^r; \quad p(t) = (e^{t\theta_1 i}, \dots, e^{t\theta_r i}),$$

with $p(0) = (1, \dots, 1)$ and $p(1) = (z_1, \dots, z_r)$. So \mathbf{T}^r is path connected and therefore every torus is path connected. \square

Our next result characterises which compact Lie groups are tori.

Theorem 10.4

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 10.3. So we need to show that when H is connected and abelian it is a torus since by Proposition 9.8 it would be path connected.

Suppose that $\dim H = r$ and let \mathfrak{h} be the Lie algebra of H ; then $\dim \mathfrak{h} = r$. Let $X, Y \in \mathfrak{h}$. From the definition of the Lie bracket in the proof of Theorem 3.18,

$$[X, Y] = \frac{d}{ds}_{|s=0} \frac{d}{dt}_{|t=0} \exp(sX) \exp(tY) \exp(-sX) = 0$$

since $\exp(sX), \exp(tY) \in H$. Because H is abelian, this gives

$$\exp(sX) \exp(tY) \exp(-sX) = \exp(tY),$$

showing that all Lie brackets in \mathfrak{h} are zero.

Now consider the exponential map $\exp: \mathfrak{h} \rightarrow H$. Propositions 2.2 and 2.1 imply that for all $X, Y \in \mathfrak{h}$,

$$\exp(X) \exp(Y) = \exp(X + Y), \quad \exp(-X) = \exp(X)^{-1}.$$

So $\exp \mathfrak{h} = \text{im } \exp \subseteq H$ is a subgroup. By Proposition 2.4, $\exp \mathfrak{h}$ also contains a neighbourhood of the identity so by Proposition 9.21 we have $\exp \mathfrak{h} = H$.

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 $\{v_1, \dots, v_s\}$ for some $s \leq r$. Extending this to an \mathbb{R} -basis $\{v_1, \dots, v_s, v_{s+1}, \dots, v_r\}$ of \mathfrak{h} we obtain isomorphisms of Lie groups

$$\exp(\mathfrak{h}) \cong \mathfrak{h}/K \cong \mathbb{R}^s / \mathbb{Z}^s \times \mathbb{R}^{r-s}.$$

The right-hand term is only compact if $s = r$, therefore K contains a basis of \mathfrak{h} and

$$\mathbb{R}^r / \mathbb{Z}^r \cong \mathfrak{h}/K \cong H.$$

Since $T \cong \mathbb{R}/\mathbb{Z}$, this shows that H is a torus. \square

In this proof we made use of the essential idea contained in the next result.

Proposition 10.5

Let T be a torus of rank r . Then the exponential map $\exp: \mathfrak{t} \rightarrow T$ 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 7.36, we met the idea of a topological generator of the circle group. It turns out that all tori have such generators.

Definition 10.6

Let G be a Lie group. Then an element $g \in G$ is called a *topological generator* or just a *generator* of G if the cyclic subgroup $\langle g \rangle \leq G$ is dense in G , i.e., $\overline{\langle g \rangle} = G$.

Proposition 10.7

Every torus T has a generator.

Proof

Without loss of generality we can assume that $T = \mathbb{R}^r/\mathbb{Z}^r$ and will write elements in the form $[x_1, \dots, x_r] = (x_1, \dots, x_r) + \mathbb{Z}^r$. The group operation is then addition. Let U_1, U_2, U_3, \dots be a countable base for the topology on T .

A cube of side $\epsilon > 0$ in T is a subset of the form

$$C([u_1, \dots, u_r], \epsilon) = \{[x_1, \dots, x_r] \in T : |x_k - u_k| < \epsilon/2 \forall k\},$$

for some $[u_1, \dots, u_r] \in T$. Such a cube is the image of a cube in \mathbb{R}^r under the quotient map $\mathbb{R}^r \rightarrow T$.

Let $C_0 \subseteq T$ be a cube of side $\epsilon > 0$. 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 $0 \leq k \leq m$, there is an integer N_k satisfying $N_k \epsilon_k > 1$ and $N_k C_k \subseteq U_k$. 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 $z = [z_1, \dots, z_r] \in \bigcap_{k \geq 1} C_k$, we have $N_k z \in C_k$ for each k , hence the powers of z are dense in T , so z is a generator of T . \square

10.2 Maximal Tori in Compact Lie Groups

Definition 10.8

Let G be a Lie group and $T \leq G$ be a closed subgroup which is a torus. Then T is *maximal* or a *maximal torus* in G if the only torus $T' \leq G$ for which $T \leq T'$ is T .

Remark 10.9

It is easy to see that every torus $T \leq G$ is contained in a maximal torus $\bar{T} \leq G$ where $\dim T \leq \dim \bar{T} \leq \dim G$. Later we will also see that if G is connected then every element $g \in G$ is contained in a maximal torus.

We will now describe some important examples of maximal tori in compact connected matrix groups. 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 $n \geq 1$, and $\theta_i \in [0, 2\pi)$ ($i = 1, \dots, n$), let

$$R_{2n}(\theta_1, \dots, \theta_n) = \begin{bmatrix} R(\theta_1) & O & \cdots & \cdots & \cdots & O \\ O & R(\theta_2) & O & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ O & \cdots & \cdots & \cdots & O & R(\theta_n) \end{bmatrix},$$

$$R_{2n+1}(\theta_1, \dots, \theta_n) = \begin{bmatrix} R(\theta_1) & O & \cdots & \cdots & \cdots & \cdots & O \\ O & R(\theta_2) & O & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & O & R(\theta_n) & O \\ O & \cdots & \cdots & \cdots & \cdots & O & 1 \end{bmatrix},$$

where each entry marked O is an appropriately sized block so that these are matrices of sizes $2n \times 2n$ and $(2n + 1) \times (2n + 1)$ respectively. Then $R_{2n}(\theta_1, \dots, \theta_n) \in \mathrm{SO}(2n)$ and $R_{2n+1}(\theta_1, \dots, \theta_n) \in \mathrm{SO}(2n + 1)$.

It is obvious that $\mathbf{T} \cong \mathrm{U}(1)$ as Lie groups. After first identifying \mathbb{C} with \mathbb{R}^2 as real vector spaces using the bases $\{1, i\}$ and $\{e_1, e_2\}$, we also obtain an isomorphism of Lie groups

$$\mathbf{T} \longrightarrow \mathrm{SO}(2); \quad e^{\theta i} \longmapsto R_2(\theta).$$

The maximal tori listed in the next result are usually regarded as the *standard maximal tori* for the corresponding groups.

Proposition 10.10

Each of the following is a maximal torus in the stated group:

$$\begin{aligned} \{R_{2n}(\theta_1, \dots, \theta_n) : \forall k \theta_k \in [0, 2\pi)\} &\leqslant \mathrm{SO}(2n), \\ \{R_{2n+1}(\theta_1, \dots, \theta_n) : \forall k \theta_k \in [0, 2\pi)\} &\leqslant \mathrm{SO}(2n+1), \\ \{\mathrm{diag}(z_1, \dots, z_n) : \forall k |z_k| = 1\} &\leqslant \mathrm{U}(n), \\ \{\mathrm{diag}(z_1, \dots, z_n) : \forall k |z_k| = 1, z_1 \cdots z_n = 1\} &\leqslant \mathrm{SU}(n), \\ \{\mathrm{diag}(z_1, \dots, z_n) : \forall k z_k \in \mathbb{C}, |z_k| = 1\} &\leqslant \mathrm{Sp}(n), \end{aligned}$$

We now begin to study the structure of compact Lie groups in terms of their maximal tori. Throughout the rest of the section, let G be a compact connected Lie group and $T \leqslant G$ be a maximal torus. The next result is fundamental to understanding the relationship between the different maximal tori in G .

Theorem 10.11

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 of this depends on the powerful and important *Lefschetz Fixed Point Theorem* from Algebraic Topology and we only give a sketch indicating it is used. For details on this result, the interested reader might usefully consult an introductory book on Algebraic Topology such as [9, 20].

The quotient space G/T is a compact space and each element $g \in G$ gives rise to a continuous map

$$\mu_g: G/T \longrightarrow G/T; \quad \mu_g(xT) = (gx)T = gxT.$$

Since G is path connected, there is a continuous map

$$p: [0, 1] \times G/T \longrightarrow G/T;$$

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 that the Euler characteristic $\chi(G/T)$ is non-zero. It can indeed 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 $g \in xTx^{-1}$. \square

Theorem 10.12

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 10.7, T' has a generator t say. By Theorem 10.11, there is a $y \in G$ such that $T' \leqslant yTy^{-1}$. As T' is a maximal torus and yTy^{-1} is a torus, we must have $T' = yTy^{-1}$. \square

Our next result gives some important special cases related to the examples of Proposition 10.10. Notice that if $A \in \mathrm{SO}(m)$, $A^{-1} = A^T$, while if $B \in \mathrm{U}(m)$, $B^{-1} = B^*$.

Theorem 10.13 (Principle Axis Theorem)

In each of the following matrix groups every element is conjugate to one of the stated form.

$$\mathrm{SO}(2n): \quad R_{2n}(\theta_1, \dots, \theta_n), \quad \forall k \theta_k \in [0, 2\pi);$$

$$\mathrm{SO}(2n+1): \quad R_{2n+1}(\theta_1, \dots, \theta_n), \quad \forall k \theta_k \in [0, 2\pi);$$

$$\mathrm{U}(n): \quad \mathrm{diag}(z_1, \dots, z_n), \quad \forall k z_k \in \mathbb{C}, |z_k| = 1;$$

$$\mathrm{SU}(n): \quad \mathrm{diag}(z_1, \dots, z_n), \quad \forall k z_k \in \mathbb{C}, |z_k| = 1, z_1 \cdots z_n = 1;$$

$$\mathrm{Sp}(n): \quad \mathrm{diag}(z_1, \dots, z_n), \quad \forall k z_k \in \mathbb{C}, |z_k| = 1.$$

There are related 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

$$\mathrm{Ad}_g: G \longrightarrow \mathfrak{g}; \quad \mathrm{Ad}_g(t) = g t g^{-1}.$$

Proposition 10.14

Suppose that $g \in G$ and $H, K \leqslant G$ are Lie subgroups with Lie algebras $\mathfrak{h}, \mathfrak{k} \leqslant \mathfrak{g}$. If $gHg^{-1} = K$, then $\mathrm{Ad}_g \mathfrak{h} = \mathfrak{k}$.

Proof

By definition, for each $x \in \mathfrak{h}$ there is a curve $\gamma: (-\varepsilon, \varepsilon) \longrightarrow H$ with $\gamma(0) = 1$ and $\gamma'(0) = x$. Since $t \mapsto g\gamma(t)g^{-1}$ is a curve in K ,

$$\mathrm{Ad}_g(x) = \frac{d}{dt} g\gamma(t)g^{-1} \Big|_{t=0} \in \mathfrak{k},$$

so $\text{Ad}_g \mathfrak{h} \leqslant \mathfrak{k}$. Since Ad_g is injective and $\dim \mathfrak{h} = \mathfrak{k}$, we must have $\text{Ad}_g \mathfrak{h} = \mathfrak{k}$. \square

If $x, y \in \mathfrak{g}$ and $y = \text{Ad}_g(x)$ we will sometimes say that x is *conjugate* to y with respect to G . Being conjugate in this sense gives an equivalence relation on \mathfrak{g} .

For $t \in \mathbb{R}$, let

$$R'(t) = \begin{bmatrix} 0 & -t \\ t & 0 \end{bmatrix}$$

and

$$R'_{2n}(t_1, \dots, t_n) = \begin{bmatrix} R'(t_1) & O & \cdots & \cdots & \cdots & O \\ O & R'(t_2) & O & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ O & \cdots & \cdots & \cdots & O & R'(t_n) \end{bmatrix},$$

$$R'_{2n+1}(t_1, \dots, t_n) = \begin{bmatrix} R'(t_1) & O & \cdots & \cdots & \cdots & \cdots & O \\ O & R'(t_2) & O & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & O & R'(t_n) & O \\ O & \cdots & \cdots & \cdots & \cdots & O & 1 \end{bmatrix}.$$

Then $R'_{2n}(t_1, \dots, t_n) \in \mathfrak{so}(2n)$ and $R'_{2n+1}(t_1, \dots, t_n) \in \mathfrak{so}(2n+1)$.

Theorem 10.15 (Principle Axis Theorem for Lie algebras)

For each of the following Lie algebras \mathfrak{g} , every element $x \in \mathfrak{g}$ is conjugate in G to one of the stated form.

- | | | |
|-------------------------|-------------------------------------|--|
| $\mathfrak{so}(2n)$: | $R'_{2n}(t_1, \dots, t_n),$ | $\forall k \theta_k \in [0, 2\pi);$ |
| $\mathfrak{so}(2n+1)$: | $R'_{2n+1}(t_1, \dots, t_n),$ | $\forall k \theta_k \in [0, 2\pi);$ |
| $\mathfrak{u}(n)$: | $\text{diag}(t_1 i, \dots, t_n i),$ | $\forall k t_k \in \mathbb{R};$ |
| $\mathfrak{su}(n)$: | $\text{diag}(t_1 i, \dots, t_n i),$ | $\forall k t_k \in \mathbb{R}, t_1 + \dots + t_n = 1;$ |
| $\mathfrak{sp}(n)$: | $\text{diag}(t_1 i, \dots, t_n i),$ | $\forall k t_k \in \mathbb{R}.$ |

We can now give an important result which we have already seen is true for many familiar examples.

Theorem 10.16

For a compact connected Lie group G , the exponential map $\exp: \mathfrak{g} \rightarrow G$ is surjective.

Proof

Let $T \leq G$ be a maximal torus. By Theorem 10.11, every element $g \in G$ is conjugate to an element $xgx^{-1} \in T$. By Proposition 10.5, $xgx^{-1} = \exp(t)$ for some $t \in \mathfrak{t}$, hence

$$g = x^{-1} \exp(t)x = \exp(\text{Ad}_x(t)),$$

where $\text{Ad}_x(t) \in \mathfrak{g}$. So $g \in \exp \mathfrak{g}$. Therefore $\exp \mathfrak{g} = G$. \square

10.3 The Normaliser and Weyl Group of a Maximal Torus

Given Theorem 10.11, we can continue to develop the general theory for a compact connected Lie group G .

Proposition 10.17

Let $A \leq G$ be a compact abelian Lie group and suppose that $A_1 \leq A$ 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 10.4, hence it has a generator a_0 by Proposition 10.7. 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 form $(gb)^{kd+r}$ 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 $T \leq G$ be a maximal torus. By Theorem 10.11, 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. \square

Proposition 10.18

Let $A \leq G$ be a connected abelian subgroup and let $g \in G$ commute with all the elements of A . Then there is a torus $T \leq G$ containing the subgroup $\langle A, g \rangle \leq G$ 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 10.4. Now consider the abelian subgroup $\langle A, g \rangle \leq G$ generated by A and g , whose closure $B \leq G$ is again compact and abelian. If the connected component of the identity is $B_1 \leq B$ then B_1 has finitely many cosets by compactness, and these are of the form $g^r B_1$ ($r = 0, 1, \dots, d - 1$) for some d . By Proposition 10.17, $\langle A, g \rangle$ is contained in a torus. \square

Theorem 10.19

Let $T \leq G$ be a maximal torus and let $T \leq A \leq G$ where A is abelian. Then $A = T$. Equivalently, every maximal torus is a maximal abelian subgroup.

Proof

For each element $g \in A$, Proposition 10.18 implies that there is a torus containing $\langle T, g \rangle$, but by the maximality of T this must equal T . Hence $A = T$. \square

We have now established that every maximal torus is also a maximal abelian subgroup, and that any two maximal tori are conjugate in G . Next we focus on the relationship of a maximal torus $T \leq G$ to the rest of G .

Recall that for a subgroup $H \leq G$, the *normaliser* of H in G is the smallest subgroup of G in which H is normal,

$$N_G(H) = \{g \in G : gHg^{-1} = H\}.$$

Then $N_G(H) \leq G$ is a closed subgroup of G and so is also compact. It also contains H and its closure in G as normal subgroups. There is a continuous left action of $N_G(H)$ on H by conjugation, i.e., for $g \in N_G(H)$ 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 $W_G(T) = N_G(T)/T$ is known as the *Weyl group* of T in G . Notice that the connected component of the identity in $N_G(T)$ contains T ; in fact it agrees with T as we will soon see.

Lemma 10.20

Let $T \leq G$ be a torus and let $Q \leq N_G(T)$ 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 10.5, the exponential map is a surjective group homomorphism $\exp: t \rightarrow T$ whose kernel is a discrete subgroup. In fact, there is a commutative diagram

$$\begin{array}{ccccc} \ker \exp & \longrightarrow & t & \longrightarrow & T \\ \cong \downarrow & & \cong \downarrow & & \cong \downarrow \\ \mathbb{Z}^r & \longrightarrow & \mathbb{R}^r & \longrightarrow & \mathbb{R}^r / \mathbb{Z}^r \end{array}$$

in which all the maps are the evident ones.

Now a Lie group automorphism $\alpha: T \rightarrow T$ lifts to homomorphism $\tilde{\alpha}: t \rightarrow t$ restricting to an isomorphism $\tilde{\alpha}_0: \ker \exp \rightarrow \ker \exp$. Indeed, since each element of $\ker \exp \cong \mathbb{Z}^r$ is uniquely divisible in $t \cong \mathbb{R}^r$, continuity implies that $\tilde{\alpha}_0$ determines $\tilde{\alpha}$ on t . But the automorphism group $\text{Aut}(\ker \exp) \cong \text{Aut}(\mathbb{Z}^r)$ 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 $\text{Aut}(\ker \exp)$. Hence the action of Q on T is trivial. \square

This result shows that $N_G(T)_1$, the connected component of the identity in $N_G(T)$, acts trivially on the torus T . In fact, if $g \in N_G(T)$ acts trivially on T then it commutes with all the elements of T , so by Theorem 10.19 g is in T . Thus T consists of all the elements of G with this property, i.e.,

$$T = \{g \in G : gxg^{-1} = x \ \forall x \in T\}. \quad (10.1)$$

In particular, we have $N_G(T)_1 = T$. We can now state an important omnibus result on the Weyl group.

Theorem 10.21

Let $T \leqslant G$ be a maximal torus.

i) The Weyl group $W_G(T)$ is isomorphic to the group of components of the normaliser,

$$W_G(T) = N_G(T)/T \cong \pi_0 N_G(T).$$

Hence $W_G(T)$ is finite.

ii) $W_G(T)$ acts on T by conjugation,

$$gT \cdot x = gxg^{-1}.$$

This action on T is faithful, i.e., the coset $gT \in N_G(T)/T$ acts trivially on T if and only if $g \in T$.

Proof

- i) $N_G(T)$ has finitely many cosets of T since it is closed, hence compact, so each coset is clopen.
- ii) The faithfulness of the action follows from Equation (10.1). \square

Proposition 10.22

Let $T \leq G$ be a maximal torus and $x, y \in T$. If x, y are conjugate in G then they are conjugate in $N_G(T)$, hence there is an element $w \in W_G(T)$ for which $y = w \cdot x$.

Proof

Suppose that $y = gxg^{-1}$. Then the centraliser $Z_G(y) \leq G$ 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 $H = Z_G(y)_1$, the connected component of the identity in $Z_G(y)$; this is a closed subgroup of G since it is closed in $Z_G(y)$. Then as T, gTg^{-1} are connected subgroups of $Z_G(y)$ 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 10.11 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 $h^{-1}g \in N_G(T)$ and

$$(h^{-1}g)x(h^{-1}g)^{-1} = h^{-1}yh = y.$$

Now setting $w = h^{-1}g \in W_G(T)$ we obtain the desired result. \square

10.4 The Centre of a Compact Connected Lie Group

Recall that the centre of G is the closed normal subgroup

$$Z(G) = \{g \in G : \forall x \in G, gx = xg\} \triangleleft G.$$

Theorem 10.23

The centre of G is the intersection of all the maximal tori in G ,

$$Z(G) = \bigcap_{\substack{T \leq G \\ \text{a maximal torus}}} T.$$

Proof

If $c \in Z(G)$ then $c \in T$ for some maximal torus $T \leq G$. So for $g \in G$,

$$c = gcg^{-1} \in gTg^{-1},$$

and by Theorem 10.12,

$$c \in \bigcap_{\substack{T \leq G \\ \text{a maximal torus}}} T.$$

Hence

$$Z(G) \leq \bigcap_{\substack{T \leq G \\ \text{a maximal torus}}} T.$$

Conversely, if

$$c \in \bigcap_{\substack{T \leq G \\ \text{a maximal torus}}} T,$$

then for any $x \in G$, $x \in T'$ for some maximal torus $T' \leq G$, hence c, x are both elements of the abelian group T' . So c and x commute. This shows that $c \in Z(G)$. \square

Theorem 10.23 often proves useful when finding the centre of a compact connected Lie group. For example, the centres of $SO(2n)$ and $SO(2n+1)$ lie in the maximal tori of Proposition 10.10, and it is easy to find which elements of these are central in the whole groups. We illustrate this by determining the centre of $SO(4)$.

Consider the element

$$R_4(\theta_1, \theta_2) = \begin{bmatrix} \cos \theta_1 & -\sin \theta_1 & 0 & 0 \\ \sin \theta_1 & \cos \theta_1 & 0 & 0 \\ 0 & 0 & \cos \theta_2 & -\sin \theta_2 \\ 0 & 0 & \sin \theta_2 & \cos \theta_2 \end{bmatrix} \in Z(SO(4)).$$

For $R_4(\theta_1, \theta_2)$ to commute with $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ we must have $\sin \theta_1 = 0 = \sin \theta_2$, so $\cos \theta_1 = \pm 1$ and $\cos \theta_2 = \pm 1$. Hence

$$R_4(\theta_1, \theta_2) = \begin{bmatrix} \pm I_2 & O_2 \\ O_2 & \pm I_2 \end{bmatrix}.$$

Commutativity with $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ forces $R_4(\theta_1, \theta_2) = \pm I_4$. This also illustrates the case $n = 4$ of Proposition 5.31.

EXERCISES

- 10.1. Show that there are exactly two Lie isomorphisms $T \rightarrow T$, but infinitely many Lie isomorphisms $T^r \rightarrow T^r$ when $r \geq 2$.
- 10.2. Prove that in a Lie group G every torus $T \leq G$ is contained in a maximal torus in G , thus verifying the first part of Remark 10.9. Explain why the second part of this remark is false when G is not connected, i.e., not every element of G is contained in a torus.
- 10.3. a) Show that the subgroup $T \leq U(2)$ consisting of all the diagonal matrices is a maximal torus of $U(2)$. Determine the normaliser $N_{U(2)}(T)$, its group of components $\pi_0 N_{U(2)}(T)$ and its action by conjugation on T .
 b) Show that the subgroup of diagonal matrices in $SU(3)$ is a maximal torus and determine its normaliser and group of path components and describe the conjugation action of the latter on this torus.
 c) Let

$$T_2 = \{\text{diag}(u, v) \in \text{Sp}(2) : u, v \in \mathbb{C}\} \leq \text{Sp}(2).$$

Show that T_2 is a maximal torus of $\text{Sp}(2)$. Making use of Example 9.20, determine $N_{\text{Sp}(2)}(T_2)$ and $\pi_0 N_{\text{Sp}(2)}(T_2)$ and describe its conjugation action on T_2 .

- 10.4. Show that the group

$$A = \{(\cos \theta_1 + \sin \theta_1 e_1 e_2)(\cos \theta_2 + \sin \theta_2 e_3 e_4) \cdots (\cos \theta_n + \sin \theta_n e_{2n-1} e_{2n}) : \theta_1, \dots, \theta_n \in [0, 2\pi)\}$$

is a maximal torus in each of the spinor groups $\text{Spin}(2n)$ and $\text{Spin}(2n+1)$.

For small values of n , determine the normalisers and Weyl groups of A in $\text{Spin}(2n)$ and $\text{Spin}(2n+1)$. Find the conjugation action of each Weyl group on A .

How are these maximal tori related under the double covering maps $\rho: \text{Spin}(n) \rightarrow \text{SO}(m)$ of Section 5.3 to the maximal tori of $\text{SO}(2n)$ and $\text{SO}(2n+1)$ given by Proposition 10.10?

