Brian C. Hall

Lie Groups, Lie Algebras, and Representations

An Elementary Introduction

With 31 Illustrations

Springer

Brian C. Hall
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556-4618
USA
bhall@nd.edu

Editorial Board

S. Axler
Mathematics Department
San Francisco State
University
San Francisco, CA 94132
USA
axler@sfsu.edu

F.W. Gehring
Mathematics Department
East Hall
University of Michigan
Ann Arbor, MI 48109
USA
fgehring@math.lsa.umich.edu

K.A. Ribet
Mathematics Department
University of California,
Berkeley
Berkeley, CA 94720-3840
USA
ribet@math.berkeley.edu

Mathematics Subject Classification (2000): 22-01, 14L35, 20G05

Library of Congress Cataloging-in-Publication Data
Hall, Brian C.
Lie groups, Lie algebras, and representations : an elementary introduction / Brian C. Hall.
p. cm. — (Graduate texts in mathematics ; 222)
Includes index.
ISBN 0-387-40122-9 (acid-free paper)
1. Lie groups. 2. Lie algebras. 3. Representations of groups. 4. Representations of algebras. I. Title. II. Series.
QA387.H34 2003
512′.55—dc21 2003054237

ISBN 0-387-40122-9 Printed on acid-free paper.

© 2003 Springer-Verlag New York, Inc.
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, Inc., 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now know or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks and similar terms, even if the are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed in the United States of America. (EB)

9 8 7 6 5 4 3 2 (Corrected second printing, 2004) SPIN 11002024

Springer-Verlag is a part of *Springer Science+Business Media*

springeronline.com

Preface

This book provides an introduction to Lie groups, Lie algebras, and representation theory, aimed at graduate students in mathematics and physics. Although there are already several excellent books that cover many of the same topics, this book has two distinctive features that I hope will make it a useful addition to the literature. First, it treats Lie groups (not just Lie algebras) in a way that minimizes the amount of manifold theory needed. Thus, I neither assume a prior course on differentiable manifolds nor provide a condensed such course in the beginning chapters. Second, this book provides a gentle introduction to the machinery of semisimple groups and Lie algebras by treating the representation theory of $\mathsf{SU}(2)$ and $\mathsf{SU}(3)$ in detail before going to the general case. This allows the reader to see roots, weights, and the Weyl group "in action" in simple cases before confronting the general theory.

The standard books on Lie theory begin immediately with the general case: a smooth manifold that is also a group. The Lie algebra is then defined as the space of left-invariant vector fields and the exponential mapping is defined in terms of the flow along such vector fields. This approach is undoubtedly the right one in the long run, but it is rather abstract for a reader encountering such things for the first time. Furthermore, with this approach, one must either assume the reader is familiar with the theory of differentiable manifolds (which rules out a substantial part of one's audience) or one must spend considerable time at the beginning of the book explaining this theory (in which case, it takes a long time to get to Lie theory proper).

My way out of this dilemma is to consider only matrix groups (i.e., closed subgroups of $\mathsf{GL}(n;\mathbb{C})$). (Others before me have taken such an approach, as discussed later.) Every such group is a Lie group, and although not every Lie group is of this form, most of the interesting examples are. The exponential of a matrix is then defined by the usual power series, and the Lie algebra $\mathfrak{g}$ of a closed subgroup G of $\mathsf{GL}(n;\mathbb{C})$ is defined to be the set of matrices X such that $\exp(tX)$ lies in G for all real numbers t. One can show that $\mathfrak{g}$ is, indeed, a Lie algebra (i.e., a vector space and closed under commutators). The usual elementary results can all be proved from this point of view: the image of the

exponential mapping contains a neighborhood of the identity; in a connected group, every element is a product of exponentials; every continuous group homomorphism induces a Lie algebra homomorphism. (These results show that every matrix group is a smooth embedded submanifold of $\mathsf{GL}(n;\mathbb{C})$, and hence a Lie group.)

I also address two deeper results: that in the simply-connected case, every Lie algebra homomorphism induces a group homomorphism and that there is a one-to-one correspondence between subalgebras $\mathfrak{h}$ of $\mathfrak{g}$ and connected Lie subgroups H of G. The usual approach to these theorems makes use of the Frobenius theorem. Although this is a fundamental result in analysis, it is not easily stated (let alone proved) and it is not especially Lie-theoretic. My approach is to use, instead, the Baker–Campbell–Hausdorff theorem. This theorem is more elementary than the Frobenius theorem and arguably gives more intuition as to why the above-mentioned results are true. I begin with the technically simpler case of the Heisenberg group (where the Baker–Campbell–Hausdorff series terminates after the first commutator term) and then proceed to the general case.

Appendix C gives two examples of Lie groups that are not matrix Lie groups. Both examples are constructed from matrix Lie groups: One is the universal cover of $\mathsf{SL}(n;\mathbb{R})$ and the other is the quotient of the Heisenberg group by a discrete central subgroup. These examples show the limitations of working with matrix Lie groups, namely that important operations such as the of taking quotients and covers do not preserves the class of matrix Lie groups. In the long run, then, the theory of matrix Lie groups is not an acceptable substitute for general Lie group theory. Nevertheless, I feel that the matrix approach is suitable for a first course in the subject not only because most of the interesting examples of Lie groups are matrix groups but also because all of the theorems I will discuss for the matrix case continue to hold for general Lie groups. In fact, most of the proofs are the same in the general case, *except* that in the general case, one needs to spend a lot more time setting up the basic notions before one can begin.

In addressing the theory of semisimple groups and Lie algebras, I use representation theory as a motivation for the structure theory. In particular, I work out in detail the representation theory of $\mathsf{SU}(2)$ (or, equivalently, $\mathsf{sl}(2;\mathbb{C})$) and $\mathsf{SU}(3)$ (or, equivalently, $\mathsf{sl}(3;\mathbb{C})$) before turning to the general semisimple case. The $\mathsf{sl}(3;\mathbb{C})$ case (more so than just the $\mathsf{sl}(2;\mathbb{C})$ case) illustrates in a concrete way the significance of the Cartan subalgebra, the roots, the weights, and the Weyl group. In the general semisimple case, I keep the representation theory at the fore, introducing at first only as much structure as needed to state the theorem of the highest weight. I then turn to a more detailed look at root systems, including two- and three-dimensional examples, Dynkin diagrams, and a discussion (without proof) of the classification. This portion of the text includes numerous images of the relevant structures (root systems, lattices of dominant integral elements, and weight diagrams) in ranks two and three.

I take full advantage, in treating the semisimple theory, of the correspondence established earlier between the representations of a simply-connected group and the representations of its Lie algebra. So, although I treat things from the point of view of complex semisimple Lie algebras, I take advantage of the characterization of such algebras as ones isomorphic to the complexification of the Lie algebra of a compact simply-connected Lie group K. (Although, for the purposes of this book, we could take this as the definition of a complex semisimple Lie algebra, it is equivalent to the usual algebraic definition.) Having the compact group at our disposal simplifies several issues. First and foremost, it implies the complete reducibility of the representations. Second, it gives a simple construction of Cartan subalgebras, as the complexification of any maximal abelian subalgebra of the Lie algebra of K. Third, it gives a more transparent construction of the Weyl group, as $W = N(T)/T$, where T is a maximal torus in K. This description makes it evident, for example, why the weights of any representation are invariant under the action of W. Thus, my treatment is a mixture of the Lie algebra approach of Humphreys (1972) and the compact group approach of Bröcker and tom Dieck (1985) or Simon (1996).

This book is intended to supplement rather than replace the standard texts on Lie theory. I recommend especially four texts for further reading: the book of Lee (2003) for manifold theory and the relationship between Lie groups and Lie algebras, the book of Humphreys (1972) for the Lie algebra approach to representation theory, the book of Bröcker and tom Dieck (1985) for the compact-group approach to representation theory, and the book of Fulton and Harris (1991) for numerous examples of representations of the classical groups. There are, of course, many other books worth consulting; some of these are listed in the Bibliography.