10.5. a) For $n \geq 1$, show that

$$\{\text{diag}(\varepsilon_1, \dots, \varepsilon_n) : \varepsilon_1, \dots, \varepsilon_n = \pm 1\} \leq O(n)$$

is a maximal abelian subgroup of $O(n)$.

b) Let $T_{2n} \leq \text{SO}(2n)$ and $T_{2n+1} \leq \text{SO}(2n+1)$ be the maximal tori of Proposition 10.10 and

$$T'_{2n} = T_{2n}, \quad T'_{2n+1} = T_{2n+1} \cup \text{diag}(1, \dots, 1, -1)T_{2n+1}.$$

Show that $T'_{2n} \leq O(2n)$ and $T'_{2n+1} \leq O(2n+1)$ are maximal abelian subgroups.

c) Explain why these results are compatible with those of Chapter 10.

Semi-simple Factorisation

In this chapter we begin our discussion of the classification theory of compact connected Lie groups. Our aim is to introduce the main ideas involved and eventually to describe the outcome of this classification in terms of root systems and Dynkin diagrams. The reader interested in seeing more details might usefully consult some of the books [1, 2, 4, 7, 13, 14, 25] which focus on various aspects of this theory.

Assumption 11.1

Throughout this chapter, unless otherwise specified, G will denote a compact connected Lie group with Lie algebra \mathfrak{g} . By Theorem 10.1 we may assume that $G \leq U(n)$ for some n and $\mathfrak{g} \leq u(n)$ is a Lie subalgebra. This allows us to give proofs valid for matrix subgroups of unitary groups.

11.1 An Invariant Inner Product

In order to discuss the semi-simple decomposition of a compact connected Lie group and later of its adjoint representation, we first need to introduce an \mathbb{R} -bilinear inner product on its Lie algebra. We do this by first defining such an inner product on the real Lie algebra $u(n) = \text{Sk-Herm}_n$, then restricting it to a subalgebra.

We define $(\cdot | \cdot)$ by

$$(X | Y) = -\text{tr}(XY) = \text{tr}(X^*Y) \in \mathbb{C} \quad (X, Y \in \mathfrak{u}(n)). \quad (11.1)$$

Then $(\cdot | \cdot)$ is symmetric since $\text{tr}(YX) = \text{tr}(XY)$. Using standard properties of the trace together with the definition of the hermitian conjugate $(\cdot)^*$, we find that

$$\begin{aligned} \overline{(X | Y)} &= \overline{\text{tr}(XY)} = \text{tr}(\overline{X}\overline{Y}) \\ &= \text{tr}(\overline{X}\overline{Y})^T \\ &= -\text{tr}(Y^*X^*) \\ &= -\text{tr}(YX) \\ &= (Y | X) \\ &= (X | Y), \end{aligned}$$

hence $(X | Y)$ is real. Finally,

$$(X | X) = -\text{tr}(X^2) = \text{tr}(X^*X),$$

and the latter is the sum of the eigenvalues of the positive definite hermitian matrix X^*X . Since these eigenvalues are real and non-negative, $(\cdot | \cdot)$ is real and positive definite. This is still true for the restriction of $(\cdot | \cdot)$ to an \mathbb{R} -Lie subalgebra of $\mathfrak{u}(n)$, in particular to the Lie subalgebra of a Lie subgroup of $U(n)$. In the case $SU(2) \leqslant U(2)$, $(\cdot | \cdot)$ agrees with the inner product introduced in Section 3.5.

For $A \in U(n)$,

$$\begin{aligned} (AXA^{-1} | AYA^{-1}) &= \text{tr}(AXA^{-1}AYA^{-1}) \\ &= \text{tr}(AXYA^{-1}) \\ &= \text{tr}(XY), \end{aligned}$$

giving

$$(\text{Ad}_A(X) | \text{Ad}_A(Y)) = (X | Y). \quad (11.2)$$

This shows that $\text{Ad}_A: \mathfrak{u}(n) \rightarrow \mathfrak{u}(n)$ is a linear isometry with respect to the inner product $(\cdot | \cdot)$. If we choose an orthonormal basis for $\mathfrak{u}(n)$ then the adjoint representation becomes a homomorphism $\text{Ad}: U(n) \rightarrow \text{SO}(\dim \mathfrak{u}(n))$ since $U(n)$ is connected.

When $A = \exp(tH)$ with $t \in \mathbb{R}$ and $H \in \mathfrak{u}(n)$,

$$\begin{aligned} \frac{d}{dt}(AXA^{-1} | AYA^{-1})_{t=0} &= \frac{d}{dt} \text{tr}(AXYA^{-1})_{t=0} \\ &= \text{tr}(HXY - XYH) \\ &= \text{tr}([H, X]Y + X[H, Y]) \\ &= ([H, X] | Y) + (X | [H, Y]), \end{aligned}$$

and so

$$(\text{ad}_H(X) | Y) + (X | \text{ad}_H(Y)) = 0. \quad (11.3)$$

Because of Equations (11.2) and (11.3), the inner product $(\cdot | \cdot)$ is said to be *invariant* with respect to Ad , ad or the Lie bracket $[\cdot, \cdot]$. On choosing an orthonormal basis for $u(n)$ with respect to the inner product $(\cdot | \cdot)$, each associated matrix $\text{Ad}_H \in \text{GL}_{\dim U(n)}(\mathbb{R})$ is special orthogonal and $\text{ad}_H \in M_{\dim U(n)}(\mathbb{R})$ is skew-symmetric, hence the adjoint representation becomes a homomorphism of Lie algebras

$$\text{ad}: u(n) \longrightarrow \mathfrak{so}(\dim U(n)) = \text{Sk-Sym}_{\dim U(n)}(\mathbb{R}).$$

Here is the more general result obtained by restricting the inner product $(\cdot | \cdot)$ to the subspace \mathfrak{g} .

Theorem 11.2

Let $G \leqslant U(n)$ be a compact connected Lie subgroup. On choosing an orthonormal \mathbb{R} -basis for \mathfrak{g} with respect to the inner product $(\cdot | \cdot)$, the adjoint representations of G and \mathfrak{g} take the form

$$\text{Ad}: G \longrightarrow \text{SO}(\dim G), \quad \text{ad}: \mathfrak{g} \longrightarrow \mathfrak{so}(\dim G),$$

which are homomorphisms of groups and Lie algebras respectively.

Recall from Definition 3.31 the notion of complexification of a Lie algebra. We can extend the inner product $(\cdot | \cdot)$ to a \mathbb{C} -bilinear inner product on the complexification $u(n)_{\mathbb{C}} = \mathfrak{gl}_n(\mathbb{C})$. The adjoint action of $U(n)$ on $u(n)$ extends to a \mathbb{C} -linear action on $u(n)_{\mathbb{C}}$ and this extension is Ad -invariant.

Proposition 11.3

Let $\mathfrak{g} \leqslant u(n)$. Then there is a unique extension of $(\cdot | \cdot)$ on \mathfrak{g} to a \mathbb{C} -bilinear inner product $(\cdot | \cdot)_{\mathbb{C}}$ on the complexification $\mathfrak{g}_{\mathbb{C}}$. Furthermore, $(\cdot | \cdot)_{\mathbb{C}}$ is both Ad -invariant and ad -invariant.

Proof

Let $\{U_1, \dots, U_d\}$ be an \mathbb{R} -basis of \mathfrak{g} , which is also a \mathbb{C} -basis of $\mathfrak{g}_{\mathbb{C}}$. Taking as the definition

$$\left(\sum_i z_i U_i \mid \sum_j w_j U_j \right)_{\mathbb{C}} = \sum_i \sum_j z_i w_j (U_i \mid U_j),$$

it is easily seen that $(\cdot | \cdot)_{\mathbb{C}}$ is indeed the unique \mathbb{C} -bilinear extension of $(\cdot | \cdot)$ to the complex vector space $\mathfrak{g}_{\mathbb{C}}$. The invariance is straightforward to check. \square

Example 11.4

For $n \geq 1$, the \mathbb{C} -basis of $\mathfrak{u}(n)_{\mathbb{C}} = \mathfrak{gl}_n(\mathbb{C})$ of Equation (3.27) satisfies

$$(E^{k\ell} | E^{rs}) = -\delta_{kr}\delta_{\ell s},$$

while for $\mathfrak{su}(n)_{\mathbb{C}} = \mathfrak{sl}_n(\mathbb{C})$,

$$(E^{kk} - E^{(k-1)(k-1)} | E^{rs}) = 0 \quad \text{if } r \neq s, \quad (11.4a)$$

$$(E^{kk} - E^{(k-1)(k-1)} | E^{rr} - E^{(r-1)(r-1)}) = \begin{cases} -2 & \text{if } r = k, \\ 1 & \text{if } |r - k| = 1. \end{cases} \quad (11.4b)$$

11.2 The Centre and its Lie Algebra

In Section 10.4 we described the centre $Z(G)$ of G in terms of the maximal tori of G . We can determine the dimension of $Z(G)$ by considering its Lie algebra.

Theorem 11.5

Let G be a compact connected Lie group and let \mathfrak{z} be the Lie algebra of its centre $Z(G)$. Then $\mathfrak{z} = z(\mathfrak{g})$ and so $\dim Z(G) = \dim z(\mathfrak{g})$.

Proof

Let $\gamma: (-\varepsilon, \varepsilon) \rightarrow Z(G)$ be a smooth curve with $\gamma(0) = I$. By Theorem 7.32, for all $x \in \mathfrak{g}$ and $s \in \mathbb{R}$, $\exp(sx) \in G$ and so for all $t \in \mathbb{R}$,

$$\gamma(t) \exp(sx) \gamma(t)^{-1} = \exp(sx).$$

Then we have

$$\begin{aligned} [\gamma'(0), x] &= \frac{d}{dt} \gamma(t) x \gamma(t)^{-1} \Big|_{t=0} \\ &= \frac{d}{dt} \gamma(t) \left(\frac{d}{ds} \exp(sx) \Big|_{s=0} \right) \gamma(t)^{-1} \Big|_{t=0} \\ &= \frac{d}{dt} \frac{d}{ds} (\gamma(t) \exp(sx) \gamma(t)^{-1}) \Big|_{s=0|t=0} \\ &= \frac{d}{dt} \frac{d}{ds} (\exp(sx)) \Big|_{s=0|t=0} \\ &= \frac{d}{dt} x \Big|_{t=0} = 0, \end{aligned}$$

so $\gamma'(0) \in z(\mathfrak{g})$. This shows that $\mathfrak{z} \subseteq z(\mathfrak{g})$.

Conversely, let $z \in z(\mathfrak{g})$ and $t \in \mathbb{R}$. Then by Theorem 10.16, every element $g \in G$ has the form $g = \exp(x)$ for some $x \in \mathfrak{g}$, therefore making use of Proposition 2.2 we have

$$\begin{aligned}\exp(tz)g\exp(tz)^{-1} &= \exp(tz)\exp(x)\exp(tz)^{-1} \\ &= \exp(tz+x)\exp(-tz) \\ &= \exp(tz+x-tz) \\ &= \exp(x),\end{aligned}$$

which shows that $\exp(tz) \in Z(G)$ and hence $z \in \mathfrak{z}$ by Equation (7.15). So $z(\mathfrak{g}) \subseteq \mathfrak{z}$.

Combining these two inclusions we obtain $z(\mathfrak{g}) = \mathfrak{z}$ and so

$$\dim Z(G) = \dim \mathfrak{z} = \dim z(\mathfrak{g}).$$

□

Corollary 11.6

$\dim Z(G) = 0$ if and only if \mathfrak{g} is centreless.

A Lie group G for which Corollary 11.6 holds can still have non-trivial centre; this is illustrated in Propositions 5.31 and 5.32. Here is a more general observation in this vein.

Proposition 11.7

Let G be a compact Lie group and $H \leq G$ be a closed subgroup. Then $\dim H = 0$, if and only if H is a finite subgroup.

More generally, H has finitely many path components, each of which is compact.

Proof

When $\dim H = 0$, H is a discrete subgroup of the compact group G so it must be finite as shown in the exercises for Chapter 1.

If $H \leq G$ is a closed subset then it is compact by Proposition 1.26. The path components form an open cover of H by disjoint sets, and this must be finite by compactness. □

Now we can give another description of the centre of G .

Proposition 11.8

The adjoint action of G on \mathfrak{g} gives a Lie homomorphism

$$\text{Ad}: G \longrightarrow \text{SO}(\dim G); \quad g \longmapsto \text{Ad}_g,$$

with compact connected image and kernel $\ker \text{Ad} = Z(G)$.

Proof

Let $g \in Z(G)$. Then if $x \in \mathfrak{g}$ and $t \in \mathbb{R}$, $g \exp(tx)g^{-1} = \exp(tx)$ and differentiating with respect to t at 0 gives $gxg^{-1} = x$. Hence $g \in \ker \text{Ad}$.

Conversely, if $g \in \ker \text{Ad}$ then for $x \in \mathfrak{g}$, $gxg^{-1} = x$ and by applying the exponential we obtain

$$g \exp(x)g^{-1} = \exp(gxg^{-1}) = \exp(x).$$

Since every element of G has the form $\exp(x)$, we have $g \in Z(G)$. \square

11.3 Lie Ideals and the Adjoint Action

We would like to generalise Theorem 11.5 to arbitrary closed normal subgroups. We will make use of ideas and notation introduced in Section 1.8. First we need to investigate the adjoint action further and in particular its behaviour with respect to Lie ideals in the Lie algebra \mathfrak{g} .

Proposition 11.9

Let G be a compact connected Lie group and $\mathfrak{n} \subseteq \mathfrak{g}$ be an \mathbb{R} -subspace. Then $\mathfrak{n} \triangleleft \mathfrak{g}$ if and only if the adjoint action of G on \mathfrak{g} restricts to an action on \mathfrak{n} ,

$$\text{Ad}_g: \mathfrak{n} \longrightarrow \mathfrak{n}; \quad \text{Ad}_g(x) = gxg^{-1}.$$

Proof

Let $\mathfrak{n} \triangleleft \mathfrak{g}$. Then for every $u \in \mathfrak{g}$,

$$\text{ad}_u(x) = [u, x] \in \mathfrak{n} \quad (x \in \mathfrak{n}).$$

This shows that the adjoint action of \mathfrak{g} on \mathfrak{n} assigns to each $u \in \mathfrak{g}$ a linear transformation $\text{ad}_u \in \text{End}_{\mathbb{R}}(\mathfrak{n})$. Now consider the differential equation

$$\varphi'(t) = \varphi(t) \text{ad}_u, \tag{11.5}$$

where $\varphi: \mathbb{R} \rightarrow \mathrm{GL}_{\mathbb{R}}(n)$ is supposed to be a smooth function and $\varphi(t) \mathrm{ad}_u$ means the composition of these elements $\mathrm{End}_{\mathbb{R}}(n)$. By the general theory of such equations given in Theorem 2.12, this has a unique solution. To see this solution explicitly, for each $t \in \mathbb{R}$ consider the function

$$\mathrm{Ad}_{\exp(tu)}: g \longrightarrow g; \quad \mathrm{Ad}_{\exp(tu)}(x) = \exp(tu)x \exp(-tu).$$

On differentiating with respect to t we find

$$\begin{aligned} \frac{d}{dt} \exp(tu)x \exp(-tu) &= \exp(tu)[u, x] \exp(-tu) \\ &= \mathrm{Ad}_{\exp(tu)}(\mathrm{ad}_u(x)), \end{aligned}$$

hence this satisfies Equation (11.5) and so its unique solution is the function φ for which

$$\varphi(t) = \mathrm{Ad}_{\exp(tu)}: n \longrightarrow n.$$

Since G is compact and connected, by Theorem 10.16, every element $g \in G$ has the form $g = \exp(u)$ for some $u \in g$, hence Ad_g restricts to a map $\mathrm{Ad}_g: n \longrightarrow n$. This means that n is indeed closed under the action of each Ad_g , i.e., $\mathrm{Ad}_g \in \mathrm{GL}_{\mathbb{R}}(n)$.

Conversely, suppose that $n \subseteq g$ is closed under the action of Ad_g for every $g \in G$, then for $u \in g$, $x \in n$ and $t \in \mathbb{R}$, $\mathrm{Ad}_{\exp(tu)}(x) \in n$ and differentiating at $t = 0$ gives $[u, x] \in n$. Hence $n \triangleleft g$. \square

Now suppose that there is a connected normal Lie subgroup $N \triangleleft G$ with Lie algebra $n \leqslant g$. For each $x \in n$ and $u \in g$ we have $\mathrm{Ad}_{\exp(tu)} \exp(sx) \in G$ for all $s, t \in \mathbb{R}$. Differentiating with respect to s at 0, we obtain

$$\frac{d}{ds} \mathrm{Ad}_{\exp(tu)} \exp(sx)|_{s=0} = \mathrm{Ad}_{\exp(tu)}(x) \in n.$$

Differentiating this with respect to t at 0, now gives $[u, x] \in n$. Hence the Lie algebra n of a normal subgroup of G is a Lie ideal, $n \triangleleft g$.

Proposition 11.10

Let G be a compact connected Lie group. Let $N \triangleleft G$ be a closed normal subgroup with Lie subalgebra $n \leqslant g$. Then $n \triangleleft g$ and the adjoint action of G on g restricts to an action on n ,

$$\mathrm{Ad}_g: n \longrightarrow n; \quad \mathrm{Ad}_g(x) = gxg^{-1}.$$

Corollary 11.11

Let G be a compact connected Lie group. If \mathfrak{g} is simple then G has no closed normal subgroups of dimension greater than 0.

Definition 11.12

A compact connected Lie group with a simple Lie algebra is called *simple*.

Remark 11.13

This terminology conflicts with the usual usage in group theory; however, it does make sense in that much of the structure of a compact connected Lie group is captured by its Lie algebra. For example, as we will see, the classification of simple compact connected Lie groups is essentially accomplished by first classifying their Lie algebras up to isomorphism, then determining their associated groups which differ by surjective Lie homomorphisms with finite kernels, i.e., *finite covering homomorphisms*.

The following converse result is true, but we will not give a proof.

Proposition 11.14

Let G be a compact connected Lie group. If G has no closed normal subgroups of dimension greater than 0, then the Lie algebra \mathfrak{g} is simple.

Recall the *orthogonal complement* of \mathfrak{n} in \mathfrak{g} ,

$$\mathfrak{n}^\perp = \{x \in \mathfrak{g} : \forall y \in \mathfrak{n}, (x | y) = 0\}.$$

Clearly \mathfrak{n}^\perp is a vector subspace of \mathfrak{g} and by standard arguments of linear algebra

$$\mathfrak{g} = \mathfrak{n} + \mathfrak{n}^\perp, \tag{11.6a}$$

$$\mathfrak{n} \cap \mathfrak{n}^\perp = \{0\}, \tag{11.6b}$$

$$(\mathfrak{n}^\perp)^\perp = \mathfrak{n}. \tag{11.6c}$$

Proposition 11.15

Let \mathfrak{g} be a finite dimensional \mathbb{R} -Lie algebra with a real positive definite *ad*-invariant inner product $(|)$. Let $\mathfrak{n} \triangleleft \mathfrak{g}$ be any Lie ideal.

i) The orthogonal complement \mathfrak{n}^\perp of \mathfrak{n} in \mathfrak{g} is a Lie ideal, $\mathfrak{n}^\perp \triangleleft \mathfrak{g}$ which satisfies

the Equation (11.6).

- ii) There is an isomorphism of Lie algebras $n^\perp \cong g/n$.

Proof

- (i) If $x \in n^\perp$ and $y \in n$ and $z \in g$, then

$$([z, x] | y) = -(x | [z, y]) = 0,$$

by Equation (11.3) and the fact that n is an ideal.

- (ii) The quotient homomorphism $g \rightarrow g/n$ restricts to a homomorphism of Lie algebras $n^\perp \rightarrow g/n$ which is injective since $n \cap n^\perp = \{0\}$. Since

$$\dim n^\perp = \dim g/n = \dim g - \dim n,$$

it is also surjective. \square

Here are some further results of this type.

Proposition 11.16

Let g be a finite dimensional \mathbb{R} -Lie algebra equipped with a real positive definite ad-invariant inner product $(\cdot | \cdot)$.

- i) g' and $z(g)$ are mutually orthogonal, i.e.,

$$([x, y] | z) = 0 \quad (x, y \in g, z \in z(g)).$$

Hence $g' \subseteq z(g)^\perp$,

- ii) There is an isomorphism of Lie algebras $g^\perp \cong g/z(g)$.
 iii) The centre of $g/z(g)$ is trivial.

Proof

- (i) If $x, y \in g$ and $z \in z(g)$, then by Equation (11.3),

$$([x, y] | z) = -(y | [x, z]) = 0,$$

since $[x, z] = 0$. So g' and $z(g)$ are indeed mutually orthogonal.

- (ii) This is a special case of Proposition 11.15(ii).

- (iii) The isomorphism

$$z(g/z(g)) \cong z(g^\perp) = \{0\}$$

and the identity

$$z(g^\perp) = z(g) \cap g^\perp = \{0\}$$

give this result. \square

11.4 Semi-simple Decompositions

Let G be a compact connected Lie group. The centre $Z(G)$ of G is a compact normal subgroup, which need not be connected. By Theorem 11.5, the centre of the Lie algebra \mathfrak{g} is an ideal $z(\mathfrak{g}) \triangleleft \mathfrak{g}$. By Proposition 11.8, the quotient $G/Z(G)$ is a compact connected Lie subgroup of $SO(\dim G)$ and the quotient homomorphism $G \rightarrow G/Z(G)$ is a Lie homomorphism whose derivative coincides with the quotient homomorphism of Lie algebras $\mathfrak{g} \rightarrow \mathfrak{g}/z(\mathfrak{g})$. This suggests that we should study G through $G/Z(G)$ and its Lie algebra. In particular, by Proposition 11.16(iii), $\dim Z(G/Z(G)) = 0$ and $Z(G/Z(G))$ is a finite group. In fact the centre of $G/Z(G)$ is trivial as the next result shows.

Proposition 11.17

Let G be a compact connected Lie group. Then the quotient group $G/Z(G)$ has trivial centre, i.e., $Z(G/Z(G)) = \{1\}$.

Proof

This follows from Proposition 11.8 together with Proposition 11.16(ii) and the factorisation of the adjoint map Ad as

$$Ad: G \rightarrow G/Z(G) \xrightarrow{\overline{Ad}} SO(\dim G),$$

from which it follows that $\ker Ad$ agrees with the kernel of the induced action of G on $\mathfrak{g}^\perp \cong \mathfrak{g}/z(\mathfrak{g})$. This in turn agrees with the inverse image of $Z(G/Z(G))$ under the quotient homomorphism $G \rightarrow G/Z(G)$. \square

Assumption 11.18

Because of Proposition 11.17, we may as well replace G by $G/Z(G)$, so from now on in this section, we will assume that $Z(G) = \{1\}$. Such a Lie group is called *semi-simple*.

The adjoint action of G on \mathfrak{g} gives rise to an orthogonal decomposition of \mathfrak{g} into Lie ideals. The proof of this requires Haar measure on G and is closely related to Theorem 10.1. For readers familiar with the representation theory of finite groups, we remark that this kind of result is analogous to *Maschke's Theorem*.

Recall from Proposition 11.9 that a Lie ideal of \mathfrak{g} is essentially the same thing as an Ad -closed \mathbb{R} -subspace of \mathfrak{g} .

Theorem 11.19

There are simple ideals $\mathfrak{g}_i \triangleleft \mathfrak{g}$ ($i = 1, \dots, r$) which decompose \mathfrak{g} as an orthogonal direct sum,

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r, \quad \mathfrak{g}_i \subseteq \mathfrak{g}_j^\perp \quad \text{if } i \neq j.$$

For each $i = 1, \dots, r$, the ideal

$$\mathfrak{g}(i) = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_{i-1} \oplus \mathfrak{g}_{i+1} \oplus \cdots \oplus \mathfrak{g}_r \triangleleft \mathfrak{g}$$

is the orthogonal complement of \mathfrak{g}_i , i.e.,

$$\mathfrak{g}(i) = \mathfrak{g}_i^\perp.$$

Notice that if $i \neq j$, $x \in \mathfrak{g}_i$ and $y \in \mathfrak{g}_j$, then since $[x, y] \in \mathfrak{g}_i \cap \mathfrak{g}_j = \{0\}$, we must have $[x, y] = 0$.

The adjoint action of G on \mathfrak{g} restricts to an action on $\mathfrak{g}(i)$. Let

$$G_i = \{g \in G : \text{Ad}_g = \text{Id}_{\mathfrak{g}(i)} : \mathfrak{g}(i) \longrightarrow \mathfrak{g}(i)\}.$$

Then G_i is clearly a closed subgroup of G and so is a compact Lie subgroup. If $x \in \mathfrak{g}_i$ then for $y \in \mathfrak{g}(i)$, $[x, y] = 0$. Hence for all $t \in \mathbb{R}$,

$$\text{Ad}_{\exp(tx)}(y) = y.$$

This shows that $\exp \mathfrak{g}_i \subseteq G_i$.

On the other hand, if $x \in \mathfrak{g}$ is in the Lie subalgebra of G_i , then for every $t \in \mathbb{R}$, $\exp(tx) \in G_i$ so then for every $y \in \mathfrak{g}(i)$,

$$\exp(tx)y\exp(-tx) = y.$$

On differentiating at $t = 0$ we see that $[x, y] = 0$. Writing $x = x_i + x'_i$ with $x_i \in \mathfrak{g}_i$ and $x'_i \in \mathfrak{g}(i)$ we find that $[x'_i, y] = 0$ for every $y \in \mathfrak{g}(i)$, so $x'_i \in Z(\mathfrak{g}) = \{0\}$. Therefore $x \in \mathfrak{g}_i$ which shows that the Lie algebra of G_i is \mathfrak{g}_i .

It is easy to verify that $G_i \triangleleft G$ and in fact every pair of elements $g_i \in G_i$ and $g_j \in G_j$ commutes if $i \neq j$. This implies that $Z(G_i) \leq Z(G)$, giving $Z(G_i) = \{1\}$.

Theorem 11.20

For each $i = 1, \dots, r$ there is a closed normal subgroup $G_i \triangleleft G$ with Lie subalgebra $\mathfrak{g}_i \triangleleft \mathfrak{g}$. Moreover, there is a direct product decomposition

$$G = G_1 \times \cdots \times G_r.$$

So to understand the structure of an arbitrary compact connected Lie group G , we first factor out its centre $Z(G)$ to form $G/Z(G)$, which is semi-simple, so is a product of simple groups. This naturally leads us to consider the structure of simple compact connected Lie groups. In order to do this and to explain the classification of such simple groups we need to study the adjoint representation in greater detail since this provides the key ingredients.

11.5 The Structure of the Adjoint Representation

Now let G be a compact connected Lie group of dimension $d = \dim G$ and $T \leq G$ be a maximal torus. We will use the notation $d = 2m$ if d is even and $d = 2m + 1$ if d is odd.

By Theorem 10.1, we may assume that G is a subgroup of some unitary group, say $G \leq U(n)$. Also, by Proposition 10.7, T has a topological generator g_0 and on choosing an orthonormal basis for \mathfrak{g} , the adjoint action of g_0 gives a linear transformation $\text{Ad}_{g_0} : \mathfrak{g} \rightarrow \mathfrak{g}$ which is equivalent to a matrix $\text{Ad}_{g_0} \in SO(d)$; this matrix determines the adjoint action of T on \mathfrak{g} . The Principle Axis Theorem 10.13 implies that Ad_{g_0} is conjugate in $SO(d)$ to a matrix of the form $R_d(\theta_1, \dots, \theta_m)$. The eigenvalues of Ad_{g_0} (with repetitions) must then be of the form

$$\begin{cases} e^{\theta_1 i}, e^{-\theta_1 i}, \dots, e^{\theta_m i}, e^{-\theta_m i} & \text{if } d \text{ is even,} \\ e^{\theta_1 i}, e^{-\theta_1 i}, \dots, e^{\theta_m i}, e^{-\theta_m i}, 1 & \text{if } d \text{ is odd.} \end{cases}$$

If we view $SO(d)$ as a subgroup of $U(d)$, then Ad_{g_0} is conjugate in $U(d)$ to the diagonal matrix

$$\begin{cases} \text{diag}(e^{\theta_1 i}, e^{-\theta_1 i}, \dots, e^{\theta_m i}, e^{-\theta_m i}) & \text{if } d \text{ is even,} \\ \text{diag}(e^{\theta_1 i}, e^{-\theta_1 i}, \dots, e^{\theta_m i}, e^{-\theta_m i}, 1) & \text{if } d \text{ is odd.} \end{cases}$$

These matrices give the action of Ad_{g_0} on the complexification \mathfrak{g}_C relative to the chosen basis of \mathfrak{g} thought of as a C -basis for \mathfrak{g}_C .

For a power g_0^r ($r \in \mathbb{Z}$), $\text{Ad}_{g_0^r} = (\text{Ad}_{g_0})^r$ is conjugate to

$$\begin{cases} R_{\dim G}(r\theta_1, \dots, r\theta_m) & \text{in } SO(d), \\ \text{diag}(e^{r\theta_1 i}, e^{-r\theta_1 i}, \dots, e^{r\theta_m i}, e^{-r\theta_m i}) & \text{in } U(d) \text{ if } d \text{ is even,} \\ \text{diag}(e^{r\theta_1 i}, e^{-r\theta_1 i}, \dots, e^{r\theta_m i}, e^{-r\theta_m i}, 1) & \text{in } U(d) \text{ if } d \text{ is odd.} \end{cases} \quad (11.7)$$

This implies that there is an orthonormal \mathbb{R} -basis $\mathcal{B} = \{v_1, \dots, v_d\}$ of \mathfrak{g} for which

$$\begin{aligned}\text{Ad}_{g_0^r}(v_{2k-1}) &= \cos(r\theta_k)v_{2k-1} + \sin(r\theta_k)v_{2k}, \\ \text{Ad}_{g_0^r}(v_{2k}) &= -\sin(r\theta_k)v_{2k-1} + \cos(r\theta_k)v_{2k},\end{aligned}$$

whenever $k \leq m$; if d is odd we also have

$$\text{Ad}_{g_0^r}(v_d) = v_d.$$

So for $k = 1, \dots, m$, there are continuous homomorphisms

$$\tilde{\gamma}_k: T \longrightarrow \text{SO}(2) \cong \mathbf{T}; \quad \tilde{\gamma}_k(g_0^r) = R(r\theta_k).$$

For each $\tilde{\gamma}_k$, the inverse map given by

$$\tilde{\gamma}_k^{-1}(g_0^r) = R(r\theta_k)^{-1} = R(-r\theta_k)$$

corresponds to replacing the pair of basis vectors (v_{2k-1}, v_{2k}) by the pair $(-v_{2k-1}, v_{2k})$. The homomorphisms $\tilde{\gamma}_k, \tilde{\gamma}_k^{-1}$ ($k = 1, \dots, m$) are called the *roots* of the maximal torus T in G , the non-constant ones being called the *non-trivial roots*.

Notice that for $x \in \mathfrak{t} \subseteq \mathfrak{g}$, $\text{Ad}_{g_0}(x) = 0$, since this comes from the adjoint action of the abelian group T . Hence if we first choose an orthonormal basis for \mathfrak{t} , we can extend it to one for \mathfrak{g} so that it has the properties above. From now on we will assume that \mathcal{B} has been chosen in this way so that $v_1, \dots, v_{\dim T}$ is an orthonormal basis for \mathfrak{t} .

For each root $\tilde{\gamma}: T \longrightarrow \mathbf{T}$ we take its derivative and by identifying the Lie algebra of \mathbf{T} with \mathbb{R} this gives an \mathbb{R} -linear transformation $\gamma: \mathfrak{t} \longrightarrow \mathbb{R}$ which is also known as a *root*.

In general, if $\gamma: \mathfrak{t} \longrightarrow \mathbb{R}$ is a non-trivial root, we define $\mathfrak{g}(\pm\gamma) \subseteq \mathfrak{g}$ to be the maximal subspace for which there is a basis $\{u_1, \dots, u_{2k}\}$ such that for $x \in \mathfrak{t}$ and $j = 1, \dots, k$,

$$\text{Ad}_{\exp(x)}(u_{2j-1}) = \cos \gamma(x) u_{2j-1} + \sin \gamma(x) u_{2j}, \quad (11.8a)$$

$$\text{Ad}_{\exp(x)}(u_{2j}) = -\sin \gamma(x) u_{2j-1} + \cos \gamma(x) u_{2j}. \quad (11.8b)$$

This is called the *root space* associated to the *root pair* $\pm\gamma$ and by definition it is non-trivial. When γ is non-trivial we have $\dim \mathfrak{g}(\pm\gamma) = 2m_{\pm\gamma}$, for some natural number $m_{\pm\gamma} \geq 1$ called the *multiplicity* of the root pair $\pm\gamma$. We also set

$$\mathfrak{g}(0) = \{x \in \mathfrak{g} : \forall g \in T, \text{Ad}_g(x) = x\},$$

which is the *null root space*. Then $\mathfrak{t} \subseteq \mathfrak{g}(0)$, so $\mathfrak{g}(0) \neq \{0\}$ if $\dim G > 0$; the number $m_0 = \dim \mathfrak{g}(0)$ is the multiplicity of the trivial root.

For practical purposes it is often more useful to work entirely with the Lie algebra and replace Equation (11.8) by the following equations which are obtained by differentiating the Ad-action at the identity,

$$\text{ad}_x(u_{2j-1}) = \gamma(x) u_{2j}, \quad (11.9a)$$

$$\text{ad}_x(u_{2j}) = -\gamma(x) u_{2j-1}. \quad (11.9b)$$

Theorem 11.21

There is an \mathbb{R} -linear direct sum decomposition which is orthogonal with respect to the invariant inner product $(\cdot | \cdot)$,

$$\mathfrak{g} = \bigoplus_{\pm\gamma \text{ a root pair}} \mathfrak{g}(\pm\gamma) = \mathfrak{g}(0) \oplus \bigoplus_{\pm\gamma \text{ a non-trivial root pair}} \mathfrak{g}(\pm\gamma).$$

Proof

This follows easily from the orthogonality of the eigenspaces associated with distinct conjugate pairs of eigenvalues of an orthogonal matrix. \square

The decomposition in this result is called the *root space decomposition* of \mathfrak{g} with respect to the maximal torus T . There is a complex version of this also called a *root space decomposition* or *Cartan decomposition* of the complexification \mathfrak{g}_C . The adjoint action of T extends to a \mathbb{C} -linear action on \mathfrak{g}_C and each of the complexified root spaces $\mathfrak{g}(\pm\gamma)_C$ for non-trivial γ splits into a direct sum of two \mathbb{C} -subspaces

$$\mathfrak{g}(\pm\gamma)_C = \mathfrak{g}_\gamma \oplus \mathfrak{g}_{-\gamma},$$

where

$$\begin{aligned} \mathfrak{g}_\gamma &= \{v \in \mathfrak{g}_C : \forall x \in \mathfrak{t}, \text{Ad}_{\exp(x)}(v) = e^{\gamma(x)i} v\} \\ &= \{v \in \mathfrak{g}_C : \forall x \in \mathfrak{t}, \text{ad}_x(v) = \gamma(x)i v\}. \end{aligned}$$

We also set $\mathfrak{g}_0 = \mathfrak{g}(0)_C$. When γ is non-trivial, $\dim_{\mathbb{C}} \mathfrak{g}_\gamma = m_{\pm\gamma}$, so we usually denote this by m_γ .

Theorem 11.22

There is a \mathbb{C} -linear direct sum decomposition,

$$\mathfrak{g}_C = \mathfrak{g}_0 \oplus \bigoplus_{\gamma \text{ a non-trivial root}} \mathfrak{g}_\gamma.$$

Such root space decompositions contain all the essential information required to classify simple compact connected Lie groups. In order to explain the result of this classification we will need to investigate the geometry of root systems and their Weyl groups and we do this in Chapter 12. To finish this introduction we explain how the Weyl group acts on roots and then give some examples which are elucidated further in Section 12.3.

For a maximal torus $T \leq G$, we know that the Weyl group $W_G(T) = N_G(T)/T$ acts by conjugation on the Lie algebra \mathfrak{t} of T . There is an associated *contragredient action* on the dual vector space \mathfrak{t}^* consisting of all \mathbb{R} -linear transformations $\mathfrak{t} \rightarrow \mathbb{R}$. This action is specified by taking ${}^{gT}\gamma = gT \cdot \gamma$ to be

$${}^{gT}\gamma: \mathfrak{t} \rightarrow \mathbb{R}; \quad {}^{gT}\gamma(x) = \gamma(\text{Ad}_{g^{-1}}(x)) = \gamma(\text{Ad}_g^{-1}(x)).$$

It is easy to see that if γ is a root, then so is ${}^{gT}\gamma$, and this leads to the following result.

Proposition 11.23

Under the action of $W_G(T)$ on \mathfrak{t}^* , the roots of G with respect to T are permuted.

Roots also have the properties given in the following result whose proof can be found in [13, 25].

Proposition 11.24

Let α, β be roots of G with respect to the maximal torus T .

- i) For non-trivial α and $w \in W_G(T)$, the multiplicities of $w\alpha$ and α agree, i.e., $m_{w\alpha} = m_\alpha$; in fact these multiplicities are 1.
- ii) For non-trivial α , the only roots which are non-zero multiples of α are $\pm\alpha$.
- iii) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \{0\}$ unless $\alpha + \beta$ is a root, and then

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$$

if $\alpha + \beta$ is non-trivial, while

$$\{0\} \neq [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \leq \mathfrak{g}_0.$$

By Theorems 10.23 and 11.5, we have $z(g) \leq \mathfrak{t}$. It is clear that for any root α we have $z(g) \subseteq \ker \alpha$; in fact this can be strengthened.

Proposition 11.25

Let T be a maximal torus of G . Then

$$z(\mathfrak{g}) = \bigcap_{\alpha \text{ a root}} \ker \alpha,$$

where the intersection is taken over roots of G with respect to T .

The roots of G with respect to T are best thought of as lying in the subspace

$$z(\mathfrak{g})^\perp = \{\gamma \in \mathfrak{t} : z(\mathfrak{g}) \subseteq \ker \gamma\} \subseteq \mathfrak{t}^*.$$

Recall that for $\theta \in \mathbb{R}$,

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

Example 11.26

$\mathrm{SO}(3)$ has the maximal torus

$$T = \left\{ \begin{bmatrix} R(\theta) & O_{2,1} \\ O_{1,2} & 1 \end{bmatrix} : \theta \in \mathbb{R} \right\} \leqslant \mathrm{SO}(3),$$

and the lie algebras of $\mathrm{SO}(3)$ and T are

$$\mathfrak{so}(3) = \left\{ \begin{bmatrix} 0 & -u & -v \\ u & 0 & -w \\ v & w & 0 \end{bmatrix} : u, v, w \in \mathbb{R} \right\},$$

$$\mathfrak{t} = \left\{ \begin{bmatrix} 0 & -t & 0 \\ t & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} : t \in \mathbb{R} \right\} \leqslant \mathfrak{so}(3).$$

The elements

$$U = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$$

form a basis of $\mathfrak{so}(3)$ with $U \in \mathfrak{t}$. Then for $t \in \mathbb{R}$,

$$\exp(tU) = \begin{bmatrix} R(t) & O_{2,1} \\ O_{1,2} & 1 \end{bmatrix}$$

and

$$\begin{aligned}\exp(tU)U\exp(tU)^{-1} &= U, \\ \exp(tU)V\exp(tU)^{-1} &= \cos t V + \sin t W, \\ \exp(tU)W\exp(tU)^{-1} &= -\sin t V + \cos t W.\end{aligned}$$

In terms of the ad-action we have

$$[U, U] = 0, \quad [U, V] = W, \quad [U, W] = -V.$$

So there are two non-trivial roots $\pm\beta$, where

$$\beta(tP) = t \quad (t \in \mathbb{R}).$$

From Example 9.19, the normaliser of T in $\mathrm{SO}(3)$ is

$$N_{\mathrm{SO}(3)}(T) = T \cup \left\{ \begin{bmatrix} -\cos \theta & \sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & -1 \end{bmatrix} : \theta \in \mathbb{R} \right\} = T \cup ZT,$$

where $Z = \mathrm{diag}(-1, 1, -1)$. The Weyl group

$$W_G(T) = \{IT, ZT\} \cong \{\pm 1\}$$

acts on the roots by

$$z^T(\pm\beta) = \mp\beta.$$

Example 11.27

$\mathrm{SU}(2)$ has the maximal torus

$$T = \{\mathrm{diag}(z, \bar{z}) : |z| = 1\} \leqslant \mathrm{SU}(2),$$

which has Lie algebra

$$\mathfrak{t} = \{\mathrm{diag}(ti, -ti) : t \in \mathbb{R}\} \leqslant \mathfrak{su}(2).$$

Starting with the orthonormal basis $\{\hat{H}, \hat{E}, \hat{F}\}$ of $\mathfrak{su}(2)$ for which

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

we find that for $t \in \mathbb{R}$,

$$\exp(t\hat{H}) = \mathrm{diag}(e^{ti}, e^{-ti})$$

and

$$\begin{aligned}\exp(t\hat{H})\hat{H}\exp(t\hat{H})^{-1} &= \hat{H}, \\ \exp(t\hat{H})\hat{E}\exp(t\hat{H})^{-1} &= \cos 2t\hat{E} + \sin 2t\hat{F}, \\ \exp(t\hat{H})\hat{F}\exp(t\hat{H})^{-1} &= -\sin 2t\hat{E} + \cos 2t\hat{F}.\end{aligned}$$

In terms of the ad-action we have

$$[\hat{H}, \hat{H}] = 0, \quad [\hat{H}, \hat{E}] = 2\hat{F}, \quad [\hat{H}, \hat{F}] = -2\hat{E}.$$

So here there are two non-trivial roots $\pm\alpha_1$ where

$$\alpha_1(\text{diag}(ti, -ti)) = \alpha_1(t\hat{H}) = 2t \quad (t \in \mathbb{R}).$$

The normaliser of T in $SU(2)$ is

$$N_{SU(2)}(T) = T \cup \left\{ \begin{bmatrix} 0 & -\bar{z} \\ z & 0 \end{bmatrix} : |z| = 1 \right\} = T \cup JT,$$

where

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

The Weyl group

$$W_{SU(2)}(T) = \{IT, JT\} \cong \{\pm 1\}$$

acts on the roots by

$$JT(\pm\alpha_1) = \mp\alpha_1.$$

Example 11.28

$U(2)$ has the maximal torus

$$T = \{\text{diag}(z, w) : |z| = 1 = |w|\} \leq U(2),$$

which has Lie algebra

$$\mathfrak{t} = \{\text{diag}(si, ti) : s, t \in \mathbb{R}\} \leq \mathfrak{u}(2).$$

We have the orthonormal basis $\{\hat{H}_1, \hat{H}_2, \hat{E}, \hat{F}\}$ of $\mathfrak{u}(2)$ for which

$$\hat{H}_1 = \begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{H}_2 = \begin{bmatrix} 0 & 0 \\ 0 & i \end{bmatrix}, \quad \hat{E} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \hat{F} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Then for $s, t \in \mathbb{R}$,

$$\exp(s\hat{H}_1 + t\hat{H}_2) = \text{diag}(e^{st}, e^{ti}).$$

Then

$$\exp(s\hat{H}_1 + t\hat{H}_2)(x\hat{H}_1 + y\hat{H}_2)\exp(s\hat{H}_1 + t\hat{H}_2)^{-1} = (x\hat{H}_1 + y\hat{H}_2),$$

$$\exp(s\hat{H}_1 + t\hat{H}_2)\hat{E}\exp(s\hat{H}_1 + t\hat{H}_2)^{-1} = \cos(s-t)\hat{E} + \sin(s-t)\hat{F},$$

$$\exp(s\hat{H}_1 + t\hat{H}_2)\hat{F}\exp(s\hat{H}_1 + t\hat{H}_2)^{-1} = -\sin(s-t)\hat{E} + \cos(s-t)\hat{F}$$

In terms of the ad-action we have

$$[s\hat{H}_1 + t\hat{H}_2, x\hat{H}_1 + y\hat{H}_2] = 0,$$

$$[s\hat{H}_1 + t\hat{H}_2, \hat{E}] = (s-t)\hat{F}, \quad [s\hat{H}_1 + t\hat{H}_2, \hat{F}] = -(s-t)\hat{E}.$$

So there are two non-trivial roots $\pm\alpha$ where

$$\alpha(s\hat{H}_1 + t\hat{H}_2) = s - t \quad (s, t \in \mathbb{R}).$$

Notice that

$$\ker \alpha = \{s(\hat{H}_1 + \hat{H}_2) : s \in \mathbb{R}\} = z(u(2)).$$

The normaliser of T in $U(2)$ is

$$N_{U(2)}(T) = T \cup \left\{ \begin{bmatrix} 0 & -\bar{z} \\ z & 0 \end{bmatrix} : |z| = 1 \right\} = T \cup JT,$$

with J as in Example 11.27; the Weyl group

$$W_{U(2)}(T) = \{IT, JT\} \cong \{\pm 1\}$$

acts on the roots by

$$J^T(\pm\alpha) = \mp\alpha,$$

as indicated in Figure 11.1.

Example 11.29

$SU(3)$ has maximal torus

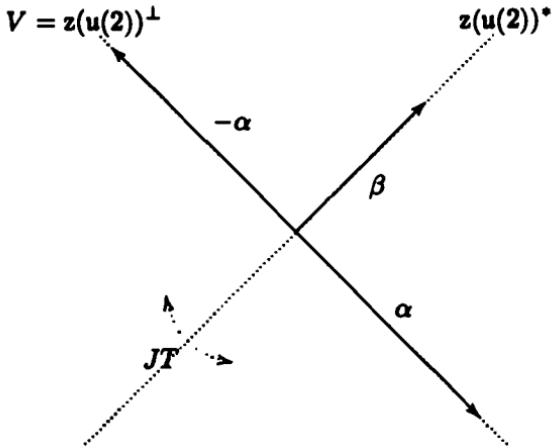
$$T = \{\text{diag}(z, w, \bar{zw}) : |z| = |w| = 1\} \leqslant SU(3),$$

with Lie algebra

$$\mathfrak{t} = \{\text{diag}(si, ti, -si - ti) : s, t \in \mathbb{R}\}.$$

In the Lie algebra $su(2)_C$ we have

$$H_1 = \begin{bmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix} \in \mathfrak{t},$$

Figure 11.1 The roots of $U(2)$ in $u(2)^*$

and

$$E^{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E^{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E^{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$E^{21} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E^{31} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad E^{32} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

These six elements form a \mathbb{C} -basis of $su(2)_\mathbb{C}$ and with respect to the invariant inner product $(\cdot | \cdot)$ we have

$$(H_1 | H_1) = (H_2 | H_2) = 2, \quad (H_1 | H_2) = -1.$$

For the ad-action we have

$$[sH_1 + tH_2, E^{12}] = (2s - t)iE^{12}, \quad [sH_1 + tH_2, E^{21}] = -(2s - t)iE^{12},$$

$$[sH_1 + tH_2, E^{13}] = (s + t)iE^{12}, \quad [sH_1 + tH_2, E^{31}] = -(s + t)iE^{12},$$

$$[sH_1 + tH_2, E^{23}] = (-s + 2t)iE^{23}, \quad [sH_1 + tH_2, E^{32}] = -(-s + 2t)iE^{32},$$

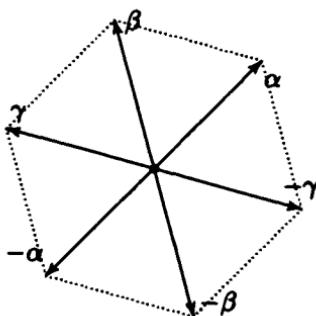
showing that there are six non-trivial roots $\pm\alpha, \pm\beta, \pm\gamma$, where

$$\alpha(sH_1 + tH_2) = 2s - t, \quad \beta(sH_1 + tH_2) = s + t, \quad \gamma(sH_1 + tH_2) = -s + 2t.$$

Using the inner product we can identify these as elements of $\mathfrak{t} \cong \mathfrak{t}^*$,

$$H_1 \longleftrightarrow \alpha, \quad H_1 + H_2 \longleftrightarrow \beta, \quad H_2 \longleftrightarrow \gamma.$$

The Weyl group $W_{SU(3)}(T)$ consists of 6 elements which act on the roots as the symmetries of the regular hexagon whose edges are indicated with dotted lines in Figure 11.2.