I hope that by keeping the mathematical prerequisites to a minimum, I have made this book accessible to students in physics as well as mathematics. Although much of the material in the book is widely used in physics, physics students are often expected to pick up the material by osmosis. I hope that they can benefit from a treatment that is elementary but systematic and mathematically precise. In Appendix A, I provide a quick introduction to the theory of groups (not necessarily Lie groups), which is not as standard a part of the physics curriculum as it is of the mathematics curriculum.

The main prerequisite for this book is a solid grounding in linear algebra, especially eigenvectors and the notion of diagonalizability. A quick review of the relevant material is provided in Appendix B. In addition to linear algebra, only elementary analysis is needed: limits, derivatives, and an occasional use of compactness and the inverse function theorem.

There are, to my knowledge, five other treatments of Lie theory from the matrix group point of view. These are (in order of publication) the book *Linear Lie Groups*, by Hans Freudenthal and H. de Vries, the book *Matrix Groups*, by Morton L. Curtis, the article "Very Basic Lie Theory," by Roger Howe, and the recent books *Matrix Groups: An Introduction to Lie Group Theory*,

by Andrew Baker, and *Lie Groups: An Introduction Through Linear Groups*, by Wulf Rossmann. (All of these are listed in the Bibliography.) The book of Freudenthal and de Vries covers a lot of ground, but its unorthodox style and notation make it rather inaccessible. The works of Curtis, Howe, and Baker overlap considerably, in style and content, with the first two chapters of this book, but do not attempt to cover as much ground. For example, none of them treats representation theory or the Baker–Campbell–Hausdorff formula. The book of Rossmann has many similarities with this book, including the use of the Baker–Campbell–Hausdorff formula. However, Rossmann's book is a bit different at the technical level, in that he considers arbitrary subgroups of $\mathrm{GL}(n;\mathbb{C})$, with no restriction on the topology.

Although the organization of this book is, I believe, substantially different from that of other books on the subject, I make no claim to originality in any of the proofs. I myself learned most of the material here from books listed in the Bibliography, especially Humphreys (1972), Bröcker and tom Dieck (1985), and Miller (1972).

I am grateful to many who made corrections, large and small, to the text before publication, including Ed Bueler, Wesley Calvert, Tom Goebeler, Ruth Gornet, Keith Hubbard, Wicharn Lewkeeratiyutkul, Jeffrey Mitchell, Ambar Sengupta, and Erdinch Tatar. I am grateful as well to those who have pointed out errors in the first printing (which have been corrected in this, the second printing), including Moshe Adrian, Kamthorn Chailuek, Paul Gibson, Keith Hubbard, Dennis Muhonen, Jason Quinn, Rebecca Weber, and Reed Wickner.

I also thank Paul Hildebrant for assisting with the construction of models of rank-three root systems using Zome, Judy Hygema for taking digital photographs of the models, and Charles Albrecht for rendering the color images. Finally, I especially thank Scott Vorthmann for making available to the vZome software and for assisting me in its use.

I welcome comments by e-mail at bhall@nd.edu. Please visit my web site at http://www.nd.edu/~bhall/ for more information, including an up-to-date list of corrections and many more color pictures than could be included in the book.

Notre Dame, Indiana
May 2004

Brian C. Hall

Contents

Part I General Theory

1 Matrix Lie Groups 3
1.1 Definition of a Matrix Lie Group 3
1.1.1 Counterexamples 4
1.2 Examples of Matrix Lie Groups 4
1.2.1 The general linear groups $\mathsf{GL}(n;\mathbb{R})$ and $\mathsf{GL}(n;\mathbb{C})$ 4
1.2.2 The special linear groups $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$ 5
1.2.3 The orthogonal and special orthogonal groups, $\mathsf{O}(n)$ and $\mathsf{SO}(n)$ 5
1.2.4 The unitary and special unitary groups, $\mathsf{U}(n)$ and $\mathsf{SU}(n)$ 6
1.2.5 The complex orthogonal groups, $\mathsf{O}(n;\mathbb{C})$ and $\mathsf{SO}(n;\mathbb{C})$. 6
1.2.6 The generalized orthogonal and Lorentz groups 7
1.2.7 The symplectic groups $\mathsf{Sp}(n;\mathbb{R})$, $\mathsf{Sp}(n;\mathbb{C})$, and $\mathsf{Sp}(n)$ ... 7
1.2.8 The Heisenberg group H 8
1.2.9 The groups $\mathbb{R}^*$, $\mathbb{C}^*$, S^1, $\mathbb{R}$, and $\mathbb{R}^n$ 9
1.2.10 The Euclidean and Poincaré groups $\mathsf{E}(n)$ and $\mathsf{P}(n;1)$ ... 9
1.3 Compactness 11
1.3.1 Examples of compact groups 11
1.3.2 Examples of noncompact groups 11
1.4 Connectedness 12
1.5 Simple Connectedness 15
1.6 Homomorphisms and Isomorphisms 17
1.6.1 Example: $\mathsf{SU}(2)$ and $\mathsf{SO}(3)$ 18
1.7 The Polar Decomposition for $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$ 19
1.8 Lie Groups 20
1.9 Exercises 23

2 Lie Algebras and the Exponential Mapping 27
2.1 The Matrix Exponential 27
2.2 Computing the Exponential of a Matrix 30

2.2.1 Case 1: X is diagonalizable 30
2.2.2 Case 2: X is nilpotent 31
2.2.3 Case 3: X arbitrary 32
2.3 The Matrix Logarithm 32
2.4 Further Properties of the Matrix Exponential 35
2.5 The Lie Algebra of a Matrix Lie Group 38
2.5.1 Physicists' Convention 39
2.5.2 The general linear groups 39
2.5.3 The special linear groups 40
2.5.4 The unitary groups 40
2.5.5 The orthogonal groups 40
2.5.6 The generalized orthogonal groups 41
2.5.7 The symplectic groups 41
2.5.8 The Heisenberg group 41
2.5.9 The Euclidean and Poincaré groups 42
2.6 Properties of the Lie Algebra 43
2.7 The Exponential Mapping 48
2.8 Lie Algebras 53
2.8.1 Structure constants 56
2.8.2 Direct sums 56
2.9 The Complexification of a Real Lie Algebra 56
2.10 Exercises 58

3 The Baker–Campbell–Hausdorff Formula 63
3.1 The Baker–Campbell–Hausdorff Formula for the Heisenberg Group 63
3.2 The General Baker–Campbell–Hausdorff Formula 67
3.3 The Derivative of the Exponential Mapping 70
3.4 Proof of the Baker–Campbell–Hausdorff Formula 73
3.5 The Series Form of the Baker–Campbell–Hausdorff Formula 74
3.6 Group Versus Lie Algebra Homomorphisms 76
3.7 Covering Groups 80
3.8 Subgroups and Subalgebras 82
3.9 Exercises 88

4 Basic Representation Theory 91
4.1 Representations 91
4.2 Why Study Representations? 94
4.3 Examples of Representations 95
4.3.1 The standard representation 95
4.3.2 The trivial representation 96
4.3.3 The adjoint representation 96
4.3.4 Some representations of $\mathsf{SU}(2)$ 97
4.3.5 Two unitary representations of $\mathsf{SO}(3)$ 99
4.3.6 A unitary representation of the reals 100

4.3.7 The unitary representations of the Heisenberg group ... 100
4.4 The Irreducible Representations of $\mathfrak{su}(2)$ ... 101
4.5 Direct Sums of Representations ... 106
4.6 Tensor Products of Representations ... 107
4.7 Dual Representations ... 112
4.8 Schur's Lemma ... 113
4.9 Group Versus Lie Algebra Representations ... 115
4.10 Complete Reducibility ... 118
4.11 Exercises ... 121