Figure 11.2 The roots of $SU(3)$ in $su(3)^*$ **EXERCISES**

11.1. For $n \geq 2$, consider the unitary group $U(n)$ and its Lie algebra $u(n)$ with its standard invariant inner product $(\cdot | \cdot)$.

a) Find the orthogonal complement of the centre of $u(n)$,

$$z(u(n))^\perp = \{X \in u(n) : \forall Z \in z(u(n)) (X | Z) = 0\}$$

b) Show that there is a Lie homomorphism $\varphi: U(n) \rightarrow Z(U(n))/Z$ whose kernel $\ker \varphi \triangleleft U(n)$ has Lie algebra $z(u(n))^\perp \leq u(n)$ and where $Z \triangleleft Z(U(n))$ is a finite normal subgroup.

c) Deduce that there is a surjective Lie homomorphism with finite kernel,

$$Z(U(n)) \times \ker \varphi \longrightarrow U(n).$$

11.2. Derive Equations (11.9) from Equations (11.8).

11.3. Verify that the Weyl group of Example 11.29 acts on the roots as claimed.

11.4. Determine the roots of each of the groups $Sp(2)$, and $SO(4)$ for the standard maximal torus. In each case, find the action of the Weyl group on the roots.

11.5. a) Find a maximal torus T_1 for the group

$$G_1 = \left\{ \begin{bmatrix} A & O_{2,2} \\ O_{2,2} & B \end{bmatrix} : A, B \in U(2), \det A \det B = 1 \right\} \leqslant SU(4).$$

Determine the roots of G_1 with respect to T_1 . Find the normaliser of T_1 in G_1 and the action of the Weyl group on the roots.

b) Do the same for the group

$$G_2 = \left\{ \begin{bmatrix} A & O_{2,2} \\ O_{2,2} & B \end{bmatrix} : A, B \in U(2), \det A = \det B \right\} \leqslant U(4).$$

Roots Systems, Weyl Groups and Dynkin Diagrams

In this chapter we discuss the theory of root systems and their Weyl groups and outline the classification of the irreducible ones. We indicate how these are associated with families of compact connected simple Lie groups. Accessible introductions to this material can be found in many books, e.g., [4, 25, 10]; [14] contains a technical account of the study of Lie algebras from the perspective of root systems, including certain important infinite dimensional examples known as Kac Moody algebras. We merely sketch the highlights without giving proofs.

12.1 Inner Products and Duality

In this section we recall some standard facts about inner products and duality. Throughout, let V be an ℓ -dimensional \mathbb{R} -vector space.

The \mathbb{R} -linear dual of V is the \mathbb{R} -vector space V^* consisting of all \mathbb{R} -linear transformations $V \rightarrow \mathbb{R}$. Given a basis $\{v_1, \dots, v_\ell\}$ for V , define elements $v_i^* \in V^*$ for $i = 1, \dots, \ell$ by

$$v_i^*(v_j) = \delta_{ij}.$$

Proposition 12.1

If $\{v_1, \dots, v_\ell\}$ is a basis of V , then $\{v_1^*, \dots, v_\ell^*\}$ is a basis of V^* .

$\{v_1^*, \dots, v_\ell^*\}$ is called the *dual basis* determined by $\{v_1, \dots, v_\ell\}$.

Suppose that $\langle \cdot, \cdot \rangle$ is a positive definite (but not necessarily symmetric) inner product on V . As usual, the length of a vector $\alpha \in V$ is defined to be $|\alpha| = \sqrt{\langle \alpha, \alpha \rangle}$. There is an \mathbb{R} -linear transformation

$$\nu: V \longrightarrow V^*; \quad \nu(u) = (u, \cdot),$$

where

$$\nu(u)(v) = \langle u, v \rangle \quad (v \in V).$$

The following result is standard.

Proposition 12.2

The \mathbb{R} -linear transformation $\nu: V \longrightarrow V^*$ is an isomorphism.

V^* inherits an inner product, which we will also write $\langle \cdot | \cdot \rangle$, defined by

$$\langle \alpha | \beta \rangle = \langle \nu^{-1}(\alpha) | \nu^{-1}(\beta) \rangle.$$

This is a positive definite inner product on V^* .

Proposition 12.3

Let $\{v_1, \dots, v_\ell\}$ be a basis for V and

$$A = [\langle v_i | v_j \rangle] \in M_\ell(\mathbb{R}).$$

Then the matrix $\bar{A} = [\langle v_i^* | v_j^* \rangle]$ associated with the inner product on $\langle \cdot | \cdot \rangle$ on V^* is given by

$$\bar{A} = (A^T)^{-1}.$$

Proof

This is a standard piece of linear algebra. Since

$$\nu(v_i) = \sum_{r=1}^{\ell} (\nu(v_i) | v_r) v_r^*,$$

on setting $B = [b_{ij}] = A^{-1}$, we obtain

$$\nu^{-1}(v_i) = \sum_{r=1}^{\ell} b_{ir} v_r,$$

which yields

$$\begin{aligned}(v_i^* \mid v_j^*) &= (\nu^{-1}(v_i)^* \mid \nu^{-1}(v_j^*)) = \sum_{r=1}^{\ell} \sum_{s=1}^{\ell} b_{ir} b_{js} (v_r \mid v_s) \\ &= \sum_{s=1}^{\ell} \delta_{is} b_{js} \\ &= b_{ji}.\end{aligned}$$

This implies the matrix identity

$$[(v_i^* \mid v_j^*)] = (A^T)^{-1}$$

which is the desired result. \square

12.2 Roots systems and their Weyl groups

Let V be an ℓ -dimensional real vector space with a positive definite (but not necessarily symmetric) inner product $\langle \cdot, \cdot \rangle$; as usual, the length of a vector $\alpha \in V$ is defined to be $|\alpha| = \sqrt{\langle \alpha, \alpha \rangle}$.

If $u \in V - \{0\}$, there is an associated \mathbb{R} -linear transformation $s_u: V \rightarrow V$ for which

$$s_u(x) = \begin{cases} -x & \text{if } x = tu, \\ x & \text{if } \langle u, x \rangle = 0. \end{cases}$$

It is easy to see that for $x \in V$,

$$s_u(x) = x - \frac{2 \langle u, x \rangle}{\langle u, u \rangle} u. \quad (12.1)$$

Geometrically, s_u is a *hyperplane reflection* in the hyperplane

$$H_u^\perp = \{x \in V : \langle u, x \rangle = 0\} \subseteq V.$$

Proposition 12.4

If $u, v \in V - \{0\}$, the following identities are satisfied:

$$\begin{aligned}s_u^2 &= s_u \circ s_u = \text{Id}_V, \\ \begin{cases} s_v = s_u & \text{if } v = tu \text{ for some } t \in \mathbb{R}, \\ s_v \neq s_u & \text{otherwise.} \end{cases}\end{aligned}$$

Definition 12.5

A subset $\Phi \subseteq V$ is a (*reduced*) *root system* if it has the following properties:

- Φ spans V and $0 \notin \Phi$;
- if $u \in \Phi$ then $-u \in \Phi$ and no other multiples of u are in Φ ;
- if $u, v \in \Phi$ then $2\langle u, v \rangle / \langle u, u \rangle \in \mathbb{Z}$ and $s_u(v) \in \Phi$.

An element of Φ is called a *root*.

The final condition is equivalent to the requirement that for $u, v \in \Phi$ and some $k \in \mathbb{Z}$,

$$s_u(v) - v = ku.$$

Each s_u acts on Φ as a permutation and it can be viewed as an element of the group of all permutations of Φ , S_Φ .

Definition 12.6

The *Weyl group* of Φ is the subgroup $W(\Phi) \leq S_\Phi$ generated by the s_u ($u \in \Phi$).

Remark 12.7

By definition $W(\Phi)$ is finite with order $|W(\Phi)|$ dividing $|\Phi|!$.

Definition 12.8

A subset $\Pi \subseteq \Phi$ is a set of *fundamental* or *simple roots* if it satisfies

- Π is an \mathbb{R} -basis of V ;
- every $u \in \Phi$ can be uniquely expressed in the form

$$u = \sum_{w \in \Pi} c_w w$$

where $c_w \in \mathbb{Z}$ either all satisfy $c_w \geq 0$ or all satisfy $c_w \leq 0$.

The choice of such a fundamental system is not unique, however given Π there is a partition $\Phi = \Phi^+ \cup \Phi^-$, where

$$\Phi^+ = \left\{ \sum_{w \in \Pi} c_w w : c_w \geq 0 \right\}, \quad \Phi^- = \left\{ \sum_{w \in \Pi} c_w w : c_w \leq 0 \right\}.$$

Notice that

$$-\Phi^+ = \{-u : u \in \Phi^+\} = \Phi^-, \quad -\Phi^- = \{-u : u \in \Phi^-\} = \Phi^+.$$

The elements of Π are called (*positive*) *simple roots*, while the elements of Φ^+ or Φ^- are called *positive* or *negative roots*.

Remark 12.9

It is easy to see that $W(\Phi)$ is generated by the reflections s_u ($u \in \Pi$).

12.3 Some Examples of Root Systems

We begin with an already familiar example from Section 11.5.

Example 12.10 (The root system of $U(2)$)

This builds on Example 11.28. Consider t^* with orthonormal basis $\{\varepsilon_1, \varepsilon_2\}$ dual to the orthonormal basis $\{\hat{H}_1, \hat{H}_2\}$ of $t \cong t^{**}$. Here we take $V = z(u(2))^\perp \subseteq t^*$. The root $\alpha = \varepsilon_1 - \varepsilon_2 \in V$ has length $|\alpha| = \sqrt{2}$. If $\beta = \varepsilon_1 + \varepsilon_2 \in t^*$, then $\alpha \cdot \beta = 0$, and $R(\beta) = z(u(2))^*$. So for this root system we have

$$\Pi = \{\alpha\}, \quad \Phi = \{-\alpha, \alpha\},$$

with Weyl group $W(\Phi) = S_\Phi \cong S_2$. The generator JT of the Weyl group acts on t^* by reflection in the line $z(u(2))^*$ as indicated in Figure 11.1.

In the remaining examples, \mathbb{R}^t and its subspaces have the usual positive definite inner product $\langle x | y \rangle = x \cdot y$. In the diagrams, the solid arrows indicate a choice of positive simple roots, while dashed arrows indicate the remaining roots.

Example 12.11 (The root system A_1)

Let $V = \mathbb{R}$ and

$$\Pi = \{\alpha_1\}, \quad \Phi = \{-\alpha_1, \alpha_1\},$$

where $\alpha_1 \in \mathbb{R}$ is non-zero. This root system has Weyl group $W(\Phi) = S_\Phi \cong S_2$.

Figure 12.1 The roots of A_1

Example 12.12 (The root system A_2)

Let $V = \mathbb{R}^2$ and

$$\Pi = \{\alpha_1, \alpha_2\}, \quad \Phi = \{\alpha_1, -\alpha_1, \alpha_2, -\alpha_2, \alpha_1 + \alpha_2, -\alpha_1 - \alpha_2\},$$

where α_1, α_2 are two vectors for which

$$|\alpha_1| = |\alpha_2|, \quad \alpha_1 \cdot \alpha_2 = -\frac{|\alpha_1| |\alpha_2|}{2}.$$

This root system has the hyperplane reflections $s_{\alpha_1}, s_{\alpha_2}, s_{\alpha_1+\alpha_2}$ and Weyl group $S_\Phi \cong S_3$.

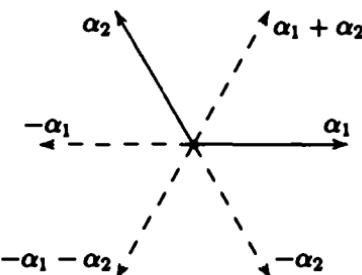


Figure 12.2 The roots of A_2

The preceding examples generalise to give the root systems A_n .

Example 12.13 (The root system A_n for $n \geq 1$)

Consider

$$V = \{x \in \mathbb{R}^{n+1} : x \cdot (e_1 + \cdots + e_{n+1}) = 0\} \subseteq \mathbb{R}^{n+1},$$

and $\alpha_k = e_k - e_{k+1}$ for $1 \leq k \leq n$. Then V is a vector space of dimension n with basis $\{\alpha_1, \dots, \alpha_n\}$. Notice that

$$\alpha_i \cdot \alpha_j = \begin{cases} 2 & \text{if } i = j, \\ -1 & \text{if } |i - j| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Taking

$$\Pi = \{\alpha_1, \dots, \alpha_n\},$$

$$\begin{aligned} \Phi &= \{\pm(\alpha_r + \alpha_{r+1} + \cdots + \alpha_s) : 1 \leq r \leq s \leq n\} \\ &= \{\pm(e_r - e_{s+1}) : 1 \leq r < s \leq n\}, \end{aligned}$$

we obtain a root system with Weyl group $S_\Phi \cong S_n$.

Example 12.14 (The root system B_n for $n \geq 2$)

Take $V = \mathbb{R}^n$, $\alpha_i = e_i - e_{i+1}$ ($i = 1, \dots, n-1$) and $\alpha_n = e_n$. Then if

$$\Pi = \{\alpha_1, \dots, \alpha_n\},$$

$$\Phi = \{\pm(\alpha_r + \dots + \alpha_s) : 1 \leq r \leq s \leq n\}$$

$$\cup \{\pm(\alpha_r + \dots + \alpha_{s-1} + 2\alpha_s + \dots + 2\alpha_n) : 1 \leq r < s \leq n\}$$

$$= \{\pm e_r : 1 \leq r \leq n\} \cup \{\pm(e_r + e_s), \pm(e_r - e_s) : 1 \leq r < s \leq n\},$$

this is a root system whose Weyl group is the group of $n \times n$ generalised permutation matrices, which has order $|W(\Phi)| = n! 2^n$.

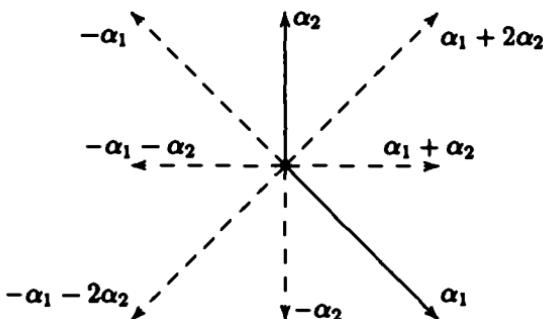


Figure 12.3 The roots of B_2

Example 12.15 (The root system C_n for $n \geq 2$)

Take $V = \mathbb{R}^n$, $\alpha_i = e_i - e_{i+1}$ ($i = 1, \dots, n-1$) and $\alpha_n = 2e_n$. If

$$\Pi = \{\alpha_1, \dots, \alpha_n\},$$

$$\Phi = \{\pm(\alpha_r + \dots + \alpha_s) : 1 \leq r \leq s \leq n\}$$

$$\cup \{\pm(\alpha_r + \dots + \alpha_{s-1} + 2\alpha_s + \dots + 2\alpha_{n-1} + \alpha_n) : 1 \leq r < s < n\}$$

$$\cup \{\pm(2\alpha_r + \dots + 2\alpha_{n-1} + \alpha_n) : 1 \leq r < n\}$$

$$= \{\pm 2e_r : 1 \leq r \leq n\} \cup \{\pm(e_r + e_s), \pm(e_r - e_s) : 1 \leq r < s \leq n\},$$

this is a root system whose Weyl group is again the group of $n \times n$ generalised permutation matrices, which has order $|W(\Phi)| = n! 2^n$.

A careful examination of the roots for B_2 and C_2 shows that apart from the actual lengths of the root vectors and the way they are named, these are essentially the same root systems.

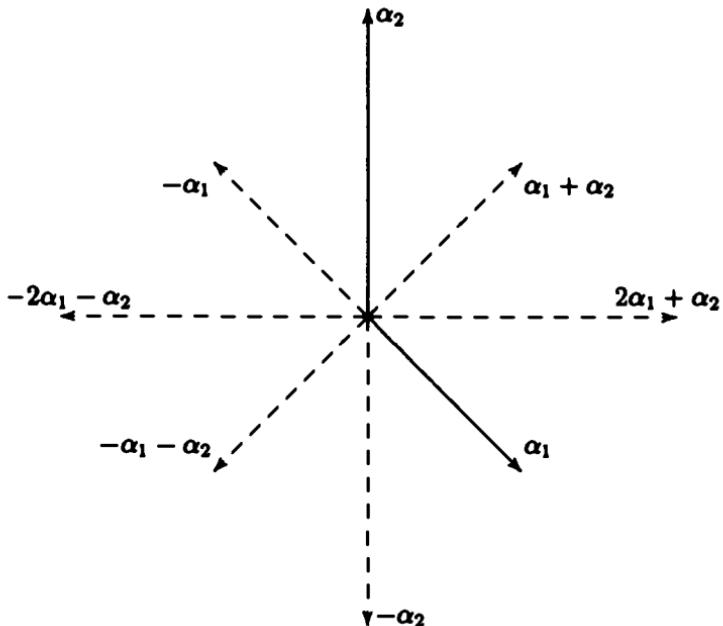


Figure 12.4 The roots of C_2

Example 12.16 (The root system D_n for $n \geq 4$)

Take $V = \mathbb{R}^n$, $\alpha_i = e_i - e_{i+1}$ ($i = 1, \dots, n-1$) and $\alpha_n = e_{n-1} + e_n$. Then if

$$\Pi = \{\alpha_1, \dots, \alpha_n\},$$

$$\Phi = \{\pm(\alpha_r + \dots + \alpha_s) : 1 \leq r \leq s \leq n\}$$

$$\cup \{\pm(\alpha_r + \dots + \alpha_{n-2} + \alpha_n) : 1 \leq r \leq n-2\}$$

$$\cup \{\pm(\alpha_r + \dots + \alpha_{s-1} + 2\alpha_s + \dots + 2\alpha_{n-2} + \alpha_{n-1} + \alpha_n) : 1 \leq r < s \leq n-2\}$$

$$= \{\pm(e_r + e_s), \pm(e_r - e_s) : 1 \leq r < s \leq n\},$$

this is a root system whose Weyl group is the group of $n \times n$ generalised permutation matrices with an even number of sign changes, and this has order $|W(\Phi)| = n! 2^{n-1}$.

12.4 The Dynkin Diagram of a Root System

Let Π be a fundamental system of the root system $\Phi \subseteq V$. Choosing an ordering $\alpha_1, \dots, \alpha_\ell$ for the elements of Π , there are integers

$$a_{ij} = \frac{2 \langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle} \quad (1 \leq i, j \leq \ell).$$

The $\ell \times \ell$ matrix $[a_{ij}]$ with integer entries is called the *Cartan matrix* of the root system Φ .

Defining the angle $\theta_{ij} \in [0, \pi]$ by

$$\theta_{ij} = \cos^{-1} \frac{\langle \alpha_i, \alpha_j \rangle}{|\alpha_i| |\alpha_j|},$$

we have

$$a_{ji} = \frac{2|\alpha_j|}{|\alpha_i|}$$

and so

$$a_{ij} a_{ji} = 4 \cos^2 \theta_{ij}. \quad (12.2)$$

Because a_{ij}, a_{ji} are integers and $0 \leq \cos^2 \theta_{ij} \leq 1$, the expression $4 \cos^2 \theta_{ij}$ has to take one of the values $0, 1, 2, 3, 4$, where the final case only occurs when $\cos \theta_{ij} = \pm 1$ and so $j = i$. In the situation where $j \neq i$, the following seven outcomes are the only ones possible.

$$a_{ij} = 0, \quad a_{ji} = 0, \quad \varphi_{ij} = \frac{\pi}{2}, \quad |\alpha_j| = |\alpha_i|.$$

$$a_{ij} = 1, \quad a_{ji} = 1, \quad \varphi_{ij} = \frac{\pi}{3}, \quad |\alpha_j| = |\alpha_i|.$$

$$a_{ij} = -1, \quad a_{ji} = -1, \quad \varphi_{ij} = \frac{2\pi}{3}, \quad |\alpha_j| = \sqrt{2} |\alpha_i|.$$

$$a_{ij} = 1, \quad a_{ji} = 2, \quad \varphi_{ij} = \frac{\pi}{4}, \quad |\alpha_j| = \sqrt{2} |\alpha_i|.$$

$$a_{ij} = -1, \quad a_{ji} = -2, \quad \varphi_{ij} = \frac{3\pi}{4}, \quad |\alpha_j| = \sqrt{2} |\alpha_i|.$$

$$a_{ij} = 1, \quad a_{ji} = 3, \quad \varphi_{ij} = \frac{\pi}{6}, \quad |\alpha_j| = \sqrt{3} |\alpha_i|.$$

$$a_{ij} = -1, \quad a_{ji} = -3, \quad \varphi_{ij} = \frac{5\pi}{6}, \quad |\alpha_j| = \sqrt{3} |\alpha_i|.$$

Notice that except for the first case where $\varphi_{ij} = \pi/2$, the value of φ_{ij} determines the (unordered) set of numbers

$$\left\{ \frac{|\alpha_i|}{|\alpha_j|}, \frac{|\alpha_j|}{|\alpha_i|} \right\}.$$

We will now show how to associate to a root system $\Pi \subseteq \Phi$ a *Dynkin diagram*. The diagram has one vertex for each simple root labelled by the root, with some bonds joining pairs of them. For distinct simple roots $\alpha_i, \alpha_j \in \Pi$, there are d_{ij} bonds between the corresponding vertices, where

$$d_{ij} = \begin{cases} 0 & \text{if } \varphi_{ij} = \frac{\pi}{2}, \\ 1 & \text{if } \varphi_{ij} = \frac{\pi}{3}, \frac{2\pi}{3}, \\ 2 & \text{if } \varphi_{ij} = \frac{\pi}{4}, \frac{3\pi}{4}, \\ 3 & \text{if } \varphi_{ij} = \frac{\pi}{6}, \frac{5\pi}{6}. \end{cases}$$

In the cases where $d_{ij} = 2, 3$, an inequality sign $>$ is added indicating which of the roots is larger. Here are the possible connections between nodes that can occur.

	Labels: α_i α_j
$d_{ij} = 0$	○ ○
$d_{ij} = 1$	○ — ○
$d_{ij} = 2$ $ \alpha_i < \alpha_j $	○ ← → ○
$d_{ij} = 3$ $ \alpha_i < \alpha_j $	○ ← ← → ○

The whole Dynkin diagram is obtained by filling in the bonds between the vertices for all the simple roots. Such a diagram is *irreducible* or *connected* if it has only one path component.

12.5 Irreducible Dynkin Diagrams

The connected *irreducible Dynkin diagrams* are given in Tables 12.1 and 12.2. Each of the four classical series A_n , B_n , C_n and D_n consists of infinitely many distinct examples, while there are finitely many exceptional ones E_6 , E_7 , E_8 , F_4 , G_2 (the root system of G_2 is shown in Figure 12.5). The classical Dynkin diagrams are related to familiar matrix groups, where the entries at the bottom of Table 12.1 are consequences of certain special isomorphisms. Lie groups associated with the exceptional Dynkin diagrams are constructed in Adams [2]. The subscript n is the rank of the associated Lie group, and also the number of nodes in the Dynkin diagram.

Root system	Matrix group	Dynkin diagram
A_n ($n \geq 1$)	$SU(n)$	
B_n ($n \geq 2$)	$\text{Spin}(2n+1), \text{SO}(2n+1)$	
C_n ($n \geq 3$)	$Sp(n)$	
D_n ($n \geq 4$)	$\text{Spin}(2n), \text{SO}(2n)$	
A_1	$SU(2) \cong \text{Spin}(3), \text{SO}(3)$	
$B_2 = C_2$	$Sp(2) \cong \text{Spin}(5), \text{SO}(5)$	

Table 12.1 Classical Dynkin diagrams and associated matrix groups

Root system	Dynkin diagram
E_6	
E_7	
E_8	
F_4	
G_2	

Table 12.2 Exceptional Dynkin diagrams

12.6 From Root Systems to Lie Algebras

Now assume that $\Pi = \{\alpha_1, \dots, \alpha_\ell\}$ is a set of simple roots for a root system $\Phi \subseteq \mathfrak{t}^*$, the linear dual of an \mathbb{R} -vector space \mathfrak{t} of dimension ℓ , equipped with a positive definite inner product $(\cdot | \cdot)$. Using the duality isomorphism $\nu: \mathfrak{t} \rightarrow \mathfrak{t}^*$,

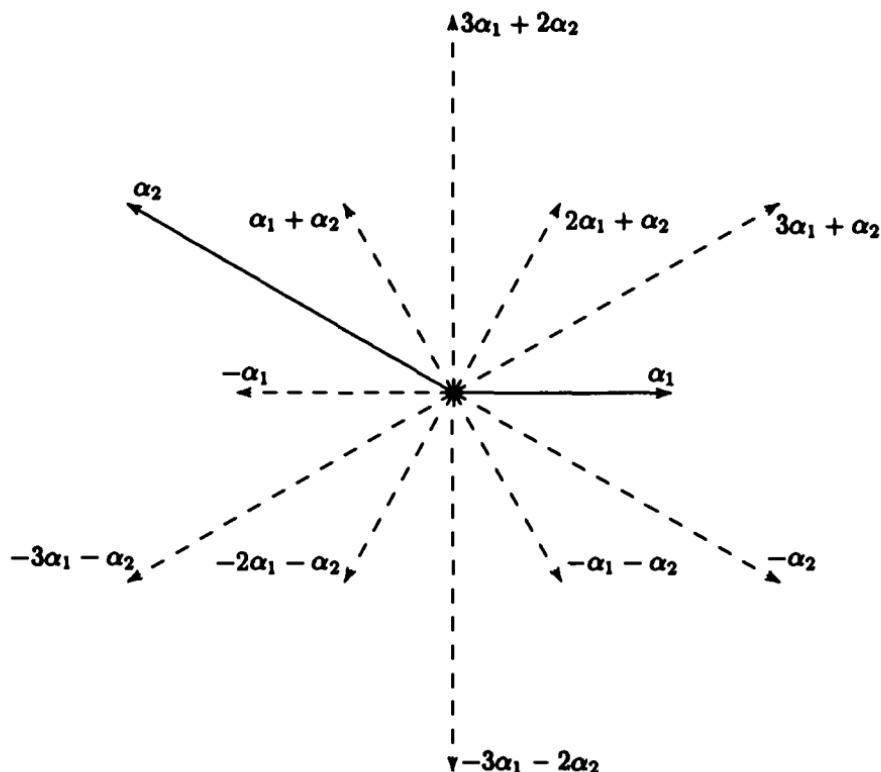


Figure 12.5 The roots of G_2

we can define elements

$$\check{\alpha} = \frac{2}{(\alpha | \alpha)} \nu^{-1}(\alpha) \in \mathfrak{t} \quad (\alpha \in \Phi).$$

These are called the *coroots*. We have the formula

$$\begin{aligned} \alpha(\check{\beta}) &= \frac{2}{(\beta | \beta)} (\nu^{-1}(\alpha) | \nu^{-1}(\beta)) \\ &= \frac{2}{(\beta | \beta)} (\alpha | \beta). \end{aligned} \tag{12.3}$$

We write

$$\check{\Phi} = \{\check{\alpha} : \alpha \in \Phi\}, \quad \check{\Pi} = \{\check{\alpha} : \alpha \in \Pi\},$$

for the sets of all coroots and *simple coroots* respectively. For each coroot $\check{\alpha} \in \check{\Phi}$, define $s_{\check{\alpha}} t$ by the formula analogous to Equation (12.1).

Proposition 12.17

$\check{\Phi} \subseteq t$ is a root system with $\check{\Pi}$ a set of simple roots.

The root system $\check{\Pi} \subseteq \check{\Phi}$ is called the *dual root system* to $\Pi \subseteq \Phi$.

Recall that for a finite dimensional vector space, $V \cong V^{**}$ and given a positive definite inner product such an isomorphism is obtained as the composite

$$V \xrightarrow{\nu} V^* \xrightarrow{\nu'} V^{**},$$

where $\nu': V^* \rightarrow V^{**}$ is obtained using the above inner product on V^* . It is easy to verify that given a root system in $\Phi' \subseteq V^{**}$, for each $\alpha \in \Phi'$,

$$\check{\alpha} = \nu^{-1}(\nu')^{-1}(\alpha). \quad (12.4)$$

This means that we can start with a root system in V or in V^* and pass back and forth by taking coroots.

The following important result is proved in [14]. We will illustrate it with the examples provided by the root systems A_n .

Theorem 12.18

Given a root system $\Phi \subseteq t^*$ with simple roots $\Pi = \{\alpha_1, \dots, \alpha_\ell\}$, there is a \mathbb{C} -Lie algebra g associated to Φ for which

- $t_{\mathbb{C}} \subseteq g$ is a maximal abelian \mathbb{C} -Lie subalgebra;
- for each root $\alpha \in \Phi$, there are elements $e_\alpha, f_\alpha \in g$ satisfying

$$[e_\alpha, f_\beta] = \delta_{ij} \check{\alpha}, \quad [t, e_\alpha] = \alpha(t) e_\alpha, \quad [t, f_\alpha] = -\alpha(t) f_\alpha,$$

for all $t \in t_{\mathbb{C}}$;

- The elements $\alpha_1, \dots, \alpha_\ell, e_{\alpha_1}, f_{\alpha_1}, \dots, e_{\alpha_\ell}, f_{\alpha_\ell}$ form a basis of g .

If g' is a second such Lie algebra, there is a unique isomorphism of Lie algebras $g \rightarrow g'$ which extends the identity function $t_{\mathbb{C}} \rightarrow t_{\mathbb{C}} \subseteq g'$.

Because of the essential uniqueness provided by this result, we will write $g(\Phi)$ for the Lie algebra associated with a root system Φ .

Example 12.19

Consider the root system A_n of Example 12.13. Then the complexification of $\mathfrak{su}(n+1)$ is $\mathfrak{sl}_{n+1}(\mathbb{C})$. We take

$$t = \{\text{diag}(t_1 i, \dots, t_{n+1} i) : t_k \in \mathbb{R}\},$$

and

$$\check{\alpha}_k = iE^{k \cdot k} - iE^{(k+1) \cdot (k+1)} \quad (k = 1, \dots, n).$$

Then the coroots are the matrices of the form

$$\check{\alpha} = \sum_{k=1}^n c_k \check{\alpha}_k \quad (c_k = 0, \pm 1),$$

satisfying the inner product formulæ of Equation (11.4). The dual roots α satisfy

$$\alpha(\alpha_k) = (\alpha | \alpha_k) = (\check{\alpha} | \check{\alpha}_k).$$

For each root α , we take

$$e_\alpha = \sum_{k=1}^n c_k (E^{k \cdot (k+1)}), \quad f_\alpha = \sum_{k=1}^n c_k (E^{(k+1) \cdot k}).$$

Then there is an isomorphism of \mathbb{C} -Lie algebras $\mathfrak{g}(A_n) \cong \mathfrak{sl}_{n+1}(\mathbb{C})$.

EXERCISES

- 12.1. Verify Remark 12.9.
- 12.2. Let Φ_{B_2} and Φ_{C_2} be the sets of roots for B_2 and C_2 . Find a bijection $f: \Phi_{B_2} \rightarrow \Phi_{C_2}$ which preserves angles but not lengths.
- 12.3. $\Delta \Delta$ Using the ideas of Section 11.5, verify as many as you can of the identifications of root systems of Table 12.1.
- 12.4. $\Delta \Delta$ For $n \geq 3$, determine the roots of $U(n)$ with respect to the standard maximal torus

$$T_n = \{\text{diag}(z_1, \dots, z_n) : |z_1| = \dots = |z_n| = 1\} \leq U(n).$$

Find the normaliser of T_n in $U(n)$ and the action of the Weyl group on the roots.

Hints and Solutions to Selected Exercises

Chapter 1

1.1. Using the formula of Remark 1.4 we obtain

$$\left\| \begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix} \right\| = \max\{|u|, |v|\}, \quad \left\| \begin{bmatrix} u & 1 \\ 0 & u \end{bmatrix} \right\| = \frac{\sqrt{2|u|^2 + 1 + \sqrt{4|u|^2 + 1}}}{\sqrt{2}},$$
$$\left\| \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \right\| = 1, \quad \left\| \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix} \right\| = \max\{e^t, e^{-t}\}.$$

1.2. (a) From the definition of the matrix norm,

$$\begin{aligned} \|BAB^{-1}\| &= \max \left\{ \frac{|BAB^{-1}x|}{|x|} : x \neq 0 \right\} = \max \left\{ \frac{|Ax|}{|Bx|} : x \neq 0 \right\} \\ &= \max \left\{ \frac{|Ax|}{|x|} : x \neq 0 \right\} = \|A\|. \end{aligned}$$

(b) The inequality $\|CAC^{-1}\| \leq \|C\| \|A\| \|C^{-1}\|$ always holds.

1.3. (a) We have $|x|^2 = x^*x$ by definition, so

$$\|A\|^2 = \sup\{|Ax|^2 : x \in \mathbb{C}^n, |x| = 1\}.$$

Since the unit sphere $\{x \in \mathbb{C}^n : |x| = 1\}$ is closed and bounded, the continuous function $x \mapsto |Ax|^2$ defined on it attains its supremum, hence

$$\|A\|^2 = \max\{|Ax|^2 : x \in \mathbb{C}^n, |x| = 1\}.$$

(b) The matrix A^*A is hermitian and for any $x \in \mathbb{C}^n$, $x^*A^*Ax = |Ax|^2$, which is real and non-negative; it also vanishes precisely when $Ax = 0$. By the general theory of hermitian matrices, there is an orthonormal basis $\{u_1, \dots, u_n\}$ consisting of eigenvalues $\lambda_1, \dots, \lambda_n$ which are all real. For each j ,

$$u_j^* A^* A u_j = \lambda_j u_j^* u_j = \lambda_j$$

since $\mathbf{u}_j^* \mathbf{u}_j = 1$; hence $\lambda_j \geq 0$. If $A \neq O$ there must be at least one non-zero eigenvalue of $A^* A$ since otherwise $|Ax|^2$ would always vanish; so in this case we have a largest positive eigenvalue λ which we might as well assume is $\lambda = \lambda_1$. Now if $|\mathbf{x}| = 1$ we can write $\mathbf{x} = t_1 \mathbf{u}_1 + \cdots + t_n \mathbf{u}_n$, where $t_j \in \mathbb{C}$ and $|t_1|^2 + \cdots + |t_n|^2 = 1$. Then

$$\begin{aligned}\mathbf{x}^* A^* A \mathbf{x} &= \mathbf{x}^* (\lambda_1 t_1 \mathbf{u}_1 + \cdots + \lambda_n t_n \mathbf{u}_n) = \lambda_1 |t_1|^2 + \cdots + \lambda_n |t_n|^2 \\ &\leq \lambda_1 |t_1|^2 + \cdots + \lambda_1 |t_n|^2 = \lambda_1,\end{aligned}$$

hence $\|A\|^2 \leq \lambda_1$ and in fact we have equality since $\mathbf{u}_1^* A^* A \mathbf{u}_1 = \lambda_1$.

1.4. (a) Let

$$\ell = \frac{1}{2} \left(1 + \lim_{r \rightarrow \infty} \frac{\|A_{r+1}\|}{\|A_r\|} \right) < 1.$$

For large enough m we have

$$\begin{aligned}\|A_m + \cdots + A_{m+k}\| &\leq \|A_m\| + \cdots + \|A_{m+k}\| \\ &= \|A_m\| \left(1 + \frac{\|A_{m+1}\|}{\|A_m\|} + \frac{\|A_{m+1}\| \|A_{m+2}\|}{\|A_m\| \|A_{m+1}\|} + \cdots + \frac{\|A_{m+1}\| \cdots \|A_{m+k}\|}{\|A_m\| \cdots \|A_{m+k-1}\|} \right) \\ &< \|A_m\| \left(1 + \ell + \ell^2 + \cdots + \ell^{k-1} \right) = \|A_m\| \frac{(1 - \ell^k)}{(1 - \ell)} \rightarrow \frac{\|A_m\|}{1 - \ell} \quad \text{as } k \rightarrow \infty.\end{aligned}$$

Also, if m_0 is large enough, then for $m > m_0$, $\frac{\|A_m\|}{\|A_{m-1}\|} < \ell$, so

$$\|A_m\| = \frac{\|A_m\| \cdots \|A_{m_0+1}\|}{\|A_{m-1}\| \cdots \|A_{m_0}\|} \|A_{m_0}\| < \ell^{m-m_0} \|A_{m_0}\| \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

Combining these we see that the partial sums of the series form a Cauchy sequence and hence the series converges.

(b) Here $\|A_m\| \rightarrow \infty$ as $m \rightarrow \infty$ and the series diverges.

(c) Any of the usual tests based on absolute convergence works.

1.5. (a) For $m, n \geq 1$,

$$\begin{aligned}\left\| \sum_{r=m}^{m+n} A^r \right\| &= \|A^m + A^{m+1} + \cdots + A^{m+n}\| \leq \|A\|^m + \|A\|^{m+1} + \cdots + \|A\|^{m+n} \\ &= \|A\|^m \frac{(1 - \|A\|^{n+1})}{(1 - \|A\|)} \rightarrow 0 \quad \text{as } m, n \rightarrow \infty.\end{aligned}$$

So the sequence of partial sums is a Cauchy sequence. This can also be done using the ratio test of Exercise 1.4.

(b) For any $n \geq 1$,

$$(I - A)(I + A + A^2 + \cdots + A^n) = I - A^{n+1} = (I + A + A^2 + \cdots + A^n)(I - A),$$

so

$$\|(I - A)(I + A + A^2 + \cdots + A^n) - I\| = \|A^{n+1}\| \leq \|A\|^{n+1} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

and similarly for $\|(I + A + A^2 + \cdots + A^n)(I - A) - I\|$. Hence $I - A$ has inverse

$$(I - A)^{-1} = \sum_{r=0}^{\infty} A^r.$$

(c) If $A^k = O$, then

$$(I - A)^{-1} = \sum_{r=0}^{k-1} A^r, \quad \exp(A) = \sum_{r=0}^{k-1} \frac{1}{r!} A^r.$$

1.6. (a) The subset

$$\mathrm{O}(n) = \{A \in M_n(\mathbb{R}) : A^T A - I = O\} \subseteq M_n(\mathbb{R})$$

is bounded since for $A \in \mathrm{O}(n)$, $\|A\| = 1$. It is also closed since if a sequence $A_r \in \mathrm{O}(n)$ converges to $A \in M_n(\mathbb{R})$ then

$$A^T A - I = \lim_{r \rightarrow \infty} (A_r^T A_r - I) = \lim_{r \rightarrow \infty} O = O.$$

Therefore $\mathrm{O}(n)$ is compact.

(b) The subset

$$\mathrm{U}(n) = \{A \in M_n(\mathbb{C}) : A^* A - I = O\} \subseteq M_n(\mathbb{C})$$

is compact by a similar argument to (a).

(c) Consider the $n \times n$ diagonal matrices

$$A_k = \text{diag}(k, 1/k, 1, \dots, 1) \quad (k \geq 1).$$

Then $\det A_k = 1$ so $A_k \in \mathrm{SL}_n(\mathbb{k}) \leqslant \mathrm{GL}_n(\mathbb{k})$. But $\|A_k\| = k \rightarrow \infty$ as $k \rightarrow \infty$, so this sequence is unbounded.

1.7. (a) If $A, B \in \overline{H}$ then there are sequences A_r, B_r in H with $A_r \rightarrow A$ and $B_r \rightarrow B$ as $r \rightarrow \infty$. Hence $A_r, B_r \rightarrow AB$ and so $AB \in \overline{H}$ since each $A_r, B_r \in H$. Similarly, $A_r^{-1} \rightarrow A^{-1}$, showing that $A^{-1} \in \overline{H}$.

(b) Here is an alternative proof in terms of open sets which applies to any topological group.

Let $u, v \in \overline{H}$ and consider uv . If $uv \notin \overline{H}$, the open set $G - \overline{H}$ contains uv ; notice that $G - \overline{H}$ is the biggest open subset of G which does not intersect \overline{H} . Since the product map mult is continuous, there is an open set of the form $U \times V \subseteq G \times G$ contained in the open set $\text{mult}^{-1}(G - \overline{H})$ with $u \in U$ and $v \in V$. Notice that $U \cap H = \emptyset = V \cap H$ by the above remark. Now take $u' \in U \cap H$ and $v' \in V \cap H$ and note that $u'v' = \text{mult}(u', v') \in H$ and also $\text{mult}(u', v') \in G - \overline{H}$, giving a contradiction. So $uv \in \overline{H}$. Also, if $w \in \overline{H}$, suppose that $w^{-1} \in G - \overline{H}$. Then $\text{inv}^{-1}(G - \overline{H}) \subseteq G$ is open, where inv is the inverse map. But as $w \in \text{inv}^{-1}(G - \overline{H})$, there must be an element $h \in H \cap \text{inv}^{-1}(G - \overline{H})$ and so $h^{-1} \in G - \overline{H}$ which is impossible. Hence $\overline{H} \leqslant G$.

(c) Consider any sequence of diagonal matrices

$$A_k = \text{diag}(a_k, 1, \dots, 1) \quad (k \geq 1)$$

where $a_k \leq 1$ and $a_k \rightarrow \sqrt{2}$ as $k \rightarrow \infty$ ($\sqrt{2}$ could be replaced by any other irrational number greater than 1). Then $A_k \rightarrow \text{diag}(\sqrt{2}, 1, \dots, 1) \notin \Gamma$. We must have $\Gamma = \mathrm{GL}_n(\mathbb{R})$ since for every $B = [b_{ij}] \in \mathrm{GL}_n(\mathbb{R})$ we can find sequences $b_{ij}^{(k)} \in \mathbb{Q}$ satisfying $b_{ij}^{(k)} \rightarrow b_{ij}$ as $k \rightarrow \infty$.

1.8. By Proposition 1.26, $G \subseteq M_n(\mathbb{k})$ is a closed subset of $\mathrm{GL}_n(\mathbb{R})$.

1.9. Suppose that $G \leqslant \mathrm{GL}_n(\mathbb{R}) \subseteq M_n(\mathbb{R})$ for some n . Then $H \subseteq M_n(\mathbb{R})$ is closed and bounded since G is compact. For each $h \in H$ the set $\{h\} \subseteq H$ is open and these form

an open cover of H by disjoint open sets. By the Heine-Borel Theorem 1.20 there must be finitely many of these sets that cover H , so H is itself finite.

1.10. In each case, we have subgroups $G_{n+1} \leq \mathrm{GL}_{n+1}(\mathbf{k})$ (where $\mathbf{k} = \mathbb{R}$ or \mathbb{C}) and $G_n \leq \mathrm{GL}_n(\mathbf{k})$ where G_{n+1} is closed in $\mathrm{GL}_{n+1}(\mathbf{k})$ and $G_n = G_{n+1} \cap \mathrm{GL}_n(\mathbf{k})$. As $\mathrm{GL}_n(\mathbf{k})$ is closed in $\mathrm{GL}_{n+1}(\mathbf{k})$, G_n is closed in G_{n+1} .

1.11. (a) If $A \in \mathrm{Symp}_{2m}(\mathbb{R})$, then $A^T J_{2m} A = J_{2m}$ and as $\det J_{2m} = 1$, we have $\det A^2 = 1$, so $\det A = \pm 1$.

(b) Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{Symp}_2(\mathbb{R})$. By definition,

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & ad - bc \\ -ad + bc & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

so $A \in \mathrm{Symp}_2(\mathbb{R})$ if and only if $\det A = 1$; hence $\mathrm{Symp}_2(\mathbb{R}) \leq \mathrm{SL}_2(\mathbb{R})$.

1.12. Clearly $\rho_n U(n) \leq \rho_n \mathrm{GL}_n(\mathbb{C})$. If $Z = [z_{rs}] \in M_n(\mathbb{C})$ where $z_{rs} = x_{rs} + y_{rs}i$ with $x_{rs}, y_{rs} \in \mathbb{R}$, then

$$\rho_n(Z^*) = \rho_n([x_{sr}] - [y_{sr}i]) = \rho_n(Z)^T.$$

Hence $Z \in U(n)$ if and only if $\rho_n(Z) \in O(2n)$. If $A \in O(2n)$, then $A \in \mathrm{Symp}_{2n}(\mathbb{R})$ if and only if $AJ_{2n} = J_{2n}A$ and by Proposition 1.44 this is equivalent to $A \in \rho_n U(n)$. This shows that $\rho_n U(n) = O(2n) \cap \mathrm{Symp}_{2n}(\mathbb{R})$. If $B \in \mathrm{Symp}_{2n}(\mathbb{R})$ then $BJ_{2n} = J_{2n}B$ if and only if $B^T = B$. This shows that $\rho_n \mathrm{GL}_n(\mathbb{C}) \cap \mathrm{Symp}_{2n}(\mathbb{R}) = O(2n) \cap \mathrm{Symp}_{2n}(\mathbb{R})$. The second collection of equalities is proved in a similar way.

1.13. (a) The triangle inequality for ρ follows from the following sequence of inequalities valid for any $(x_1, x_2), (y_1, y_2), (z_1, z_2) \in X_1 \times X_2$,

$$\begin{aligned} \rho((x_1, x_2), (z_1, z_2)) &= \sqrt{\rho_1(x_1, z_1)^2 + \rho_2(x_2, z_2)^2} \\ &\leq \sqrt{(\rho_1(x_1, y_1) + \rho_1(y_1, z_1))^2} + \sqrt{(\rho_2(x_2, y_2) + \rho_2(y_2, z_2))^2} \\ &\quad [\text{by the triangle inequalities for } \rho_1 \text{ and } \rho_2] \\ &\leq \sqrt{\rho_1(x_1, y_1)^2 + \rho_2(x_2, y_2)^2} + \sqrt{\rho_1(y_1, z_1)^2 + \rho_2(y_2, z_2)^2} \\ &\quad [\text{by the usual triangle inequality for } \mathbb{R}^2] \\ &= \rho((x_1, x_2), (y_1, y_2)) + \rho((y_1, y_2), (z_1, z_2)). \end{aligned}$$

The other properties required for ρ to be a metric are straightforward to check. If $U_1 \times U_2 \subseteq X_1 \times X_2$ with $U_1 \subseteq X_1$ and $U_2 \subseteq X_2$ open, then for each $x_1 \in U_1$ and $x_2 \in U_2$ there are open discs $N_{X_1, \rho_1}(x_1; r_1) \subseteq U_1$ and $N_{X_2, \rho_2}(x_2; r_2) \subseteq U_2$; for $r = \min\{r_1, r_2\}$, we have $N_{X_1 \times X_2, \rho}((x_1, x_2); r) \subseteq U_1 \times U_2$, so $U_1 \times U_2$ is open with respect to ρ . Conversely, for an open disc $N_{X_1 \times X_2, \rho}((x_1, x_2); r)$ we have

$$N_{X_1, \rho_1}(x_1; r/\sqrt{2}) \times N_{X_2, \rho_2}(x_2; r/\sqrt{2}) \subseteq N_{X_1 \times X_2, \rho}((x_1, x_2); r).$$

Hence every subset of $X_1 \times X_2$ open with respect to ρ is a union of products of open discs in X_1 and X_2 . So the topology associated with ρ has the same open sets as the product topology.

(b) If $(x_{1,r}, x_{2,r}) \rightarrow (x_1, x_2)$ say, then for $i = 1, 2$,

$$\rho_i(x_{i,r}, x_i) \leq \rho((x_{1,r}, x_{2,r}), (x_1, x_2)),$$

so $x_{i,r} \rightarrow x_i$ as $r \rightarrow \infty$. Conversely, if $x_{1,r} \rightarrow x_1$ and $x_{2,r} \rightarrow x_2$ then

$$\rho((x_{1,r}, x_{2,r}), (x_1, x_2)) = \sqrt{\rho_1(x_{1,r}, x_1)^2 + \rho_2(x_{2,r}, x_2)^2} \rightarrow 0,$$

so $(x_{1,r}, x_{2,r}) \rightarrow (x_1, x_2)$ as $r \rightarrow \infty$.

1.14. (a) If $g \notin \text{Stab}_G(x)$ then $gx \in X - \{x\}$. Since $X - \{x\}$ is an open subset of X , $\mu^{-1}(X - \{x\}) \subseteq G \times X$ is open, so there are open sets $U \subseteq G$, $V \subseteq X$ with $g \in U$, $x \in V$ and $U \times V \subseteq \mu^{-1}(X - \{x\})$. But then for $h \in U$ we must have $hx \neq x$, hence $U \subseteq G - \text{Stab}_G(x)$. This shows that $G - \text{Stab}_G(x)$ is open in G .