Part II Semisimple Theory

5 The Representations of SU(3) ... 127
5.1 Introduction ... 127
5.2 Weights and Roots ... 129
5.3 The Theorem of the Highest Weight ... 132
5.4 Proof of the Theorem ... 135
5.5 An Example: Highest Weight $(\mathbf{1},\mathbf{1})$ ... 140
5.6 The Weyl Group ... 142
5.7 Weight Diagrams ... 149
5.8 Exercises ... 152

6 Semisimple Lie Algebras ... 155
6.1 Complete Reducibility and Semisimple Lie Algebras ... 156
6.2 Examples of Reductive and Semisimple Lie Algebras ... 161
6.3 Cartan Subalgebras ... 162
6.4 Roots and Root Spaces ... 164
6.5 Inner Products of Roots and Co-roots ... 170
6.6 The Weyl Group ... 173
6.7 Root Systems ... 180
6.8 Positive Roots ... 181
6.9 The $\mathfrak{sl}(n;\mathbb{C})$ Case ... 182
6.9.1 The Cartan subalgebra ... 182
6.9.2 The roots ... 182
6.9.3 Inner products of roots ... 183
6.9.4 The Weyl group ... 184
6.9.5 Positive roots ... 184
6.10 Uniqueness Results ... 184
6.11 Exercises ... 185

7 Representations of Complex Semisimple Lie Algebras ... 191
7.1 Integral and Dominant Integral Elements ... 192
7.2 The Theorem of the Highest Weight ... 194
7.3 Constructing the Representations I: Verma Modules ... 200

7.3.1 Verma modules 200
7.3.2 Irreducible quotient modules 202
7.3.3 Finite-dimensional quotient modules 204
7.3.4 The $\mathsf{sl}(2;\mathbb{C})$ case 208
7.4 Constructing the Representations II: The Peter–Weyl Theorem 209
7.4.1 The Peter–Weyl theorem 210
7.4.2 The Weyl character formula 211
7.4.3 Constructing the representations 213
7.4.4 Analytically integral versus algebraically integral elements 215
7.4.5 The $\mathsf{SU}(2)$ case 216
7.5 Constructing the Representations III: The Borel–Weil Construction 218
7.5.1 The complex-group approach 218
7.5.2 The setup 220
7.5.3 The strategy 222
7.5.4 The construction 225
7.5.5 The $\mathsf{SL}(2;\mathbb{C})$ case 229
7.6 Further Results 230
7.6.1 Duality 230
7.6.2 The weights and their multiplicities 232
7.6.3 The Weyl character formula and the Weyl dimension formula 234
7.6.4 The analytical proof of the Weyl character formula 236
7.7 Exercises 240

8 More on Roots and Weights 243
8.1 Abstract Root Systems 243
8.2 Duality 248
8.3 Bases and Weyl Chambers 249
8.4 Integral and Dominant Integral Elements 254
8.5 Examples in Rank Two 256
8.5.1 The root systems 256
8.5.2 Connection with Lie algebras 257
8.5.3 The Weyl groups 257
8.5.4 Duality 258
8.5.5 Positive roots and dominant integral elements 258
8.5.6 Weight diagrams 259
8.6 Examples in Rank Three 262
8.7 Additional Properties 263
8.8 The Root Systems of the Classical Lie Algebras 265
8.8.1 The orthogonal algebras $\mathsf{so}(2n;\mathbb{C})$ 265
8.8.2 The orthogonal algebras $\mathsf{so}(2n+1;\mathbb{C})$ 266
8.8.3 The symplectic algebras $\mathsf{sp}(n;\mathbb{C})$ 268
8.9 Dynkin Diagrams and the Classification 269

8.10 The Root Lattice and the Weight Lattice 273
8.11 Exercises 276

A A Quick Introduction to Groups 279
A.1 Definition of a Group and Basic Properties 279
A.2 Examples of Groups 281
A.2.1 The trivial group 282
A.2.2 The integers 282
A.2.3 The reals and $\mathbb{R}^n$ 282
A.2.4 Nonzero real numbers under multiplication 282
A.2.5 Nonzero complex numbers under multiplication 282
A.2.6 Complex numbers of absolute value 1 under multiplication 283
A.2.7 The general linear groups 283
A.2.8 Permutation group (symmetric group) 283
A.2.9 Integers **mod** n 283
A.3 Subgroups, the Center, and Direct Products 284
A.4 Homomorphisms and Isomorphisms 285
A.5 Quotient Groups 286
A.6 Exercises 289

B Linear Algebra Review 291
B.1 Eigenvectors, Eigenvalues, and the Characteristic Polynomial . 291
B.2 Diagonalization 293
B.3 Generalized Eigenvectors and the SN Decomposition 294
B.4 The Jordan Canonical Form 296
B.5 The Trace 296
B.6 Inner Products 297
B.7 Dual Spaces 299
B.8 Simultaneous Diagonalization 299

C More on Lie Groups 303
C.1 Manifolds 303
C.1.1 Definition 303
C.1.2 Tangent space 304
C.1.3 Differentials of smooth mappings 305
C.1.4 Vector fields 306
C.1.5 The flow along a vector field 307
C.1.6 Submanifolds of vector spaces 308
C.1.7 Complex manifolds 309
C.2 Lie Groups 309
C.2.1 Definition 309
C.2.2 The Lie algebra 310
C.2.3 The exponential mapping 311
C.2.4 Homomorphisms 311

C.2.5 Quotient groups and covering groups . . . 312
C.2.6 Matrix Lie groups as Lie groups . . . 313
C.2.7 Complex Lie groups . . . 313
C.3 Examples of Nonmatrix Lie Groups . . . 314
C.4 Differential Forms and Haar Measure . . . 318

D Clebsch–Gordan Theory for SU(2) and the Wigner–Eckart Theorem . . . 321
D.1 Tensor Products of $\mathsf{sl}(2;\mathbb{C})$ Representations . . . 321
D.2 The Wigner–Eckart Theorem . . . 324
D.3 More on Vector Operators . . . 328

E Computing Fundamental Groups of Matrix Lie Groups . . . 331
E.1 The Fundamental Group . . . 331
E.2 The Universal Cover . . . 332
E.3 Fundamental Groups of Compact Lie Groups I . . . 333
E.4 Fundamental Groups of Compact Lie Groups II . . . 336
E.5 Fundamental Groups of Noncompact Lie Groups . . . 342

References . . . 345

Index . . . 347

Part I

General Theory

1

Matrix Lie Groups

1.1 Definition of a Matrix Lie Group

We begin with a very important class of groups, the general linear groups. The groups we will study in this book will all be subgroups (of a certain sort) of one of the general linear groups. This chapter makes use of various standard results from linear algebra that are summarized in Appendix B. This chapter also assumes basic facts and definitions from the theory of abstract groups; the necessary information is provided in Appendix A.

Definition 1.1. *The* ***general linear group*** *over the real numbers, denoted* $\mathsf{GL}(n;\mathbb{R})$*, is the group of all* $n \times n$ *invertible matrices with real entries. The general linear group over the complex numbers, denoted* $\mathsf{GL}(n;\mathbb{C})$*, is the group of all* $n \times n$ *invertible matrices with complex entries.*

The general linear groups are indeed groups under the operation of matrix multiplication: The product of two invertible matrices is invertible, the identity matrix is an identity for the group, an invertible matrix has (by definition) an inverse, and matrix multiplication is associative.