(b) If $g \notin \text{Stab}_G(W)$ then $gW \not\subseteq W$ or $g^{-1}W \not\subseteq W$. If $gW \not\subseteq W$ then for some $w \in W$, $gw \neq w$. Now a similar argument to (a) shows that there is an open set $U \subseteq G - \text{Stab}_G(W)$ containing g . If $g^{-1}W \not\subseteq W$ then we find an open set $U' \subseteq G - \text{Stab}_G(W)$ containing g^{-1} ; applying the inverse map inv we obtain an open set $\text{inv } U' \subseteq G - \text{Stab}_G(W)$ containing g . Finally, since each stabiliser $\text{Stab}_G(w)$ is closed the intersection of any collection of such stabilisers is also closed.

1.15. (a) Taking ρ_1 to be the norm metric on $M_n(\mathbf{k})$ and ρ_2 to be the usual metric on \mathbf{k}^n , define ρ as in Exercise 1.13., i.e., set

$$\rho(A, \mathbf{x}) = \sqrt{\|A\|^2 + |\mathbf{x}|^2}.$$

If $A, B \in M_n(\mathbf{k})$ and $\mathbf{x}, \mathbf{y} \in \mathbf{k}^n$, then

$$|Ax - By| = |(Ax - Bx) + (Bx - By)| \leq \|A - B\| |\mathbf{x}| + \|B\| |\mathbf{x} - \mathbf{y}|.$$

A routine ϵ - δ argument now establishes continuity of the product map at each (A, \mathbf{x}) .

(b) These are consequences of results in the previous exercise which show that these stabiliser subgroups are closed.

(c) We have

$$\text{Stab}_{GL_n(\mathbb{R})}(e_n) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ B & 1 \end{bmatrix} : A \in GL_{n-1}(\mathbb{R}), B \in M_{1,n-1}(\mathbb{R}) \right\},$$

$$\text{Stab}_{GL_n(\mathbb{R})}(X) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ B & t \end{bmatrix} : A \in GL_{n-1}(\mathbb{R}), B \in M_{1,n-1}(\mathbb{R}), t \in \mathbb{R} \right\},$$

$$\text{Stab}_{SL_n(\mathbb{R})}(e_n) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ B & 1 \end{bmatrix} : A \in SL_{n-1}(\mathbb{R}), B \in M_{1,n-1}(\mathbb{R}) \right\},$$

$$\text{Stab}_{SL_n(\mathbb{R})}(X) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ B & \det A^{-1} \end{bmatrix} : A \in GL_{n-1}(\mathbb{R}), B \in M_{1,n-1}(\mathbb{R}) \right\},$$

$$\text{Stab}_{O(n)}(e_n) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ O_{1,n-1} & 1 \end{bmatrix} : A \in O(n-1), \right\},$$

$$\text{Stab}_{O(n)}(X) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ O_{1,n-1} & t \end{bmatrix} : A \in O(n-1), t = \pm 1 \right\},$$

$$\text{Stab}_{SO(n)}(e_n) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ O_{1,n-1} & 1 \end{bmatrix} : A \in SO(n-1), \right\},$$

$$\text{Stab}_{SO(n)}(X) = \left\{ \begin{bmatrix} A & O_{n-1,1} \\ O_{1,n-1} & t \end{bmatrix} : A \in O(n-1), t = \det A \right\}.$$

1.16. (b) We have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} x^n = (ax + cy)^n,$$

so

$$\text{Stab}_{\text{SU}(2)}(x^n) = \left\{ \begin{bmatrix} u & 0 \\ 0 & \bar{u} \end{bmatrix} : u^n = 1 \right\}.$$

(c) We have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} xy = (ax + cy)(bx + dy) = abx^2 + (ad + bc)xy + cdy^2,$$

so

$$\text{Stab}_{\text{SU}(2)}(xy) = \left\{ \begin{bmatrix} u & 0 \\ 0 & \bar{u} \end{bmatrix} : |u| = 1 \right\} \cup \left\{ \begin{bmatrix} 0 & v \\ -\bar{v} & 0 \end{bmatrix} : |v| = 1 \right\}.$$

(d) If $A \in \ker \varphi_n$, then $A \in \text{Stab}_{\text{SU}(2)}(x^n)$, so by (b),

$$A = \begin{bmatrix} u & 0 \\ 0 & \bar{u} \end{bmatrix} \quad \text{with } u^n = 1.$$

Also, for each $r = 1, \dots, n$,

$$Ax^r y^{n-r} = (ux)^r (\bar{u}y)^{n-r} = u^{2r} x^r y^{n-r} = x^r y^{n-r},$$

so this can only happen if $u^{2r} = 1$ for all such r ; hence $u^2 = 1$. If n is odd, this gives $u = 1$, while if n is even $u = \pm 1$.

Chapter 2

2.1. For $n \geq 0$ we have

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^{2n} = (-1)^n \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^{2n+1} = (-1)^n \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

hence

$$\exp \left(\begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix} \right) = \begin{bmatrix} \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} & \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} \\ \sum_{n=0}^{\infty} \frac{(-1)^{n+1} t^{2n+1}}{(2n+1)!} & \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \end{bmatrix} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}.$$

Similarly,

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{2n} = I_2, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{2n+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

hence

$$\exp \left(\begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix} \right) = \begin{bmatrix} \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} & \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} \\ \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} & \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} \end{bmatrix} = \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}.$$

Finally, by an easy induction on n ,

$$\begin{bmatrix} t & 0 \\ -2 & t \end{bmatrix}^n = \begin{bmatrix} t^n & 0 \\ -2n & t^n \end{bmatrix}.$$

hence

$$\exp\left(\begin{bmatrix} t & 0 \\ -2 & t \end{bmatrix}\right) = \begin{bmatrix} \sum_{n=0}^{\infty} \frac{t^n}{n!} & 0 \\ -2 \sum_{n=0}^{\infty} \frac{nt^{n-1}}{n!} & \sum_{n=0}^{\infty} \frac{t^n}{n!} \end{bmatrix} = \begin{bmatrix} e^t & 0 \\ -2 \sum_{n=0}^{\infty} \frac{t^n}{n!} & e^t \end{bmatrix} = \begin{bmatrix} e^t & 0 \\ -2e^t & e^t \end{bmatrix}.$$

2.2. (a) Each partial sum has the form

$$\sum_{r=0}^n \frac{1}{r!} (BAB^{-1})^r = B \left(\sum_{r=0}^n \frac{1}{r!} A^r \right) B^{-1}.$$

(b) Take $A = D$ and $B = C$ in (a) to obtain

$$\exp(D) = C \exp(\text{diag}(\lambda_1, \dots, \lambda_n)) C^{-1} = C \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) C^{-1}.$$

(c) The eigenvalues of $\begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix}$ are $\pm ti$ and the matrix $C = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$ satisfies

$$\begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix} = C \text{diag}(ti, -ti) C^{-1}.$$

Now using the identities $\cos t = (e^{ti} + e^{-ti})/2$ and $\sin t = (e^{ti} - e^{-ti})/2i$ we obtain

$$\exp\left(\begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix}\right) = C \text{diag}(e^{ti}, e^{-ti}) C^{-1} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}.$$

Similarly,

$$\begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix} = C \text{diag}(t, -t) C^{-1},$$

where $C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Using the identities $\cosh t = (e^t + e^{-t})/2$ and $\sinh t = (e^t - e^{-t})/2$ we obtain

$$\exp\left(\begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix}\right) = C \text{diag}(e^t, e^{-t}) C^{-1} = \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}.$$

2.3. (a) Every positive power of N is strictly upper triangular and N is nilpotent, say $N^k = O$. So $\exp(N)$ is upper triangular with 1's down its main diagonal.

(b) Write $N = tI + U$, where $t \in \mathbf{k}$ and U is strictly upper triangular. Then U is nilpotent, say U^ℓ for some $\ell > 0$. So for $m \geq 1$ we have

$$N^m = (tI + U)^m = \sum_{r=0}^m \binom{m}{r} t^{m-r} U^r.$$

For $m \geq \ell$ this becomes

$$N^m = \sum_{r=0}^{\ell-1} \binom{m}{r} t^{m-r} U^r.$$

Thus

$$\exp(N) = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{r=0}^{\min\{\ell-1, m\}} \binom{m}{r} t^{m-r} U^r = \sum_{r=0}^{\ell-1} \left(\sum_{m=0}^{\infty} \frac{1}{m!} \binom{m}{r} t^{m-r} \right) U^r.$$

2.4. (a) The operation of transposition is a continuous function on $M_n(\mathbb{R})$ so it commutes with taking limits of sequences. Alternatively, the partial sums of $\exp(S)$ satisfy

$$\left(\sum_{k=0}^n \frac{1}{k!} S^k \right)^T = \sum_{k=0}^n \frac{1}{k!} (S^T)^k = \sum_{k=0}^n \frac{1}{k!} (-S)^k = \exp(-S) = \exp(S)^{-1},$$

hence $\exp(S)^T = \exp(S)^{-1}$.

(b) Similarly, the hermitian conjugate of the exponential satisfies $\exp(S)^* = \exp(S)^{-1}$.

2.5. (a) Start by solving the differential equation

$$\alpha'(t) = \alpha(t) \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix}, \quad \alpha(0) = I.$$

The solution is

$$\alpha(t) = \begin{bmatrix} e^{-t} & 2(e^{-t} - e^t) \\ 0 & e^t \end{bmatrix}.$$

The desired solution is

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} e^{-t} & 2(e^{-t} - e^t) \\ 0 & e^t \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3e^{-t} - 2e^t \\ 2e^t \end{bmatrix}.$$

(b) The solutions are

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} -2e^{-t} \sin t \\ e^{-t}(\cos t - \sin t) \end{bmatrix}, \quad \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} e^{-t} \\ e^{-t} \end{bmatrix}.$$

2.6. A solution has the form $x(t) = \exp(tA)x(0)$ for $t \in \mathbb{R}$. Since $A^* = -A$ we have $\exp(tA)^T = \exp(-tA) = \exp(tA)^{-1}$, hence $\exp(tA)$ is orthogonal. So

$$|x(t)|^2 = x(t) \cdot x(t) = (\exp(tA)x(0)) \cdot (\exp(tA)x(0)) = x(0) \cdot x(0) = |x(0)|^2.$$

We also have

$$x(t) \cdot x'(t) = \frac{1}{2} \frac{d}{dt} |x(t)|^2 = 0,$$

showing that $x(t)$ and $x'(t)$ are orthogonal.

2.7. (a) Choose any sequence of matrices

$$A_n = \begin{bmatrix} \lambda_{1,n} & 1 & 0 & \dots & 0 \\ 0 & \lambda_{2,n} & 1 & \dots & \vdots \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \dots & \dots & 0 & \lambda_{r-1,n} & 1 \\ 0 & \dots & \dots & 0 & \lambda_{r,n} \end{bmatrix}$$

in which the r diagonal λ 's in $J(\lambda, r)$ have been replaced by r sequences of non-zero terms $\lambda_{1,n}, \dots, \lambda_{r,n}$ that satisfy

$$\lambda_{i,n} \neq \lambda_{j,n} \quad \text{if } i \neq j, \quad \lambda_{j,n} \rightarrow \lambda \quad \text{as } n \rightarrow \infty.$$

Then each matrix A_n has r distinct eigenvalues so is diagonalisable and $A_n \rightarrow J(\lambda, r)$ as $n \rightarrow \infty$.

(b) This follows from (a) together with Theorem 2.9.

Chapter 3

3.1. (a) Any abelian Lie algebras of the same dimension are isomorphic since they are isomorphic as vector spaces. So let \mathfrak{g} be any 2-dimensional k -Lie algebra which is not abelian. Then there are elements $x, y \in \mathfrak{g}$ for which $[x, y] \neq 0$; these elements cannot be linearly dependent so they form a basis of the k -vector space \mathfrak{g} . If $[x, y] = rx + sy$ with $r, s \in k$, we can interchange x, y if necessary to ensure that $r \neq 0$, and then replace x by $(r^{-1})x$ to ensure that $[x, y] = x + sy$. Finally, since $[x + sy, y] = x + sy$, we can replace x by $x + sy$ to ensure that $[x, y] = x$. Notice that in \mathfrak{b} we have the elements

$$U = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix},$$

which have the bracket $[U, V] = U$. Then there is an obvious isomorphism of Lie algebras $\mathfrak{g} \rightarrow \mathfrak{b}$ under which $x \mapsto U$ and $y \mapsto V$.

(b) Take

$$G = \left\{ \begin{bmatrix} z & w \\ 0 & 1 \end{bmatrix} \right\} \leqslant \mathrm{GL}_2(k).$$

3.2. (a) If $\gamma: (a, b) \rightarrow G$ is a differentiable curve then so are the curves defined by

$$(U\gamma)(t) = U\gamma(t), \quad (\gamma U)(t) = \gamma(t)U, \quad (U\gamma U^{-1})(t) = U\gamma(t)U^{-1},$$

and these have derivatives

$$(U\gamma)'(t) = U\gamma'(t), \quad (\gamma U)'(t) = \gamma'(t)U, \quad (U\gamma U^{-1})'(t) = U\gamma'(t)U^{-1}.$$

(b) Use the derivative maps at I , i.e., $\lambda_U = dL_U$, $\rho_U = dR_U$ and $\chi_U = dC_U$. The required properties follow easily.

3.3. Follow the approach of Section 3.3.

The Lie algebra of G_1 is

$$\mathfrak{g}_1 = \left\{ \begin{bmatrix} u & 0 \\ w & v \end{bmatrix} : u, v, w \in \mathbb{R} \right\}.$$

The elements $U = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $V = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ and $W = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ form a basis with the two non-trivial brackets $[U, W] = -W$ and $[V, W] = W$.

The Lie algebra of G_2 is

$$\mathfrak{g}_2 = \left\{ \begin{bmatrix} 0 & t \\ t & 0 \end{bmatrix} : t \in \mathbb{R} \right\},$$

and this is abelian.

The Lie algebra of G_3 is

$$\mathfrak{g}_3 = \left\{ \begin{bmatrix} r & 0 & u \\ v & s & w \\ -u & 0 & t \end{bmatrix} : r, s, t, u, v, w \in \mathbb{R} \right\}.$$

The Lie algebra of G_4 is

$$\mathfrak{g}_4 = \left\{ \begin{bmatrix} A & t \\ 0 & 0 \end{bmatrix} : A \in M_n(k), t \in k^n \right\},$$

with Lie bracket given by

$$\left[\begin{bmatrix} A_1 & t_1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} A_2 & t_2 \\ 0 & 0 \end{bmatrix} \right] = \begin{bmatrix} [A_1, A_2] & A_1 t_2 - A_2 t_1 \\ 0 & 0 \end{bmatrix}.$$

Consider the $2m \times 2m$ real matrix

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & \cdots & \cdots & A_{mm} \end{bmatrix}$$

where each $A_{rs} \in M_2(\mathbb{R})$. Then $A^T J_{2m} + J_{2m} A = O$ if and only if the equations

$$A_{sr}^T J + JA_{rs} = O \quad (r, s = 1, \dots, m)$$

are satisfied and these are equivalent to

$$A_{sr}^T = JA_{rs}J \quad (r, s = 1, \dots, m).$$

If we write $A_{rs} = \begin{bmatrix} a_{rs} & b_{rs} \\ c_{rs} & d_{rs} \end{bmatrix}$ then

$$a_{sr} = -d_{rs}, \quad b_{sr} = b_{rs}, \quad c_{sr} = c_{rs}, \quad d_{sr} = -a_{rs}.$$

If $s = r$, this gives $A_{rr} = \begin{bmatrix} a_{rr} & b_{rr} \\ c_{rr} & -a_{rr} \end{bmatrix}$.

3.4. We have $\mathfrak{b} \cong \mathfrak{a}/\ker \Phi$. Then \mathfrak{b} is abelian if and only if for every commutator $[x, y]$ in \mathfrak{a} , $[x, y] + \ker \Phi = 0 + \ker \Phi$, i.e., $[x, y] \in \ker \Phi$. Hence \mathfrak{b} is abelian if and only if $\mathfrak{a} \subseteq \ker \Phi$.

3.5. If $\exp(sX) \exp(tY) = \exp(tY) \exp(sX)$ for all $s, t \in \mathbb{R}$, then

$$[X, Y] = \frac{d}{ds} \Big|_{s=0} \frac{d}{dt} \Big|_{t=0} (\exp(sX) \exp(tY) - \exp(tY) \exp(sX)) = 0.$$

Conversely, if $[X, Y] = 0$ then for each $s \in \mathbb{R}$, consider the function

$$\gamma: \mathbb{R} \longrightarrow G; \quad \gamma(t) = \exp(sX) \exp(tY) \exp(-sX).$$

Then $\gamma(0) = I$ and

$$\gamma'(t) = \exp(sX)Y \exp(tY) \exp(-sX) = \exp(sX) \exp(tY) \exp(-sX)Y = \gamma(t)Y$$

since $\exp(sX)Y \exp(-sX) = Y$ by definition of the exponential function. But the differential equation

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

has the unique solution $\gamma(t) = \exp(tY)$, so for all s, t we have

$$\exp(sX) \exp(tY) \exp(-sX) = \exp(tY)$$

and therefore

$$\exp(sX) \exp(tY) = \exp(tY) \exp(sX).$$

3.6. (a) If $A \in U$ then

$$\begin{aligned} \Phi(A)^T &= (I - A^T)(I + A^T)^{-1} = (I - A^{-1})(I + A^{-1})^{-1} \\ &= (A - I)(A + I)^{-1} = -\Phi(A), \end{aligned}$$

so $\Phi(A) \in \text{Sk-Sym}_n(\mathbb{R})$. Conversely, if $B \in \text{Sk-Sym}_n(\mathbb{R})$ then

$$\begin{aligned}\Phi((I - B)(I + B)^{-1}) &= (I - (I - B)(I + B)^{-1})(I + (I - B)(I + B)^{-1})^{-1} \\ &= (I - (I - B)(I + B)^{-1})(I + (I - B)(I + B)^{-1})^{-1} \\ &= ((I + B) - (I - B))((I + B) + (I - B))^{-1} \\ &= (2B)(2I)^{-1} = B,\end{aligned}$$

and

$$\begin{aligned}((I - B)(I + B)^{-1})^T(I - B)(I + B)^{-1} &= (I - B^T)(I + B^T)^{-1}(I - B)(I + B)^{-1} \\ &= (I + B)(I - B)^{-1}(I - B)(I + B)^{-1} = I.\end{aligned}$$

Hence $(I - B)(I + B)^{-1} \in O(n)$ and $\Phi((I - B)(I + B)^{-1}) = B$. In fact, there is a path $[0, 1] \rightarrow O(n)$ given by

$$t \mapsto (I - tB)(I + tB)^{-1},$$

hence $\det(I - tB)(I + tB)^{-1} = 1$ since $\det C = \pm 1$ if $C \in O(n)$. So $(I - B)(I + B)^{-1} \in SO(n)$ and therefore $\text{im } \Phi = \text{Sk-Sym}_n(\mathbb{R})$.

(b) By the calculation in (a), $\Phi^{-1}(B) = (I - B)(I + B)^{-1}$.

(c) We have

$$\dim SO(n) = \dim \text{Sk-Sym}_n(\mathbb{R}) = 1 + 2 + \cdots + (n - 1) = \binom{n}{2}.$$

3.7. Parts (a)–(c) are very similar to the previous question and we obtain

$$\Theta^{-1}(B) = (I - B)(I + B)^{-1} \quad (B \in \text{Sk-Herm}_2(\mathbb{C})),$$

and

$$\dim U(n) = \text{Sk-Herm}_n(\mathbb{C}) = n + 2 \binom{n}{2} = n^2.$$

(d) In the case $n = 2$, if $A \in V \cap SU(2)$, then the eigenvalues of A must have the form $\lambda, \bar{\lambda} \in \mathbb{C}$ where $|\lambda| = 1$. Then the eigenvalues of $\Theta(A)$ are $(1 - \lambda)/(1 + \lambda)$ and $(1 - \bar{\lambda})/(1 + \bar{\lambda})$, hence

$$\text{tr } \Theta(A) = \frac{1 - \lambda}{1 + \lambda} + \frac{1 - \bar{\lambda}}{1 + \bar{\lambda}} = \frac{1 - \lambda}{1 + \lambda} + \frac{\lambda - 1}{\lambda + 1} = 0.$$

Conversely, if $B \in \text{Sk-Herm}_2^0(\mathbb{C})$ then B has imaginary eigenvalues $\pm ti$ for some $t \in \mathbb{R}$, so $\Theta^{-1}(B)$ has eigenvalues $(1 - ti)/(1 + ti)$ and $(1 + ti)/(1 - ti)$. Thus we have

$$\det \Theta^{-1}(B) = \frac{(1 - ti)(1 + ti)}{(1 + ti)(1 - ti)} = 1.$$

Examples show that this can be false when $n > 2$.

3.8. (a) The surjection $\det: O(n) \rightarrow \{1, -1\}$ has kernel $SO(n)$. The diagonal matrix $\text{diag}(1, \dots, 1, -1) \in O(n)$ generates a subgroup $C = \{I, \text{diag}(1, \dots, 1, -1)\}$ of order 2 for which

$$O(n) = C SO(n), \quad C \cap SO(n) = \{I\}.$$

(b) Let

$$\mathbf{T}' = \{\text{diag}(1, \dots, 1, z) \in \mathbf{U}(n) : |z| = 1\} \leqslant \mathbf{U}(n).$$

Then

$$\mathbf{U}(n) = \mathbf{T}' \mathbf{S}\mathbf{U}(n), \quad \mathbf{T}' \cap \mathbf{S}\mathbf{U}(n) = \{I\}.$$

(c) Let

$$D = \{\text{diag}(1, \dots, 1, t) \in \mathbf{GL}_n(\mathbb{R}) : t \in \mathbb{R}\} \leqslant \mathbf{GL}_n(\mathbb{R}).$$

Then

$$\mathbf{GL}_n(\mathbb{R}) = D \mathbf{SL}_n(\mathbb{R}), \quad D \cap \mathbf{SL}_n(\mathbb{R}) = \{I\}.$$

(d) Let

$$D' = \{\text{diag}(1, \dots, 1, z) \in \mathbf{GL}_n(\mathbb{C}) : z \in \mathbb{C}\} \leqslant \mathbf{GL}_n(\mathbb{C}).$$

Then

$$\mathbf{GL}_n(\mathbb{C}) = D' \mathbf{SL}_n(\mathbb{C}), \quad D' \cap \mathbf{SL}_n(\mathbb{C}) = \{I\}.$$

Chapter 4

4.1. Some elements of order 2 in A'_3 are -1 and

$$e_3(1\ 2) = \frac{1}{6}[2(1\ 2) - (1\ 3) - (2\ 3)],$$

$$e_3(1\ 3) = \frac{1}{6}[2(1\ 3) - (2\ 3) - (1\ 2)],$$

$$e_3(2\ 3) = \frac{1}{6}[2(2\ 3) - (1\ 2) - (1\ 3)].$$

As \mathbf{H} has only the element -1 of order 2 these algebras cannot be isomorphic.

4.2. (a) The subset $A_1^\times \subseteq A$ is closed and bounded, so compact.

(b) Let $u, v \in A_1^\times$. Then we have the inequalities

$$\nu(uv) \leqslant \nu(u)\nu(v) = 1,$$

$$\nu((uv)^{-1}) = \nu(v^{-1}u^{-1}) \leqslant \nu(v^{-1})\nu(u^{-1}) \leqslant 1,$$

$$1 = \nu(1) = \nu(uv(uv)^{-1}) \leqslant \nu(uv)\nu((uv)^{-1}) \leqslant 1,$$

hence $\nu(uv) = \nu((uv)^{-1}) = 1$. So $A_1^\times \leqslant A^\times$.

(c) C_1^\times is the unit circle with abelian Lie algebra $\{ti : t \in \mathbb{R}\} \leqslant \mathbf{C}$.

$$\mathbf{H}_1^\times = \{u + vj : u, v \in \mathbb{C}, |u|^2 + |v|^2 = 1\},$$

with Lie algebra

$$\{ri + sj + tk : u, v \in \mathbb{C}, r, s, t \in \mathbb{R}\} \cong \mathbf{su}(2).$$

For $\mathbf{k} = \mathbb{R}$ or \mathbb{C} , let $A = M_n(\mathbf{k})^\times$. Suppose that $|Au| < |u|$ for some $u \in \mathbf{k}^n$; then

$$1 = \frac{|u|}{|u|} = \frac{|A^{-1}(Au)|}{|u|} < \frac{|A^{-1}(Au)|}{|Au|} \leqslant \|A^{-1}\|,$$

contradicting the fact that $\|A^{-1}\| = 1$. This shows that for every $x \in \mathbf{k}^n$, $|Ax| = |x|$. By Proposition 1.38 and the analogous result for unitary groups, this shows that $A \in O(n)$ if $\mathbf{k} = \mathbb{R}$ and $A \in U(n)$ if $\mathbf{k} = \mathbb{C}$, while the corresponding Lie algebras are $\mathfrak{o}(n) \leqslant M_n(\mathbb{R})$ and $\mathfrak{u}(n) \leqslant M_n(\mathbb{C})$.

4.3. For the case $r > 0$, let $q \in \mathbb{H}$ satisfy $q^2 = r$. Using notation from the proof of Proposition 4.40, we find that $\mathbf{R}(q)$ is a field, so the number of roots of $x^2 - r$ is at most 2 and these are the real numbers $\pm\sqrt{r}$. The case where $r < 0$ can be dealt with using Proposition 4.42.

4.4. Here is what happens when $n = 1$. For $\mathbf{H}_{\mathbb{R}}$, $\Lambda: \mathbb{H} \rightarrow M_4(\mathbb{R})$ takes the values $\Lambda(1) = I_4$ and

$$\Lambda(i) = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \Lambda(j) = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \Lambda(k) = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

and so

$$\Lambda(a1 + bi + cj + dk) = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}.$$

The reduced determinant $Rdet_{\mathbb{R}}: \mathbb{H}^{\times} \rightarrow \mathbb{R}^{\times}$ takes the value

$$\begin{aligned} Rdet_{\mathbb{R}}(a1 + bi + cj + dk) &= \det \Lambda(a1 + bi + cj + dk) \\ &= a^4 + 2a^2b^2 + 2d^2a^2 + 2c^2a^2 + b^4 + 2d^2b^2 + 2c^2b^2 + c^4 + 2c^2d^2 + d^4. \end{aligned}$$

For $\mathbf{H}_{\mathbb{C}}$, $\Lambda: \mathbb{H} \rightarrow M_2(\mathbb{C})$ has the effect

$$\Lambda(1) = I_2, \quad \Lambda(j) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

so

$$\Lambda(a + bj) = \begin{bmatrix} a & -b \\ \bar{b} & \bar{a} \end{bmatrix}.$$

When $n = 1$, the reduced determinant $Rdet_{\mathbb{C}}: \mathbb{H}^{\times} \rightarrow \mathbb{C}^{\times}$ takes the value

$$Rdet_{\mathbb{C}}(a1 + bj) = \det \Lambda(a1 + bj) = |a|^2 + |b|^2.$$

4.5. (b) We have

$$\exp(r + su) = \exp(r) \exp(su) = e^r (\cos s 1 + \sin s u).$$

4.6. (a) $\mathbf{g} = M_n(\mathbb{H})$. (b) $\mathbf{g} = \{A \in M_n(\mathbb{H}) : A^* = -A\}$.

(c) When $n = 1$, taking $\mathbf{k} = \mathbb{R}$ and using Exercise 4.4, we have $a1 + bi + cj + dk \in G$ with $a, b, c, d \in \mathbb{R}$ if and only if

$$a^4 + 2a^2b^2 + 2d^2a^2 + 2c^2a^2 + b^4 + 2d^2b^2 + 2c^2b^2 + c^4 + 2c^2d^2 + d^4 = 1,$$

so

$$\mathbf{g} = \{ri + sj + tk : r, s, t \in \mathbb{R}\}.$$

For $\mathbf{k} = \mathbb{C}$, $u1 + vj \in G$ with $u, v \in \mathbb{C}$ if and only if $|u|^2 + |v|^2 = 1$, giving

$$\mathbf{g} = \{ri + zj : r \in \mathbb{R}, z \in \mathbb{C}\}.$$

(d) When $n = 1$,

$$\text{Sp}(1) = \{a1 + bi + cj + dk : a, b, c, d \in \mathbb{R}, a^2 + b^2 + c^2 + d^2 = 1\},$$

so again using Exercise 4.4, for $k = \mathbb{R}$ or \mathbb{C} we obtain

$$\mathfrak{g} = \{ri + zj : r \in \mathbb{R}, z \in \mathbb{C}\}.$$

Chapter 5

5.1. (a) First write $u = r_1e_1 + \cdots + r_ne_n$ with $r_k \in \mathbb{R}$ and $r_1^2 + \cdots + r_n^2 = 1$. Then

$$u^2 = r_1^2 e_1^2 + \cdots + r_n^2 e_n^2 = -1.$$

Also, if $w \in \mathbb{R}^n$ with $w \cdot u = 0$, then writing $w = s_1e_1 + \cdots + s_ne_n$ with $s_k \in \mathbb{R}$, we have

$$\begin{aligned} uw &= (r_1s_1e_1^2 + \cdots + r_ns_n^2e_n^2) + \sum_{1 \leq i < j \leq n} (r_is_j - r_js_i)e_ie_j \\ &= -(u \cdot w) + \sum_{1 \leq i < j \leq n} (r_is_j - r_js_i)e_ie_j = \sum_{1 \leq i < j \leq n} (r_is_j - r_js_i)e_ie_j. \end{aligned}$$

A similar calculation shows that

$$uw = \sum_{1 \leq i < j \leq n} (r_js_i - r_is_j)e_ie_j = -wu.$$

So we have $wu = uw$ whenever $w \cdot u = 0$. Using this we obtain

$$uvu = uv_1u + uv_2u = v_1 - uv_2u.$$

(c) The elements Ae_1, \dots, Ae_n form an orthonormal basis of $\mathbb{R}^n \subseteq \text{Cl}_n$, so using the universal property of Theorem 5.4, the function $\mathbb{R}^n \rightarrow \text{Cl}_n$ sending $\sum_{i=1}^n x_i e_i$ to $\sum_{i=1}^n x_i Ae_i$ induces a homomorphism $A_* : \text{Cl}_n \rightarrow \text{Cl}_n$. The inverse $A^{-1} = A^T$ induces $(A^{-1})_* : \text{Cl}_n \rightarrow \text{Cl}_n$ which is actually the inverse of A_* . A straightforward calculation shows that $A_*(e_1 \cdots e_n) = (\det A)e_1 \cdots e_n$.

5.2. (a) The image of $i'_n : \text{Cl}_n \rightarrow \text{Cl}_{n+1}$ is spanned by the monomials $e_{i_1} \cdots e_{i_r}$ with $i_k = 1, \dots, n$.

(b) Since $x \cdot e_{n+1} = 0$,

$$(xe_{n+1})^2 = xe_{n+1}xe_{n+1} = -x^2 e_{n+1}^2 = -(-|x|^2)(-1) = -|x|^2,$$

so the universal property gives the desired homomorphism. Also notice that $xe_{n+1} \in \text{Cl}_n^+$ and it is easy to see that every monomial $e_{i_1} \cdots e_{i_{2s}}$ of even length with $i_k = 1, \dots, n+1$ is in the image.

5.3. (a) For any $k = 1, \dots, n$,

$$e_k w_n = e_k e_1 \cdots e_n = (-1)^{k-1} e_1 \cdots e_{k-1} e_k^2 e_{k+1} \cdots e_n (-1)^k e_1 \cdots e_{k-1} e_{k+1} \cdots e_n,$$

while

$$w_n e_k = e_1 \cdots e_n e_k = (-1)^{n-k} e_1 \cdots e_{k-1} e_k^2 e_{k+1} \cdots e_n (-1)^{n-k+1} e_1 \cdots e_{k-1} e_{k+1} \cdots e_n.$$

Since $n-k+1 \equiv k \pmod{2}$, $e_k w_n = w_n e_k$. From this it follows that for every monomial in the e_k ,

$$e_{i_1} \cdots e_{i_r} w_n = w_n e_{i_1} \cdots e_{i_r}.$$

(b) We have

$$\omega_n^2 = e_1 \cdots e_n e_1 \cdots e_n = (-1)^{(n-1)+\cdots+1} e_1^2 \cdots e_n^2 = (-1)^{\binom{n}{2} + n} = (-1)^{\binom{n+1}{2}}.$$

But $(n+1) \equiv 0 \pmod{4}$, so $\binom{n+1}{2} \equiv 0 \pmod{2}$ and therefore $\omega_n^2 = 1$. Now

$$\Omega_{\pm}^2 = \frac{1}{4}(1 \pm 2\omega_n + \omega_n^2) = \frac{1}{4}(2 \pm 2\omega_n) = \Omega_{\pm},$$

so these are idempotents which are central and satisfy $\Omega_+ + \Omega_- = 1$.

(c) Set $\Omega_+ A_+ = Cl_n$ and $\Omega_- A_- = Cl_n$.

(d) By Exercise 5.2, $Cl_3^+ = Cl_2 = H$ and $Cl_7^+ = Cl_6 = M_8(\mathbb{R})$. In each case multiplication by Ω_{\pm} maps the simple algebra Cl_n^+ onto A_{\pm} and therefore gives an isomorphism $Cl_n^+ \cong A_{\pm}$.

5.4. (c) Send γ to -1 and e_j to e_j .

(d) The elements $(1/2)(1 \pm \gamma) \in \mathbb{R}[ClGp_n]$ are central idempotents. Put

$$A = (1/2)(1 + \gamma)\mathbb{R}[ClGp_n], \quad B = (1/2)(1 - \gamma)\mathbb{R}[ClGp_n].$$

5.6. (b) $Z(O(n)) = Z(SO(n))$.

Chapter 6

6.1. (b) For some $Q \in \text{Lor}(3, 1)$,

$$A_t = Q \begin{bmatrix} \cos t & -\sin t & 0 & 0 \\ \sin t & \cos t & 0 & 0 \\ 0 & 0 & \cosh(-t) & \sinh(-t) \\ 0 & 0 & \sinh(-t) & \cosh(-t) \end{bmatrix} Q^{-1}.$$

(c) With the same Q as in (b), take

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

(d) A suitable equation is

$$\alpha'(t) = \alpha(t)U, \quad \alpha(0) = 1.$$

6.2. (a) We have

$$P_{24} + P_{34} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

and

$$\alpha(t) = \exp(t(P_{24} + P_{34})) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\cosh \sqrt{2}t + 2}{2} & \frac{\cosh \sqrt{2}t - 2}{2} & \frac{\sinh \sqrt{2}t}{\sqrt{2}} \\ 0 & \frac{\cosh \sqrt{2}t - 2}{2} & \frac{\cosh \sqrt{2}t + 2}{2} & \frac{\sinh \sqrt{2}t}{\sqrt{2}} \\ 0 & \frac{\sinh \sqrt{2}t}{\sqrt{2}} & \frac{\sinh \sqrt{2}t}{\sqrt{2}} & \frac{\cosh \sqrt{2}t}{2} \end{bmatrix}.$$

(b) We have

$$-i(E + F) = \begin{bmatrix} 0 & \frac{1}{2}(1-i) \\ \frac{1}{2}(1+i) & 0 \end{bmatrix}$$

and

$$\tilde{\alpha}(t) = \exp(-ti(E + F)) = \begin{bmatrix} \cosh(t/\sqrt{2}) & \frac{\sinh(t/\sqrt{2})(1-i)}{\sqrt{2}} \\ \frac{\sinh(t/\sqrt{2})(1+i)}{\sqrt{2}} & \cosh(t/\sqrt{2}) \end{bmatrix}.$$

(c) The homomorphism $\widetilde{\text{Ad}}: \text{SL}_2(\mathbb{C}) \rightarrow \text{Lor}(3,1)$ relates these two curves through the equation $\alpha = \widetilde{\text{Ad}} \circ \tilde{\alpha}$.

Chapter 7

7.1. (a) Define the function

$$F: M_n(\mathbb{R}) \times \mathbb{R} \rightarrow \mathbb{R}; \quad F(A, b) = b \det A - 1.$$

This is smooth and $M = F^{-1}0 \subseteq M_n(\mathbb{R}) \times \mathbb{R}$ is a closed subset. The derivative mapping $d = d_{(A,b)}$ at a point $(A, b) \in F^{-1}0$ has the form

$$d(X, v) = (\det A)v + d \det_A(X),$$

where we make the natural identifications $T_{(A,b)} M_n(\mathbb{R}) \times \mathbb{R} = M_n(\mathbb{R}) \times \mathbb{R}$ and $T_t \mathbb{R} = \mathbb{R}$ for any $t \in \mathbb{R}$. Since $\det A \neq 0$, d is surjective at such a point. Hence $M \subseteq M_n(\mathbb{R}) \times \mathbb{R}$ is a smooth submanifold of dimension equal to $\dim M_n(\mathbb{R}) = n^2$. We also have

$$T_{(A,b)} M = \{(X, (\det A)^{-1} d \det_A(X)) : X \in M_n(\mathbb{R})\}.$$

(b) We have $M \cong \text{GL}_n(\mathbb{R})$.

7.4. In each case we have to show that the subgroup is closed in G . In (a-c) we also use the continuous action (as introduced in Section 1.9, see the Exercises of Chapter 1 for details),

$$\mu: G \times G \rightarrow G; \quad \mu(x, y) = xyx^{-1}.$$

- (a) $Z_G(g)$ agrees with the stabiliser of g , $\text{Stab}_G(g)$ which is a closed subgroup of G .
- (b) $Z(G)$ is the intersection of all the closed sets $Z_G(g)$ for $g \in G$ and so is closed.
- (c) $N_G(g)$ is identical with the generalised stabiliser $\text{Stab}_G(H) \leqslant G$, which is a closed subgroup.

(d) Let $x \in G - \ker \varphi$. Then $\varphi(x) \neq 1$, so there is an open subset $U \subseteq H$ containing $\varphi(x)$ but not 1. Then $\varphi^{-1}U \subseteq G$ is open since φ is continuous. Clearly $\ker \varphi \cap \varphi^{-1}U = \emptyset$ and $x \in \varphi^{-1}U$. This shows that $G - \ker \varphi$ is open in G , hence $\ker \varphi$ is closed.

7.5. This is similar to the previous exercise.

7.6. Use the smooth maps L_g , R_g , and χ_g of Proposition 7.16. The tangent spaces are the images of the derivatives at g applied to the tangent space of H .

7.7. By Exercise 7.4(d), $\ker \varphi \leqslant G$ is closed. For $g \in \ker \varphi$ we have $T_g \ker \varphi = \ker d\varphi_g$.

7.8. (a) Use the continuous function

$$F: \text{GL}_{2n}(\mathbb{R}) \rightarrow M_n(\mathbb{R}); \quad F(X) = X^2 + I_{2n}$$

for which $C_{2n} = F^{-1}O$.

(b) The eigenvalues of an element $J \in C_{2n}$ are $\pm i$. Since the minimal polynomial of J over \mathbb{C} is clearly $X^2 + 1$, which has no repeated roots, J is diagonalisable over \mathbb{C} , and over \mathbb{R} has the form

$$J = B \begin{bmatrix} J_2 & O & \cdots & \cdots & O \\ O & J_2 & O & \cdots & \vdots \\ \ddots & \ddots & \ddots & \ddots & \vdots \\ \cdots & \cdots & \cdots & J_2 & O \\ O & \cdots & \cdots & O & J_2 \end{bmatrix} B^{-1}, \quad \text{where } B \in \mathrm{GL}_{2n}(\mathbb{R}) \text{ and } J_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

(c) These stabilisers are the subgroups $\rho_n \mathrm{GL}_n(\mathbb{C})$ and $\rho'_n \mathrm{GL}_n(\mathbb{C})$.

(d) For techniques useful for showing that C_{2n} is a submanifold, see Chapter 8. The tangent space to C_{2n} at J can be identified with

$$\ker dF_J = \{X \in M_{2n}(\mathbb{R}) : XJ + JX = O\}.$$

Chapter 8

8.2. (a) The obvious quotient homomorphism $\mathrm{Sp}(n) \times \mathrm{Sp}(1) \rightarrow \mathrm{Sp}(n) \tilde{\times} \mathrm{Sp}(1)$ has the discrete subgroup $\{(I_n, 1), (-I_n, -1)\}$ as its kernel, so the derivative homomorphism is an isomorphism of Lie algebras; the Lie algebra of the domain is the direct product $\mathfrak{sp}(n) \times \mathfrak{sp}(1)$.

(b) We can identify HP^{n-1} with $(\mathrm{Sp}(n) \tilde{\times} \mathrm{Sp}(1)) / (\mathrm{Sp}(n-1) \tilde{\times} \mathrm{Sp}(1))$.

Chapter 9

9.1. (a) For each connected component U of G , choose an element $u \in U$. G has a countable open covering $\{U_1, U_2, \dots\}$, hence for each u as above there is an open set $U_{j_u} \subseteq U$ containing u . This means that the components form a countable set.

(b) The components form an open covering of G . By the Heine-Borel Theorem 1.20, a finite subcollection also covers G . As these sets are disjoint there must be only finitely many of them.

9.2. (a) We have

$$\begin{aligned} \mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n) &= \left\{ \begin{bmatrix} A & O_{n-1,1} \\ B & 1 \end{bmatrix} : A \in \mathrm{SL}_{n-1}(\mathbb{R}), B \in M_{1,n-1}(\mathbb{R}) \right\}, \\ \mathrm{Orb}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n) &= \mathbb{R}^n - \{0\}. \end{aligned}$$

(b) This homogeneous space can be identified with $\mathrm{Orb}_{\mathrm{SL}_n(\mathbb{R})}(\mathbf{e}_n) = \mathbb{R}^n - \{0\}$ and this is path connected since every pair of vectors $u, v \in \mathbb{R}^n - \{0\}$ which are not parallel can be joined by a path of form

$$t \mapsto tu + (1-t)v \quad (t \in [0, 1]).$$

Now use induction on n as in Example 9.14 to show that $\mathrm{SL}_n(\mathbb{R})$ is path connected.

9.3. (a) We have

$$\begin{aligned} \mathrm{Stab}_{\mathrm{SL}_n(\mathbb{C})}(\mathbf{e}_n) &= \left\{ \begin{bmatrix} A & O_{n-1,1} \\ B & 1 \end{bmatrix} : A \in \mathrm{SL}_{n-1}(\mathbb{C}), B \in M_{1,n-1}(\mathbb{C}) \right\}, \\ \mathrm{Orb}_{\mathrm{SL}_n(\mathbb{C})}(\mathbf{e}_n) &= \mathbb{C}^n - \{0\}. \end{aligned}$$

(b) This is similar to part (b) of the previous exercise.

(c) Since $\mathrm{GL}_n(\mathbb{C})/\mathrm{SL}_n(\mathbb{C})$ is identified with \mathbb{C}^\times using the determinant, Proposition 9.11 shows that $\mathrm{GL}_n(\mathbb{C})$ is path connected.

9.4. (a) If $x \in \mathbb{R}^n$, then $x^T A A^T x = |A^T x|^2 \geq 0$, and this vanishes only when $A^T x = 0$, which happens only when $x = 0$ since A^T is invertible. The eigenvalues of S are then positive real numbers and S is diagonalisable, say $S = P \mathrm{diag}(\lambda_1, \dots, \lambda_n) P^T$ for some $P \in \mathrm{O}(n)$ and $\lambda_i > 0$. Then we take $S_1 = P \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) P^T$.

(b) We have

$$\begin{aligned}(S_1^{-1} A)^T (S_1^{-1} A) &= A^T (S_1^T)^{-1} S_1^{-1} A = A^T S^{-1} A \\ &= A^T (AA^T)^{-1} A = A^T (A^T)^{-1} A^{-1} A = I.\end{aligned}$$

(c) Since $RR^T = I = SS^T$, we have

$$Q^2 = QSS^T Q^T = QS(QS)^T = PR(PR)^T = PRR^T P^T = P^2.$$

(d) S_2 must have the same eigenvalues as $\mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$ and indeed

$$S_2 = U \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) U^T$$

for some $U \in \mathrm{O}(n)$. So

$$U \mathrm{diag}(\lambda_1, \dots, \lambda_n) U^T = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$$

and hence

$$U \mathrm{diag}(\lambda_1, \dots, \lambda_n) = \mathrm{diag}(\lambda_1, \dots, \lambda_n) U.$$

By comparing entries we find that this is only possible if either $\lambda_i = \lambda_j$ for all i, j or U is diagonal, in which case $S_2 = \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$.

(e) Use (c) and (d).

(f) First show that the set of positive definite symmetric matrices is a path connected subset of $\mathrm{GL}_n(\mathbb{R})$ by finding a path from every such matrix of the form $V \mathrm{diag}(\lambda_1, \dots, \lambda_n) V^T$ with $V \in \mathrm{O}(n)$ to I .

9.5. (a) We have

$$\mathrm{Stab}_{\mathrm{Aff}_n(\mathbb{k})}(0) = \left\{ \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} : A \in \mathrm{GL}_n(\mathbb{k}) \right\}, \quad \mathrm{Orb}_{\mathrm{Aff}_n(\mathbb{k})}(0) = \mathbb{k}^n.$$

Notice that $\mathrm{Stab}_{\mathrm{Aff}_n(\mathbb{k})}(0)$ can be identified with $\mathrm{GL}_n(\mathbb{k})$.

(b) This follows from the corresponding facts for $\mathrm{GL}_n(\mathbb{k})$.

9.6. See Section 8.5 for background on this. The spaces in (a) and (b) are connected and actually diffeomorphic.

9.7. This is similar to the previous exercise; the homogeneous spaces are diffeomorphic and connected.

9.8. (b) The Pfaffian function $\mathrm{pf}: \Sigma_{2m} \rightarrow \mathbb{R}^\times$ is surjective, so Σ_{2m} has at least two path components. In fact it has exactly two components since the diffeomorphic homogeneous space $\mathrm{GL}_{2m}(\mathbb{R})/\mathrm{Symp}_{2m}(\mathbb{R})$ has at most two components; this implies that $\mathrm{Symp}_{2m}(\mathbb{R}) \leq \mathrm{SL}_{2m}(\mathbb{R})$.

9.10. By straightforward calculation we find that

$$N_{U(2)}(T) = T \cup \left\{ \begin{bmatrix} 0 & u \\ v & 0 \end{bmatrix} : |u| = |v| = 1 \right\} \leq U(2).$$

This has two path components,

$$T, \quad \left\{ \begin{bmatrix} 0 & u \\ v & 0 \end{bmatrix} : |u| = |v| = 1 \right\},$$

and $\pi_0 N_{U(2)}(T) \cong \{1, -1\}$.

9.11. Using the previous exercise, we obtain

$$N_G(T') = \left\{ \begin{bmatrix} P & O_{2,2} \\ O_{2,2} & Q \end{bmatrix} : P, Q \in N_{U(2)}(T), \det P = \det Q \right\},$$

which has 4 path components each of which contains one of the matrices

$$I_4, \quad \begin{bmatrix} J_2 & O_{2,2} \\ O_{2,2} & I_2 \end{bmatrix}, \quad \begin{bmatrix} I_2 & O_{2,2} \\ O_{2,2} & J_2 \end{bmatrix}, \quad \begin{bmatrix} J_2 & O_{2,2} \\ O_{2,2} & J_2 \end{bmatrix},$$

where $J_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Then $\pi_0 N_G(T') \cong \{1, -1\} \times \{1, -1\}$.

Chapter 10

10.1. Given a Lie homomorphism $\varphi: T \rightarrow T$, the curve $\gamma: \mathbb{R} \rightarrow T$ given by $\gamma(t) = \varphi(e^{2\pi i t})$ is a one-parameter subgroup satisfying $\gamma'(s+t) = \gamma(s)\gamma'(t)$ for all $s, t \in \mathbb{R}$. By our general results on such curves we have $\gamma(t) = e^{2\pi c i t}$ for some $c \in \mathbb{R}$. But whenever $n \in \mathbb{Z}$, $\gamma(n) = \varphi(1) = 1$, so $c \in \mathbb{Z}$. For φ to be an isomorphism we must have $c = \pm 1$.