Definition 1.2. *Let* $M_n(\mathbb{C})$ *denote the space of all* $n \times n$ *matrices with complex entries.*

Definition 1.3. *Let* A_m *be a sequence of complex matrices in* $M_n(\mathbb{C})$*. We say that* A_m ***converges*** *to a matrix* A *if each entry of* A_m *converges (as* $m \to \infty$*) to the corresponding entry of* A *(i.e., if* $(A_m)_{kl}$ *converges to* A_{kl} *for all* $1 \leq k, l \leq n$*).*

Definition 1.4. *A* ***matrix Lie group*** *is any subgroup* G *of* $\mathsf{GL}(n;\mathbb{C})$ *with the following property: If* A_m *is any sequence of matrices in* G*, and* A_m *converges to some matrix* A *then either* $A \in G$*, or* A *is not invertible.*

The condition on G amounts to saying that G is a closed subset of $\mathsf{GL}(n;\mathbb{C})$. (This does not necessarily mean that G is closed in $M_n(\mathbb{C})$.) Thus, Definition

1.4 is equivalent to saying that a matrix Lie group is a **closed subgroup** of $\mathsf{GL}(n;\mathbb{C})$.

The condition that G be a *closed* subgroup, as opposed to merely a subgroup, should be regarded as a technicality, in that most of the *interesting* subgroups of $\mathsf{GL}(n;\mathbb{C})$ have this property. (Most of the matrix Lie groups G we will consider have the stronger property that if A_m is any sequence of matrices in G, and A_m converges to some matrix A, then $A \in G$ (i.e., that G is closed in $M_n(\mathbb{C})$).)

1.1.1 Counterexamples

An example of a subgroup of $\mathsf{GL}(n;\mathbb{C})$ which is not closed (and hence is not a matrix Lie group) is the set of all $n \times n$ invertible matrices all of whose entries are real and rational. This is in fact a subgroup of $\mathsf{GL}(n;\mathbb{C})$, but not a closed subgroup. That is, one can (easily) have a sequence of invertible matrices with rational entries converging to an invertible matrix with some irrational entries. (In fact, *every* real invertible matrix is the limit of some sequence of invertible matrices with rational entries.)

Another example of a group of matrices which is not a matrix Lie group is the following subgroup of $\mathsf{GL}(2;\mathbb{C})$. Let a be an irrational real number and let

$$G = \left\{ \left(\begin{array}{cc} e^{it} & 0 \\ 0 & e^{ita} \end{array} \right) \middle| \, t \in \mathbb{R} \right\}.$$

Clearly, G is a subgroup of $\mathsf{GL}(2,\mathbb{C})$. Because a is irrational, the matrix $-I$ is not in G, since to make e^{it} equal to -1, we must take t to be an odd integer multiple of π, in which case ta cannot be an odd integer multiple of π. On the other hand (Exercise 1), by taking $t = (2n+1)\pi$ for a suitably chosen integer n, we can make ta arbitrarily *close* to an odd integer multiple of π. Hence, we can find a sequence of matrices in G which converges to $-I$, and so G is not a matrix Lie group. See Exercise 1 and Exercise 18 for more information.

1.2 Examples of Matrix Lie Groups

Mastering the subject of Lie groups involves not only learning the general theory but also familiarizing oneself with examples. In this section, we introduce some of the most important examples of (matrix) Lie groups.

1.2.1 The general linear groups $\mathsf{GL}(n;\mathbb{R})$ and $\mathsf{GL}(n;\mathbb{C})$

The general linear groups (over $\mathbb{R}$ or $\mathbb{C}$) are themselves matrix Lie groups. Of course, $\mathsf{GL}(n;\mathbb{C})$ is a subgroup of itself. Furthermore, if A_m is a sequence of matrices in $\mathsf{GL}(n;\mathbb{C})$ and A_m converges to A, then by the definition of $\mathsf{GL}(n;\mathbb{C})$, either A is in $\mathsf{GL}(n;\mathbb{C})$, or A is not invertible.

Moreover, $\mathsf{GL}(n;\mathbb{R})$ is a subgroup of $\mathsf{GL}(n;\mathbb{C})$, and if $A_m \in \mathsf{GL}(n;\mathbb{R})$ and A_m converges to A, then the entries of A are real. Thus, either A is not invertible or $A \in \mathsf{GL}(n;\mathbb{R})$.

1.2.2 The special linear groups $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$

The **special linear group** (over $\mathbb{R}$ or $\mathbb{C}$) is the group of $n \times n$ invertible matrices (with real or complex entries) having determinant one. Both of these are subgroups of $\mathsf{GL}(n;\mathbb{C})$. Furthermore, if A_n is a sequence of matrices with determinant one and A_n converges to A, then A also has determinant one, because the determinant is a continuous function. Thus, $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$ are matrix Lie groups.

1.2.3 The orthogonal and special orthogonal groups, $\mathsf{O}(n)$ and $\mathsf{SO}(n)$

An $n \times n$ real matrix A is said to be **orthogonal** if the column vectors that make up A are orthonormal, that is, if

$$\sum_{l=1}^{n} A_{lj} A_{lk} = \delta_{jk}, \quad 1 \leq j, k \leq n.$$

(Here δ_{jk} is the Kronecker delta, equal to 1 if $j = k$ and equal to zero if $j \neq k$.) Equivalently, A is orthogonal if it preserves the inner product, namely if $\langle x, y\rangle = \langle Ax, Ay\rangle$ for all vectors x, y in $\mathbb{R}^n$. (Angled brackets denote the usual inner product on $\mathbb{R}^n$, $\langle x, y\rangle = \sum_k x_k y_k$.) Still another equivalent definition is that A is orthogonal if $A^{tr}A = I$, i.e., if $A^{tr} = A^{-1}$. (Here, A^{tr} is the transpose of A, $(A^{tr})_{kl} = A_{lk}$.) See Exercise 2.

Since $\det A^{tr} = \det A$, we see that if A is orthogonal, then $\det(A^{tr}A) = (\det A)^2 = \det I = 1$. Hence, $\det A = \pm 1$, for all orthogonal matrices A.

This formula tells us in particular that every orthogonal matrix must be invertible. However, if A is an orthogonal matrix, then

$$\left\langle A^{-1}x, A^{-1}y\right\rangle = \left\langle A(A^{-1}x), A(A^{-1}y)\right\rangle = \langle x, y\rangle .$$

Thus, the inverse of an orthogonal matrix is orthogonal. Furthermore, the product of two orthogonal matrices is orthogonal, since if A and B both preserve inner products, then so does AB. Thus, the set of orthogonal matrices forms a group.

The set of all $n \times n$ real orthogonal matrices is the **orthogonal group** $\mathsf{O}(n)$, and it is a subgroup of $\mathsf{GL}(n;\mathbb{C})$. The limit of a sequence of orthogonal matrices is orthogonal, because the relation $A^{tr}A = I$ is preserved under taking limits. Thus, $\mathsf{O}(n)$ is a matrix Lie group.

The set of $n \times n$ orthogonal matrices with determinant one is the **special orthogonal group** $\mathsf{SO}(n)$. Clearly, this is a subgroup of $\mathsf{O}(n)$, and hence of

$\mathsf{GL}(n;\mathbb{C})$. Moreover, both orthogonality and the property of having determinant one are preserved under limits, and so $\mathsf{SO}(n)$ is a matrix Lie group. Since elements of $\mathsf{O}(n)$ already have determinant ± 1, $\mathsf{SO}(n)$ is "half" of $\mathsf{O}(n)$.

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

See also Exercise 6.

1.2.4 The unitary and special unitary groups, $\mathsf{U}(n)$ and $\mathsf{SU}(n)$

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

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

Equivalently, A is unitary if it preserves the inner product, namely if $\langle x, y\rangle = \langle Ax, Ay\rangle$ for all vectors x, y in $\mathbb{C}^n$. (Angled brackets here denote the inner product on $\mathbb{C}^n$, $\langle x, y\rangle = \sum_k \overline{x_k} y_k$. We will adopt the convention of putting the complex conjugate on the left.) Still another equivalent definition is that A is unitary if $A^*A = I$, i.e., if $A^* = A^{-1}$. (Here, A^* is the **adjoint** of A, $(A^*)_{jk} = \overline{A_{kj}}$.) See Exercise 3.