For $r \geq 1$, for each integer matrix $A = [a_{ij}] \in M_r(\mathbb{Z})$, there is a Lie homomorphism

$$\varphi_A: T^r \rightarrow T^r; \quad \varphi_A(z_1, \dots, z_r) = (z_1^{a_{11}} \cdots z_1^{a_{1r}}, \dots, z_r^{a_{11}} \cdots z_r^{a_{rr}})$$

with derivative $d = A$ acting in the obvious way on $t = \mathbb{R} \times \cdots \times \mathbb{R}$. Isomorphisms come from those A with integer inverse $A^{-1} \in M_r(\mathbb{Z})$.

10.2. If $T_0 = T$ is not maximal, then it is contained in another T_1 of greater dimension. Repeating this, for dimensional reasons eventually we obtain a sequence of tori $T_0 \leq T_1 \leq \cdots \leq T_k \leq G$ where there is no torus in G which properly contains T_k .

10.3. (a) We have

$$N_{U(2)}(T) = T \cup \left\{ \begin{bmatrix} 0 & u \\ v & 0 \end{bmatrix} : |u| = |v| = 1 \right\} = T \cup \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} T, \quad \pi_0 N_{U(2)}(T) = \{1, -1\}.$$

The conjugation action is given by

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} T \cdot \text{diag}(z, w) = \text{diag}(w, z).$$

(b) and (c) are discussed in Chapters 11 and 12.

10.4. See Chapters 11 and 12.

10.5. (c) $O(n)$ is not connected.

Chapter 11

11.1. (a) By Theorem 11.5, $z(u(n))$ consists of all scalar hermitian matrices tI_n with $t \in \mathbb{R}$. Using the defining Equation (11.1), we have $X \in z(u(n))^\perp$ if and only if $\text{tr}(tiX) = 0$ for all $t \in \mathbb{R}$. Since $\text{tr}(tiX) = ti\text{tr}(X)$, this gives

$$z(u(n))^\perp = \{X \in u(n) : \text{tr } X = 0\} = su(n).$$

(b) Compose the determinant $\det: U(n) \rightarrow \mathbb{T} \cong Z(U(n))$ with the projection $Z(U(n)) \rightarrow Z(U(n))/Z$ for $Z = \{zI_N : z^n = 1\}$.

(c) Define this homomorphism by $(A, B) \mapsto AB$.

11.4. See Chapter 12 for details.

11.5. (a) Take

$$T_1 = \{\text{diag}(u, v, w, \bar{u}\bar{v}\bar{w}) : |u| = |v| = |w| = 1\}.$$

Then

$$N_{G_1}(T_1) = \left\{ \begin{bmatrix} A & O_{2,2} \\ O_{2,2} & B \end{bmatrix} : A, B \in N_{U(2)}(T), \det A \det B = 1 \right\},$$

where $N_{U(2)}(T) \leqslant U(2)$ is the normaliser of the maximal torus used in Exercise 10.3. $N_{G_1}(T_1)$ has 4 components and the Weyl group $W_{G_1}(T_1) = N_{G_1}(T_1)/T_1$ consists of the cosets of the matrices

$$I_4, \quad W_1 = \begin{bmatrix} I_2 & O_{2,2} \\ O_{2,2} & J_2 \end{bmatrix}, \quad W_2 = \begin{bmatrix} J_2 & O_{2,2} \\ O_{2,2} & I_2 \end{bmatrix}, \quad W_3 = \begin{bmatrix} J_2 & O_{2,2} \\ O_{2,2} & J_2 \end{bmatrix},$$

where $J_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$; then $W_{G_1}(T_1) \cong \{1, -1\} \times \{1, -1\}$. $W_{G_1}(T_1)$ acts on T_1 by

$$W_1 T_1 \cdot \text{diag}(u, v, w, \bar{u}\bar{v}\bar{w}) = \text{diag}(u, v, \bar{u}\bar{v}\bar{w}, w),$$

$$W_2 T_1 \cdot \text{diag}(u, v, w, \bar{u}\bar{v}\bar{w}) = \text{diag}(v, u, w, \bar{u}\bar{v}\bar{w}),$$

$$W_3 T_1 \cdot \text{diag}(u, v, w, \bar{u}\bar{v}\bar{w}) = \text{diag}(v, u, \bar{u}\bar{v}\bar{w}, w).$$

The Lie algebra of T_1 is

$$\mathfrak{t}_1 = \{\text{diag}(ri, si, ti, -(r+s+t)i) : r, s, t \in \mathbb{R}\},$$

and there are 4 non-trivial roots $\pm\alpha_1, \pm\alpha_2$, where

$$\alpha_1(ri, si, ti, -(r+s+t)i) = r - s, \quad \alpha_2(ri, si, ti, -(r+s+t)i) = r + s + 2t,$$

for $r, s, t \in \mathbb{R}$. $W_{G_1}(T_1)$ acts on the roots by

$$\begin{aligned} W_1 T_1 \cdot \alpha_1 &= \alpha_1, & W_2 T_1 \cdot \alpha_1 &= -\alpha_1, & W_3 T_1 \cdot \alpha_1 &= -\alpha_1, \\ W_1 T_1 \cdot \alpha_2 &= -\alpha_2, & W_2 T_1 \cdot \alpha_2 &= \alpha_2, & W_3 T_1 \cdot \alpha_2 &= -\alpha_2. \end{aligned}$$

(b) This is similar to (a), see also Exercise 9.11.

Chapter 12

12.2. Rotate and rescale Figure 12.3 into Figure 12.4.

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] M.F. Atiyah, R. Bott & A. Shapiro, Clifford modules, *Topology* **3** (1964) suppl. 1, 3–38.
- [4] R. Carter, G. Segal & I.G. Macdonald, *Lectures on Lie groups and Lie algebras*, Cambridge University Press (1995).
- [5] P.M. Cohn, *Classic Algebra*, Wiley (2000).
- [6] L. Conlon, *Differentiable Manifolds: A First Course*, Birkhäuser (1993).
- [7] M.L. Curtis, *Matrix Groups*, Springer-Verlag (1984).
- [8] R.W.R. Darling, *Differential Forms and Connections*, Cambridge University Press (1994).
- [9] A. Dold, *Lectures on Algebraic Topology*, Springer-Verlag (1995).
- [10] W. Ebeling, *Lattices and Codes*, Vieweg (1994).
- [11] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, American Mathematical Society (2001).
- [12] R. Howe, Very basic Lie theory, *American Mathematical Monthly* **90** (1983), 600–623; Correction, *Amer. Math. Monthly* **91** (1984), 247.
- [13] W.Y. Hsiang, *Lectures on Lie Groups*, World Scientific (2000).
- [14] V.G. Kač, *Infinite-dimensional Lie Algebras*, Cambridge University Press (1990).

- [15] S. Lang, *Algebra*, second edition, Addison-Wesley (1984).
- [16] S. Lang, *Linear Algebra*, reprint of the third edition, Springer-Verlag, (1989).
- [17] S. Lang, *Differential and Riemannian Manifolds*, third edition, Springer-Verlag (1995).
- [18] S. Lang, *Fundamentals of differential geometry*, Springer-Verlag (1999).
- [19] H.B. Lawson & M.-L. Michelsohn, *Spin Geometry*, Princeton University Press, (1989).
- [20] C.R.F. Maunder, *Algebraic Topology*, Dover Publications (1996).
- [21] M.A. Naïmark, *Normed Rings*, translated by L.F. Boron, Wolters-Noordhoff Publishing (1970).
- [22] G.K. Pedersen, *Analysis Now*, Springer-Verlag (1989).
- [23] I.R. Porteous, *Topological Geometry*, Van Nostrand Reinhold (1969).
- [24] I.R. Porteous, *Clifford Algebras and the Classical Groups*, Cambridge University Press (1995).
- [25] J.-P. Serre, *Complex Semisimple Lie Algebras*, Springer-Verlag (1987).
- [26] H. Sato, *Algebraic Topology: An Intuitive Approach*, Amer. Math. Soc. Trans. of Mathematical Monographs 183 (1996).
- [27] S. Sternberg, *Group Theory and Physics*, Cambridge University Press (1994).
- [28] G. Strang, *Linear Algebra and its Applications*, Second edition, Academic Press [Harcourt Brace Jovanovich] (1980).
- [29] F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Springer-Verlag (1983).

- LPU*-decomposition, 226
- k**-Lie algebra, 67
- k**-line in k^{n+1} , 217
- k**-matrix group, 15
- k**-norm, 7, 33, 111
- k**-projective space, 217
- \pm -grading, 136
- n*-manifold, 182

- abelian Lie algebra, 69
- action
 - adjoint, 49, 75, 86, 188
 - continuous, 37
 - group, 37
 - linear, 38
 - smooth, 208
- adjoint
 - action, 49, 75, 86, 188
 - representation, 75
- affine
 - geometry, 19
 - group, 18
 - transformation, 19
- algebra, 99
 - automorphism, 101
 - automorphisms
 - group of, 122
 - Clifford, 130
 - commutative, 100
 - division, 100
 - finite dimensional, 4, 99
 - group, 100
 - homomorphism, 101
 - isomorphism, 101
 - Lie, 59, 67
- Lorentz Clifford, 177
- norm, 111
- normed, 111
- real Clifford, 130
- semi-simple, 106
- simple, 104
- alternating group, 153
- anti-homomorphism of skew rings, 120
- atlas, 182
- automorphism
 - canonical, 134
 - inner, 124
 - of algebras, 101
 - outer, 124

- basis
 - dual, 290
- binary expansion, 131
- Borel subgroup, 18
- bracket
 - Lie, 68
- Bruhat decomposition, 226

- canonical
 - automorphism, 134
 - commutation relation, 205
- Cartan
 - decomposition, 280
 - matrix, 297
- Cauchy sequence, 10
- central
 - idempotent, 104
 - subfield, 102
- centre, 69, 151
 - of a Lie algebra, 69

- trivial, 70
- centreless, 70
- character
- irreducible
- complex, 108
- characteristic polynomial, 51
- chart, 182
- circle group, 251
- Clifford
 - algebra, 130
 - Lorentz, 177
 - real, 130
 - universal property of, 132
- group, 139
- finite, 155
- open, 17
- closed subset, 8
- closure, 41
- commutative algebra, 100
- commutator, 68
- subalgebra, 69
- commute, 68
- compact, 12, 13
- topological space, 13
- complement
- orthogonal, 274
- complete, 10
- norm, 10
- complex
 - dimension, 72
 - matrix group, 29
 - structure, 209
- complexification, 92
- component
- path, 241
- conjugate, 258
- hermitian, 27
- quaternionic, 120
- conjugation
- map, 187
- quaternionic, 120
- connected, 235
 - by a path, 241
 - Dynkin diagram, 298
 - path, 235
- continuous
 - group action, 37
 - homomorphism, 31
 - isomorphism
 - of matrix groups, 32
 - map, 7
- coordinate function, 8
- coroot, 300
- simple, 300
- covering
 - finite, 274
- cross product, 68
- curve
 - differentiable, 55, 57, 183
 - smooth, 183
- derivation property, 77
- derivative, 74
- derived subalgebra, 69
- determinant
 - reduced, 116
- diagonalisable, 51
- diffeomorphism, 181
- differentiable, 183
 - curve, 55, 57, 183
 - homomorphism, 74
 - map, 73
- Differential Geometry, 181
- dimension, 72
 - complex, 72
- direct
- isometry, 22
- product, 19, 71, 101
- disc
 - open, 7
- disjoint union, 21
- division algebra, 100
- dual
 - basis, 290
 - linear, 289
 - root
 - system, 301
- Dynkin diagram, 298
 - connected, 298
 - irreducible, 298
- elementary 2-group, 154
- Euler characteristic, 256
- exponential function, 47
- extraspecial 2-group, 154
- field
 - skew, 100
- finite
 - Clifford group, 155
 - covering, 274
 - dimensional algebra, 4, 99
- flow line, 63
- form
 - skew symmetric, 229
 - symplectic, 26, 229
- function

- coordinate, 8
- exponential, 47
- logarithm, 48
- fundamental root, 292
- future pointing light cone, 158

- Gaussian elimination, 226
- general linear group, 5, 35
- generalised permutation matrix, 132
- generated by a subset, 144
- generator, 254
 - topological, 254
- geometry
 - affine, 19
 - symplectic, 26
- Grassmannian, 222
- group
 - k-matrix, 15
 - action, 37
 - continuous, 37
 - smooth, 208
 - affine, 18
 - algebra, 100
 - alternating, 153
 - circle, 251
 - Clifford, 139
 - finite Clifford, 155
 - general linear, 5, 35
 - Heisenberg, 203
 - inner automorphism, 124
 - isometry, 22
 - Lie, 187
 - Lorentz, 26
 - matrix, 15
 - of algebra automorphisms, 122
 - of path components, 242
 - one-parameter, 57
 - orthogonal, 20
 - outer automorphism, 124
 - pinor, 143
 - projective
 - linear, 125, 219
 - quasi-symplectic, 233
 - quaternionic symplectic, 121
 - special linear, 5
 - special orthogonal, 21
 - special unitary, 28
 - spinor, 144
 - symmetric, 153
 - symplectic, 27
 - topological, 12
 - unimodular, 5
 - unitary, 27

- Weyl, 260, 292

- Hamiltonian, 118
- Hausdorff space, 14
- Heisenberg
 - algebra, 205
 - group, 203
 - Lie algebra, 205
- hermitian
 - conjugate, 27
 - skew, 83
- homeomorphism, 14
- homogeneous
 - polynomial, 38
 - space, 214
- homomorphism
 - algebra, 101
 - continuous, 31
 - differentiable, 74
 - Lie, 74
 - of algebras, 101
 - of Lie algebras, 70
 - property, 57
- hyperplane, 23
 - reflection, 23, 291

- ideal
 - Lie, 69
- idempotent, 104
 - central, 104
 - indecomposable, 105
 - orthogonal, 104
- indecomposable idempotent, 105
- indirect isometry, 22
- inner
 - automorphism, 124
 - group, 124
 - product, 227
 - invariant, 269
 - Lorentz, 157
 - non-degenerate, 227
 - positive definite, 228
 - real, 86
 - products, related, 227
- invariant, 87
 - inner product, 269
- irreducible
 - complex character, 108
 - Dynkin diagram, 298
- isometry, 21, 158
 - direct, 22
 - group, 22
 - indirect, 22
 - linear, 21

- isomorphism
 - continuous
 - of matrix groups, 32
 - of algebras, 101
 - of Lie algebras, 70, 86

- Jacobi identity, 68
- Jacobian matrix, 184
- Jordan
 - block matrix, 52
 - form, 51, 53, 79

- Kronecker symbol, 20

- Lefschetz Fixed Point Theorem, 256
- left
 - multiplication map, 187
 - regular representation, 103
- Lie
 - algebra, 59, 67
 - abelian, 69
 - Heisenberg, 205
 - homomorphism, 70
 - isomorphism, 70, 86
 - quotient, 70
 - simple, 70
 - bracket, 67, 68
 - group, 187
 - semi-simple, 276
 - simple, 274
 - homomorphism, 74
 - ideal, 69
 - proper, 69
 - subalgebra, 69
 - subgroup, 187
- light cone, 158
 - future pointing, 158
 - negative, 158
 - past pointing, 158
 - positive, 158
- linear
 - action, 38
 - dual, 289
 - isometry, 21
- local trivialisation, 214
- locally path connected, 235
- logarithm function, 48
- Lorentz
 - Clifford algebra, 177
 - group, 26
 - inner product, 157
- lower triangular
 - matrix, 171

- manifold, 182
 - smooth, 181
- map
 - conjugation, 187
 - differentiable, 73
 - left multiplication, 187
 - right multiplication, 187
 - smooth, 181, 182
- matrix
 - Cartan, 297
 - diagonalisable, 51
 - generalised permutation, 132
 - group, 15
 - complex, 29
 - Jacobian, 184
 - Jordan block, 52
 - lower triangular, 171
 - orthogonal, 20
 - permutation, 153
 - subgroup, 15, 16
 - unimodular, 4
 - unipotent, 18
 - upper triangular, 18, 171
- maximal, 255
- torus, 251
- minimal polynomial, 52

- negative
 - light cone, 158
 - root, 293
- non-degenerate inner product, 227
- non-trivial root, 279
- norm, 5–7, 33, 111
 - complete, 10
 - operator, 6, 34
 - supremum, 6
- normaliser, 260
- normed algebra, 111
- null
 - root space, 279
 - vector, 158

- one-parameter
 - group, 57
 - semigroup, 57
 - subgroup, 57
- open disc, 7
- operator norm, 6, 34
- orbit, 37
- orthogonal, 90
 - complement, 274
 - group, 20
 - special, 21

- idempotent, 104
- matrix, 20
- outer
 - automorphism, 124
 - group, 124
- parabolic subgroup, 221
- past pointing light cone, 158
- path
 - component, 241
 - connected, 235
 - locally, 235
- permutation matrix, 153
- generalised, 132
- Pfaffian, 156, 232
- phase portrait, 63
- pinor group, 143
- polynomial
 - homogeneous, 38
 - minimal, 52
- positive
 - definite, 87
 - inner product, 228
 - light cone, 158
 - root, 293
- principal bundle, 214
- product
 - cross, 68
 - direct, 19, 71, 101
 - inner, 227
 - semi-direct, 19, 71
 - topology, 11
 - vector, 68
- projective
 - linear group, 125, 219
 - quaternionic symplectic group, 219
 - space, 217
 - special unitary group, 219
 - unitary group, 219
- proper Lie ideal, 69
- pure quaternion, 119
- quadratic form, 227
- quasi-symplectic group, 233
- quaternion
 - pure, 119
 - real, 119
- quaternionic
 - conjugate, 120
 - conjugation, 120
 - symplectic group, 121
 - projective, 219
- quaternions
 - skew field of, 118
- quotient
 - Lie algebra, 70
 - topology, 33, 212
 - universal property of, 213
- real
 - Clifford algebra, 130
 - inner product, 86
 - part, 138
 - quaternion, 119
 - symmetric bilinear form, 87
 - symplectic group, 27
- reduced
 - determinant, 116
 - echelon form, 226
 - root system, 292
- reflection
 - hyperplane, 291
 - in a hyperplane, 23
- related
 - inner products, 227
 - symmetric
 - matrices, 227
- representation, 38
 - adjoint, 75
 - left regular, 103
- right multiplication map, 187
- ring anti-homomorphism, 120
- root, 279, 292
 - fundamental, 292
 - negative, 293
 - non-trivial, 279
 - positive, 293
 - simple, 292, 293
 - space, 279
 - decomposition, 280
 - null, 279
 - system, 292
 - dual, 301
 - reduced, 292
- rotation, 22
- semi-direct product
 - of groups, 19
 - of Lie algebras, 71
- semi-simple
 - algebra, 106
 - Lie group, 276
- semigroup
 - one-parameter, 57
- separable, 182
- simple, 70
 - algebra, 104

- coroot, 300
- Lie
- group, 274
- Lie algebra, 70
- root, 292, 293
- skew
 - field, 100
 - hermitian, 83
 - symmetric, 26
 - form, 229
- smooth
 - action, 208
 - curve, 183
 - manifold, 181, 182
 - map, 181, 182
- space
 - Hausdorff, 14
 - homogeneous, 214
 - projective, 217
- spacelike vector, 158
- special
 - linear group, 5
 - orthogonal group, 21
 - unitary group, 28
 - projective, 219
- sphere
 - unit, 144
- spinor group, 144
- stabiliser, 37
- standard torus, 252
- strictly upper triangular, 81
- subalgebra
 - commutator, 69
 - derived, 69
 - Lie, 69
- subfield
 - central, 102
- subgroup
 - Borel, 18
 - generated by a subset, 144
 - Lie, 187
 - matrix, 15, 16
 - one-parameter, 57
 - translation, 19
 - unipotent, 18
 - upper triangular, 18
- submanifold, 186
- subset
 - clopen, 17
 - closed, 8
- supremum norm, 6
- symmetric
 - group, 153
 - matrices, related, 227
 - real bilinear form, 87
 - skew, 26
- symplectic
 - form, 26, 229
 - geometry, 26
 - group, 27, 121
 - group, quaternionic, 121
 - group, real, 27
- tangent space, 71, 184
- timelike vector, 158
- topological
 - generator, 206, 254
 - group, 12
 - space
 - compact, 13
- topology
 - product, 11
 - quotient, 33, 212
- torus
 - maximal, 251, 255
 - of rank r , 252
 - standard, 252
- translation subgroup, 19
- transpose, 20
- trivial centre, 70
- unimodular, 4
- group, 5
- matrix, 4
- union
 - disjoint, 21
- unipotent
 - matrix, 18
 - subgroup, 18
- unit sphere, 144
- unitary
 - group, 27
 - projective, 219
- universal property, 132
 - of a Clifford algebra, 132
 - of the quotient topology, 213
- upper triangular
 - matrix, 18, 171
 - strictly, 81
 - subgroup, 18
- vector product, 68
- Weyl group, 260, 292

A SERIES OF TEXTBOOKS FOR UNDERGRADUATES
IN MATHEMATICAL SCIENCES



SPRINGER



UNDERGRADUATE



MATHEMATICS



SERIES

SUMS is a series for undergraduates in the mathematical sciences. From core foundational material to final year topics, **SUMS** books take a fresh and modern approach and are ideal for self-study for a one- or two-semester course.

SUMS books provide the student with:

- ⊕ a user-friendly approach with a clear and logical format;
- ⊕ plenty of examples, problems and fully-worked solutions;
- ⊕ up-to-date material which can be used either to support a course of lectures or for self-study.

All books are, of course, available from all good booksellers (who can order them even if they are not in stock), but if you have difficulties you can contact the publishers direct, by telephoning +44 (0) 1483 418822 (in the UK & Europe), +1/212/450/1500 (in the USA), or by emailing orders@svt.co.uk



**You are one click away
from a world of mathematics information!**

Come and visit Springer's

Mathematics Online Library

Books

- Search the Springer website catalogue
- Subscribe to our free alerting service for new books
- Look through the book series profiles

You want to order?

Email to: orders@springer.de

Journals

- Get abstracts, ToCs free of charge
- Use our powerful search engine LINK Search
- Subscribe to our free alerting service LINK Alert
- Read full-text articles (available only to subscribers of the journal)

You want to subscribe?

Email to: subscriptions@springer.de

Electronic Media

- Get more information on our software and CD-ROMs

Email to: helpdesk-em@springer.de

Bookmark now:

**[http://
www.springer.de/math/](http://www.springer.de/math/)**

Springer · Customer Service
Haberstr. 7 · D-69126 Heidelberg, Germany
Tel: +49 6221 345 200 · Fax: +49 6221 345 229
JSp: +49 6221 345 001a-1c



Springe

Aimed at advanced undergraduate and beginning graduate students, this book provides a first taste of the theory of Lie groups as an appetiser for a more substantial further course. Lie theoretic ideas lie at the heart of much of standard undergraduate linear algebra and exposure to them can inform or motivate the study of the latter.

The main focus is on matrix groups, i.e., closed subgroups of real and complex general linear groups. The first part studies examples and describes the classical families of simply connected compact groups. The second part introduces the idea of a Lie group and studies the associated notion of a homogeneous space using orbits of smooth actions.

Throughout, the emphasis is on providing an approach that is accessible to readers equipped with a standard undergraduate toolkit of algebra and analysis. Although the formal prerequisites are kept at a low level as possible, the subject matter is sophisticated and contains many of the key themes of the fully developed theory, preparing students for a more standard and abstract course in Lie theory and differential geometry.

The Springer Undergraduate Mathematics Series (SUMS) is designed for undergraduates in the mathematical sciences. From core foundational material to final year topics, SUMS books take a fresh and modern approach and are ideal for self-study or for a one- or two-semester course. Each book includes numerous examples, problems and fully-worked solutions.

$$A(B) = \frac{P}{\gamma B)$$

ISBN 1-85233-470-3

www.springer-ny.com
www.springer.co.uk
www.springer.de

ISBN 1-85233-470-3



9 781852 334703

Andrew Baker



Matrix Group

An Introduction to Lie Group Theory

$\gamma \leq \gamma$

SPF

UNDER

MATH

SI

$P_B = P \cap B$

QA
174.2
B35
2002
inger