Since $\det A^* = \overline{\det A}$, we see that if A is unitary, then $\det(A^*A) = |\det A|^2 = \det I = 1$. Hence, $|\det A| = 1$, for all unitary matrices A.

This, in particular, shows that every unitary matrix is invertible. The same argument as for the orthogonal group shows that the set of unitary matrices forms a group.

The set of all $n \times n$ unitary matrices is the **unitary group** $\mathsf{U}(n)$, and it is a subgroup of $\mathsf{GL}(n;\mathbb{C})$. The limit of unitary matrices is unitary, so $\mathsf{U}(n)$ is a matrix Lie group. The set of unitary matrices with determinant one is the **special unitary group** $\mathsf{SU}(n)$. It is easy to check that $\mathsf{SU}(n)$ is a matrix Lie group. Note that a unitary matrix can have determinant $e^{i\theta}$ for any θ, and so $\mathsf{SU}(n)$ is a smaller subset of $\mathsf{U}(n)$ than $\mathsf{SO}(n)$ is of $\mathsf{O}(n)$. (Specifically, $\mathsf{SO}(n)$ has the same dimension as $\mathsf{O}(n)$, whereas $\mathsf{SU}(n)$ has dimension one less than that of $\mathsf{U}(n)$.)

See also Exercise 8.

1.2.5 The complex orthogonal groups, $\mathsf{O}(n;\mathbb{C})$ and $\mathsf{SO}(n;\mathbb{C})$

Consider the bilinear form $(\cdot,\cdot)$ on $\mathbb{C}^n$ defined by $(x, y) = \sum_k x_k y_k$. This form is *not* an inner product (Section B.6) because, for example, it is symmetric rather than conjugate-symmetric. The set of all $n \times n$ complex matrices A which preserve this form (i.e., such that $(Ax, Ay) = (x, y)$ for all $x, y \in \mathbb{C}^n$) is the **complex orthogonal group** $\mathsf{O}(n;\mathbb{C})$, and it is a subgroup of $\mathsf{GL}(n;\mathbb{C})$. Repeating the arguments for the case of $\mathsf{SO}(n)$ and $\mathsf{O}(n)$ (but now permitting complex entries), we find that an $n \times n$ complex matrix A is in $\mathsf{O}(n;\mathbb{C})$ if and

only if $A^{tr}A = I$, that $\mathsf{O}(n;\mathbb{C})$ is a matrix Lie group, and that $\det A = \pm 1$ for all A in $\mathsf{O}(n;\mathbb{C})$. Note that $\mathsf{O}(n;\mathbb{C})$ is *not* the same as the unitary group $\mathsf{U}(n)$. The group $\mathsf{SO}(n;\mathbb{C})$ is defined to be the set of all A in $\mathsf{O}(n;\mathbb{C})$ with $\det A = 1$ and it is also a matrix Lie group.

1.2.6 The generalized orthogonal and Lorentz groups

Let n and k be positive integers, and consider $\mathbb{R}^{n+k}$. Define a symmetric bilinear form $[\cdot,\cdot]_{n,k}$ on $\mathbb{R}^{n+k}$ by the formula

$$[x,y]_{n,k} = x_1y_1 + \cdots + x_ny_n - x_{n+1}y_{n+1} - \cdots - x_{n+k}y_{n+k} \tag{1.1}$$

The set of $(n+k)\times(n+k)$ real matrices A which preserve this form (i.e., such that $[Ax,Ay]_{n,k} = [x,y]_{n,k}$ for all $x,y \in \mathbb{R}^{n+k}$) is the **generalized orthogonal group** $\mathsf{O}(n;k)$. It is a subgroup of $\mathsf{GL}(n+k;\mathbb{R})$ and a matrix Lie group (Exercise 4).

If A is an $(n+k)\times(n+k)$ real matrix, let $A^{(i)}$ denote the i^{th} column vector of A, that is,

$$A^{(i)} = \begin{pmatrix} A_{1,i} \\ \vdots \\ A_{n+k,i} \end{pmatrix}.$$

Then, A is in $\mathsf{O}(n;k)$ if and only if the following conditions are satisfied:

$$\begin{aligned} \left[A^{(l)},A^{(j)}\right]_{n,k} &= 0 \quad && l \neq j, \\ \left[A^{(l)},A^{(l)}\right]_{n,k} &= 1 \quad && 1 \leq l \leq n, \\ \left[A^{(l)},A^{(l)}\right]_{n,k} &= -1 \quad && n+1 \leq l \leq n+k. \end{aligned} \tag{1.2}$$

Let g denote the $(n+k)\times(n+k)$ diagonal matrix with ones in the first n diagonal entries and minus ones in the last k diagonal entries. Then, A is in $\mathsf{O}(n;k)$ if and only if $A^{tr}gA = g$ (Exercise 4). Taking the determinant of this equation gives $(\det A)^2 \det g = \det g$, or $(\det A)^2 = 1$. Thus, for any A in $\mathsf{O}(n;k)$, $\det A = \pm 1$.

Of particular interest in physics is the **Lorentz group** $\mathsf{O}(3;1)$. See also Exercise 7.

1.2.7 The symplectic groups $\mathsf{Sp}(n;\mathbb{R})$, $\mathsf{Sp}(n;\mathbb{C})$, and $\mathsf{Sp}(n)$

The special and general linear groups, the orthogonal and unitary groups, and the symplectic groups (which will be defined momentarily) make up the **classical groups**. Of the classical groups, the symplectic groups have the most confusing definition, partly because there are three sets of them ($\mathsf{Sp}(n;\mathbb{R})$, $\mathsf{Sp}(n;\mathbb{C})$, and $\mathsf{Sp}(n)$) and partly because they involve skew-symmetric bilinear forms rather than the more familiar symmetric bilinear forms. To further

confuse matters, the notation for referring to these groups is not consistent from author to author.

Consider the skew-symmetric bilinear form B on $\mathbb{R}^{2n}$ defined as follows:

$$B[x,y] = \sum_{k=1}^{n} x_k y_{n+k} - x_{n+k} y_k. \tag{1.3}$$

The set of all $2n \times 2n$ matrices A which preserve B (i.e., such that $B[Ax, Ay] = B[x,y]$ for all $x, y \in \mathbb{R}^{2n}$) is the **real symplectic group** $\mathsf{Sp}(n;\mathbb{R})$, and it is a subgroup of $\mathsf{GL}(2n;\mathbb{R})$. It is not difficult to check that this is a matrix Lie group (Exercise 5). This group arises naturally in the study of classical mechanics. If J is the $2n \times 2n$ matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

then $B[x,y] = \langle x, Jy \rangle$, and it is possible to check that a $2n \times 2n$ real matrix A is in $\mathsf{Sp}(n;\mathbb{R})$ if and only if $A^{tr} J A = J$. (See Exercise 5.) Taking the determinant of this identity gives $(\det A)^2 \det J = \det J$, or $(\det A)^2 = 1$. This shows that $\det A = \pm 1$, for all $A \in \mathsf{Sp}(n;\mathbb{R})$. In fact, $\det A = 1$ for all $A \in \mathsf{Sp}(n;\mathbb{R})$, although this is not obvious.

One can define a bilinear form on $\mathbb{C}^{2n}$ by the same formula (1.3). (This form involves no complex conjugates.) The set of $2n \times 2n$ complex matrices which preserve this form is the **complex symplectic group** $\mathsf{Sp}(n;\mathbb{C})$. A $2n \times 2n$ complex matrix A is in $\mathsf{Sp}(n;\mathbb{C})$ if and only if $A^{tr} J A = J$. (Note: This condition involves A^{tr}, *not* A^*.) This relation shows that $\det A = \pm 1$, for all $A \in \mathsf{Sp}(n;\mathbb{C})$. In fact, $\det A = 1$, for all $A \in \mathsf{Sp}(n;\mathbb{C})$.

Finally, we have the **compact symplectic group** $\mathsf{Sp}(n)$ defined as

$$\mathsf{Sp}(n) = \mathsf{Sp}(n;\mathbb{C}) \cap \mathsf{U}(2n).$$

See also Exercise 9. For more information and a proof that $\det A = 1$ for all $A \in \mathsf{Sp}(n;\mathbb{C})$, see Section 9.4 of Miller (1972). What we call $\mathsf{Sp}(n;\mathbb{C})$ Miller calls $\mathsf{Sp}(n)$, and what we call $\mathsf{Sp}(n)$, Miller calls $\mathsf{USp}(n)$.

1.2.8 The Heisenberg group H

The set of all 3×3 real matrices A of the form

$$A = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.4}$$

where a, b, and c are arbitrary real numbers, is the **Heisenberg group**. It is easy to check that the product of two matrices of the form (1.4) is again of that form, and, clearly, the identity matrix is of the form (1.4). Furthermore, direct computation shows that if A is as in (1.4), then

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

Thus, H is a subgroup of $\mathsf{GL}(3;\mathbb{R})$. Clearly, the limit of matrices of the form (1.4) is again of that form, and so H is a matrix Lie group.

The reason for the name "Heisenberg group" is that the Lie algebra of H gives a realization of the Heisenberg commutation relations of quantum mechanics. (See especially Chapter 4, Exercise 8.)

See also Exercise 10.

1.2.9 The groups $\mathbb{R}^*$, $\mathbb{C}^*$, S^1, $\mathbb{R}$, and $\mathbb{R}^n$

Several important groups which are not naturally groups of matrices can (and will in these notes) be thought of as such.

The group $\mathbb{R}^*$ of non-zero real numbers under multiplication is isomorphic to $\mathsf{GL}(1;\mathbb{R})$. Thus, we will regard $\mathbb{R}^*$ as a matrix Lie group. Similarly, the group $\mathbb{C}^*$ of nonzero complex numbers under multiplication is isomorphic to $\mathsf{GL}(1;\mathbb{C})$, and the group S^1 of complex numbers with absolute value one is isomorphic to $\mathsf{U}(1)$.

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

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

1.2.10 The Euclidean and Poincaré groups $\mathsf{E}(n)$ and $\mathsf{P}(n;1)$

The **Euclidean group** $\mathsf{E}(n)$ is, by definition, the group of all one-to-one, onto, distance-preserving maps of $\mathbb{R}^n$ to itself, that is, maps $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $d(f(x), f(y)) = d(x, y)$ for all $x, y \in \mathbb{R}^n$. Here, d is the usual distance on $\mathbb{R}^n$: $d(x, y) = |x - y|$. Note that we do not assume *anything* about the structure of f besides the above properties. In particular, f need not be linear. The orthogonal group $\mathsf{O}(n)$ is a subgroup of $\mathsf{E}(n)$ and is the group of all *linear* distance-preserving maps of $\mathbb{R}^n$ to itself. For $x \in \mathbb{R}^n$, define the **translation by** x, denoted T_x, by

$$T_x(y) = x + y.$$

The set of translations is also a subgroup of $\mathsf{E}(n)$.

Proposition 1.5. *Every element T of $\mathsf{E}(n)$ can be written uniquely as an orthogonal linear transformation followed by a translation, that is, in the form*

$$T = T_x R$$

with $x \in \mathbb{R}^n$ *and* $R \in \mathsf{O}(n)$.

We will not prove this. The key step is to prove that every one-to-one, onto, distance-preserving map of $\mathbb{R}^n$ to itself which fixes the origin must be linear. We will write an element $T = T_x R$ of $\mathsf{E}(n)$ as a pair $\{x, R\}$. Note that for $y \in \mathbb{R}^n$,

$$\{x, R\}\, y = Ry + x$$

and that

$$\{x_1, R_1\}\{x_2, R_2\}y = R_1(R_2 y + x_2) + x_1 = R_1 R_2 y + (x_1 + R_1 x_2).$$

Thus, the product operation for $\mathsf{E}(n)$ is the following:

$$\{x_1, R_1\}\{x_2, R_2\} = \{x_1 + R_1 x_2, R_1 R_2\}. \tag{1.5}$$

The inverse of an element of $\mathsf{E}(n)$ is given by

$$\{x, R\}^{-1} = \{-R^{-1}x, R^{-1}\}.$$

As already noted, $\mathsf{E}(n)$ is not a subgroup of $\mathsf{GL}(n;\mathbb{R})$, since translations are not linear maps. However, $\mathsf{E}(n)$ is isomorphic to a subgroup of $\mathsf{GL}(n+1;\mathbb{R})$, via the map which associates to $\{x, R\} \in \mathsf{E}(n)$ the following matrix:

$$\begin{pmatrix} & & x_1 \\ & R & \vdots \\ & & x_n \\ 0 & \cdots\ 0 & 1 \end{pmatrix}. \tag{1.6}$$

This map is clearly one-to-one, and direct computation shows that multiplication of elements of the form (1.6) follows the multiplication rule in (1.5), so that this map is a homomorphism. Thus, $\mathsf{E}(n)$ is isomorphic to the group of all matrices of the form (1.6) (with $R \in \mathsf{O}(n)$). The limit of things of the form (1.6) is again of that form, and so we have expressed the Euclidean group $\mathsf{E}(n)$ as a matrix Lie group.

We similarly define the Poincaré group $\mathsf{P}(n;1)$ to be the group of all transformations of $\mathbb{R}^{n+1}$ of the form

$$T = T_x A$$

with $x \in \mathbb{R}^{n+1}$ and $A \in \mathsf{O}(n;1)$. This is the group of affine transformations of $\mathbb{R}^{n+1}$ which preserve the Lorentz "distance" $d_L(x,y) = (x_1 - y_1)^2 + \cdots + (x_n - y_n)^2 - (x_{n+1} - y_{n+1})^2$. (An affine transformation is one of the form $x \to Ax + b$, where A is a linear transformation and b is constant.) The group product is the obvious analog of the product (1.5) for the Euclidean group.

The Poincaré group $\mathsf{P}(n;1)$ is isomorphic to the group of $(n+2)\times(n+2)$ matrices of the form

$$\begin{pmatrix} & & x_1 \\ & A & \vdots \\ & & x_{n+1} \\ 0 & \cdots\ 0 & 1 \end{pmatrix} \tag{1.7}$$

with $A \in \mathsf{O}(n;1)$. The set of matrices of the form (1.7) is a matrix Lie group.

1.3 Compactness

Definition 1.6. *A matrix Lie group G is said to be* ***compact*** *if the following two conditions are satisfied:*

1. *If A_m is any sequence of matrices in G, and A_m converges to a matrix A, then A is in G.*
2. *There exists a constant C such that for all $A \in G$, $|A_{ij}| \leq C$ for all $1 \leq i,j \leq n$.*

This is not the usual topological definition of compactness. However, the set $M_n(\mathbb{C})$ of all $n \times n$ complex matrices can be thought of as $\mathbb{C}^{n^2}$. The above definition says that G is compact if it is a *closed, bounded* subset of $\mathbb{C}^{n^2}$. It is a standard theorem from elementary analysis that a subset of $\mathbb{C}^{n^2}$ is compact if and only if it is closed and bounded.

All of our examples of matrix Lie groups except $\mathsf{GL}(n;\mathbb{R})$ and $\mathsf{GL}(n;\mathbb{C})$ have property (1). Thus, it is the boundedness condition (2) that is most important.

1.3.1 Examples of compact groups

The groups $\mathsf{O}(n)$ and $\mathsf{SO}(n)$ are compact. Property (1) is satisfied because the limit of orthogonal matrices is orthogonal and the limit of matrices with determinant one has determinant one. Property (2) is satisfied because if A is orthogonal, then the column vectors of A have norm one, and hence $|A_{kl}| \leq 1$, for all $1 \leq k,l \leq n$. A similar argument shows that $\mathsf{U}(n)$, $\mathsf{SU}(n)$, and $\mathsf{Sp}(n)$ are compact. (This includes the unit circle, $S^1 \cong \mathsf{U}(1)$.)

1.3.2 Examples of noncompact groups

All of the other examples given of matrix Lie groups are noncompact. The groups $\mathsf{GL}(n;\mathbb{R})$ and $\mathsf{GL}(n;\mathbb{C})$ violate property (1), since a limit of invertible matrices may be noninvertible. The groups $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}\,(n;\mathbb{C})$ violate (2), (except in the trivial case $n = 1$) since

$$A_m = \begin{pmatrix} m & & & & \\ & \frac{1}{m} & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

has determinant one, no matter how large m is.

The following groups also violate (2), and hence are noncompact: $\mathsf{O}(n;\mathbb{C})$ and $\mathsf{SO}(n;\mathbb{C})$; $\mathsf{O}(n;k)$ and $\mathsf{SO}(n;k)$ ($n \geq 1$, $k \geq 1$); the Heisenberg group H; $\mathsf{Sp}(n;\mathbb{R})$ and $\mathsf{Sp}(n;\mathbb{C})$; $\mathsf{E}(n)$ and $\mathsf{P}(n;1)$; $\mathbb{R}$ and $\mathbb{R}^n$; $\mathbb{R}^*$ and $\mathbb{C}^*$. It is left to the reader to provide examples to show that this is the case.

1.4 Connectedness

Definition 1.7. *A matrix Lie group G is said to be* ***connected*** *if given any two matrices A and B in G, there exists a continuous path $A(t)$, $a \leq t \leq b$, lying in G with $A(a) = A$ and $A(b) = B$.*

This property is what is called **path-connected** in topology, which is not (in general) the same as connected. However, it is a fact (not particularly obvious at the moment) that a matrix Lie group is connected if and only if it is path-connected. So, in a slight abuse of terminology, we shall continue to refer to the above property as connectedness. (See Section 1.8.)

A matrix Lie group G which is not connected can be decomposed (uniquely) as a union of several pieces, called **components**, such that two elements of the same component can be joined by a continuous path, but two elements of different components cannot.

Proposition 1.8. *If G is a matrix Lie group, then the component of G containing the identity is a subgroup of G.*

Proof. Saying that A and B are both in the component containing the identity means that there exist continuous paths $A(t)$ and $B(t)$ with $A(0) = B(0) = I$, $A(1) = A$, and $B(1) = B$. Then, $A(t)B(t)$ is a continuous path starting at I and ending at AB. Thus, the product of two elements of the identity component is again in the identity component. Furthermore, $A(t)^{-1}$ is a continuous path starting at I and ending at A^{-1}, and so the inverse of any element of the identity component is again in the identity component. Thus, the identity component is a subgroup. □

Note that because matrix multiplication and matrix inversion are continuous on $\mathsf{GL}(n;\mathbb{C})$, it follows that if $A(t)$ and $B(t)$ are continuous, then so are $A(t)B(t)$ and $A(t)^{-1}$. The continuity of the matrix product is obvious. The continuity of the inverse follows from the formula for the inverse in terms of cofactors; this formula is continuous as long as we remain in the set of invertible matrices where the determinant in the denominator is nonzero.

Proposition 1.9. *The group* $\mathsf{GL}(n;\mathbb{C})$ *is connected for all* $n \geq 1$.

Proof. Consider first the case $n = 1$. A 1×1 invertible complex matrix A is of the form $A = [\lambda]$ with λ in $\mathbb{C}^*$, the set of nonzero complex numbers. Given any two nonzero complex numbers, we can easily find a continuous path which connects them and does not pass through zero.

For the case $n \geq 2$, we will show that any element of $\mathsf{GL}(n;\mathbb{C})$ can be connected to the identity by a continuous path lying in $\mathsf{GL}(n;\mathbb{C})$. Then, any two elements A and B of $\mathsf{GL}(n;\mathbb{C})$ can be connected by a path going from A to the identity and then from the identity to B.

We make use of the result that every matrix is similar to an upper triangular matrix (Theorem B.7). That is, given any $n \times n$ complex matrix A, there exists an invertible $n \times n$ complex matrix C such that

$$A = CBC^{-1}$$

where B is upper triangular:

$$B = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

If we now assume that A is invertible, then all the λ_i's must be nonzero, since $\det A = \det B = \lambda_1 \cdots \lambda_n$. Let $B(t)$ be obtained by multiplying the part of B above the diagonal by $(1-t)$, for $0 \leq t \leq 1$, and let $A(t) = CB(t)C^{-1}$. Then, $A(t)$ is a continuous path which starts at A and ends at CDC^{-1}, where D is the diagonal matrix

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

This path lies in $\mathsf{GL}(n;\mathbb{C})$ since $\det A(t) = \lambda_1 \cdots \lambda_n = \det A$ for all t.

Now, as in the case $n = 1$, we can define $\lambda_i(t)$, which connects each λ_i to 1 in $\mathbb{C}^*$ as t goes from 1 to 2. Then, we can define $A(t)$ on the interval $1 \leq t \leq 2$ by

$$A(t) = C \begin{pmatrix} \lambda_1(t) & & 0 \\ & \ddots & \\ 0 & & \lambda_n(t) \end{pmatrix} C^{-1}.$$

This is a continuous path which starts at CDC^{-1} when $t = 1$ and ends at I $(= CIC^{-1})$ when $t = 2$. Since the $\lambda_k(t)$'s are always nonzero, $A(t)$ lies in $\mathsf{GL}(n;\mathbb{C})$. We see, then, that every matrix A in $\mathsf{GL}(n;\mathbb{C})$ can be connected to the identity by a continuous path lying in $\mathsf{GL}(n;\mathbb{C})$. □

An alternative proof of this result is given in Exercise 12.

Proposition 1.10. *The group* $\mathsf{SL}(n;\mathbb{C})$ *is connected for all* $n \geq 1$.

Proof. The proof is almost the same as for $\mathsf{GL}(n;\mathbb{C})$, except that we must be careful to preserve the condition $\det A = 1$. Let A be an arbitrary element of $\mathsf{SL}(n;\mathbb{C})$. The case $n = 1$ is trivial, so we assume $n \geq 2$. We can define $A(t)$ as before for $0 \leq t \leq 1$, with $A(0) = A$ and $A(1) = CDC^{-1}$, since $\det A(t) = \det A = 1$. Now, define $\lambda_k(t)$ as before for $1 \leq k \leq n-1$ and define $\lambda_n(t)$ to be $[\lambda_1(t) \cdots \lambda_{n-1}(t)]^{-1}$. (Note that since $\lambda_1 \cdots \lambda_n = 1$, $\lambda_n(1) = \lambda_n$.) This allows us to connect A to the identity while staying within $\mathsf{SL}(n;\mathbb{C})$. □

Proposition 1.11. *The groups* $\mathsf{U}(n)$ *and* $\mathsf{SU}(n)$ *are connected, for all* $n \geq 1$.

Proof. By a standard result of linear algebra (Theorem B.3), every unitary matrix has an orthonormal basis of eigenvectors, with eigenvalues of the form $e^{i\theta}$. It follows that every unitary matrix U can be written as

$$U = U_1 \begin{pmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n} \end{pmatrix} U_1^{-1} \tag{1.8}$$

with U_1 unitary and $\theta_i \in \mathbb{R}$. Conversely, as is easily checked, every matrix of the form (1.8) is unitary. Now, define

$$U(t) = U_1 \begin{pmatrix} e^{i(1-t)\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i(1-t)\theta_n} \end{pmatrix} U_1^{-1}.$$

As t ranges from 0 to 1, this defines a continuous path in $\mathsf{U}(n)$ joining U to I. Thus, any two elements U and V of $\mathsf{U}(n)$ can be connected to each other by a continuous path that runs from U to I and then from I to V.

A slight modification of this argument, as in the proof of Proposition 1.10, shows that $\mathsf{SU}(n)$ is connected. □

Proposition 1.12. *The group* $\mathsf{GL}(n;\mathbb{R})$ *is not connected, but has two components. These are* $\mathsf{GL}(n;\mathbb{R})^+$, *the set of* $n \times n$ *real matrices with positive determinant, and* $\mathsf{GL}(n;\mathbb{R})^-$, *the set of* $n \times n$ *real matrices with negative determinant.*

Proof. $\mathsf{GL}(n;\mathbb{R})$ cannot be connected, for if $\det A > 0$ and $\det B < 0$, then any continuous path connecting A to B would have to include a matrix with determinant zero and hence pass outside of $\mathsf{GL}(n;\mathbb{R})$.

The proof that $\mathsf{GL}(n;\mathbb{R})^+$ is connected is sketched in Exercise 15. Once $\mathsf{GL}(n;\mathbb{R})^+$ is known to be connected, it is not difficult to see that $\mathsf{GL}(n;\mathbb{R})^-$ is also connected. Let C be any matrix with negative determinant and take A and B in $\mathsf{GL}(n;\mathbb{R})^-$. Then, $C^{-1}A$ and $C^{-1}B$ are in $\mathsf{GL}(n;\mathbb{R})^+$ and can be joined by a continuous path $D(t)$ in $\mathsf{GL}(n;\mathbb{R})^+$. However, then, $CD(t)$ is a continuous path joining A and B in $\mathsf{GL}(n;\mathbb{R})^-$. □

The following table lists some matrix Lie groups, indicates whether or not the group is connected, and gives the number of components:

Group	Connected?	Components
$\mathsf{GL}(n;\mathbb{C})$	yes	1
$\mathsf{SL}(n;\mathbb{C})$	yes	1
$\mathsf{GL}(n;\mathbb{R})$	no	2
$\mathsf{SL}(n;\mathbb{R})$	yes	1
$\mathsf{O}(n)$	no	2
$\mathsf{SO}(n)$	yes	1
$\mathsf{U}(n)$	yes	1
$\mathsf{SU}(n)$	yes	1
$\mathsf{O}(n;1)$	no	4
$\mathsf{SO}(n;1)$	no	2
Heisenberg	yes	1
$\mathsf{E}(n)$	no	2
$\mathsf{P}(n;1)$	no	4

Proofs of some of these results are given in Exercises 7, 13, 14, and 15.

1.5 Simple Connectedness

Definition 1.13. *A matrix Lie group G is said to be* ***simply connected*** *if it is connected and, in addition, every loop in G can be shrunk continuously to a point in G.*

More precisely, assume that G is connected. Then, G is simply connected if given any continuous path $A(t)$, $0 \leq t \leq 1$, lying in G with $A(0) = A(1)$, there exists a continuous function $A(s,t)$, $0 \leq s,t \leq 1$, taking values in G and having the following properties: (1) $A(s,0) = A(s,1)$ for all s, (2) $A(0,t) = A(t)$, and (3) $A(1,t) = A(1,0)$ for all t.

One should think of $A(t)$ as a loop and $A(s,t)$ as a family of loops, parameterized by the variable s which shrinks $A(t)$ to a point. Condition 1 says that for each value of the parameter s, we have a loop; condition 2 says that when $s = 0$ the loop is the specified loop $A(t)$; and condition 3 says that when $s = 1$ our loop is a point.

Proposition 1.14. *The group $\mathsf{SU}(2)$ is simply connected.*

Proof. Exercise 8 shows that $\mathsf{SU}(2)$ may be thought of (topologically) as the three-dimensional sphere S^3 sitting inside $\mathbb{R}^4$. It is well known that S^3 is simply connected. □

The condition of simple connectedness is extremely important. One of our most important theorems will be that if G is simply connected, then there is a

natural one-to-one correspondence between the representations of G and the representations of its Lie algebra.

For any path-connected topological space, one can define an object called the **fundamental group**. See Appendix E for more information. A topological space is simply connected if and only if the fundamental group is the trivial group $\{1\}$. I now provide the following tables of fundamental groups, first for compact groups and then for noncompact groups. See Appendix E for the methods of proof. Here, $\mathrm{SO}_e(n;1)$ denotes the identity component of $\mathrm{SO}(n;1)$ (since one defines the fundamental group only for connected groups). In each entry, the result is understood to apply for all $n \geq 1$ unless otherwise stated.

Group	Simply connected?	Fundamental group
$\mathrm{SO}(2)$	no	$\mathbb{Z}$
$\mathrm{SO}(n)$ $(n \geq 3)$	no	$\mathbb{Z}/2$
$\mathrm{U}(n)$	no	$\mathbb{Z}$
$\mathrm{SU}(n)$	yes	$\{1\}$
$\mathrm{Sp}(n)$	yes	$\{1\}$

Group	Simply connected?	Fundamental group
$\mathrm{GL}(n;\mathbb{R})^+$ $(n \geq 2)$	no	same as $\mathrm{SO}(n)$
$\mathrm{GL}(n;\mathbb{C})$	no	$\mathbb{Z}$
$\mathrm{SL}(n;\mathbb{R})$ $(n \geq 2)$	no	same as $\mathrm{SO}(n)$
$\mathrm{SL}(n;\mathbb{C})$	yes	$\{1\}$
$\mathrm{SO}(n;\mathbb{C})$	no	same as $\mathrm{SO}(n)$
$\mathrm{SO}_e(1;1)$	yes	$\{1\}$
$\mathrm{SO}_e(n;1)$ $(n \geq 2)$	no	same as $\mathrm{SO}(n)$
$\mathrm{Sp}(n;\mathbb{R})$	no	$\mathbb{Z}$
$\mathrm{Sp}(n;\mathbb{C})$	yes	$\{1\}$

We conclude this section with a discussion of the case of $\mathrm{SO}(3)$. If v is a unit vector in $\mathbb{R}^3$, let $R_{v,\theta}$ be the element of $\mathrm{SO}(3)$ consisting of a "right-handed" rotation by angle θ in the plane perpendicular to v. Here, right-handed means that if one places the thumb of one's right hand in the v-direction, the rotation is in the direction that one's fingers curl. To say this more mathematically, let $v^\perp$ denote the plane perpendicular to v and let us choose an orthonormal basis (u_1, u_2) for $v^\perp$ in such a way that the basis (u_1, u_2, v) for $\mathbb{R}^3$ has the same orientation as the standard basis (e_1, e_2, e_3). (This means that the linear map taking (u_1, u_2, v) to (e_1, e_2, e_3) has positive determinant.) We then use the basis (u_1, u_2) to identify $v^\perp$ with $\mathbb{R}^2$, and the rotation is then in the counterclockwise direction in $\mathbb{R}^2$.

It is easily seen that $R_{-v,\theta}$ is the same as $R_{v,-\theta}$. It is also not hard to show (Exercise 16) that every element of $\mathrm{SO}(3)$ can be expressed as $R_{v,\theta}$, for some v and θ with $-\pi \leq \theta \leq \pi$. Furthermore, we can arrange that $0 \leq \theta \leq \pi$ by replacing v with $-v$ if necessary.

Now let B denote the closed ball of radius π in $\mathbb{R}^3$ and consider the map $\Phi : B \to \mathrm{SO}(3)$ given by

$$\Phi(u) = R_{\hat{u},\|u\|}, \quad u \neq 0,$$
$$\Phi(0) = I.$$

Here, $\hat{u} = u/\|u\|$ is the unit vector in the u-direction. The map Φ is continuous, even at I, since $R_{v,\theta}$ approaches the identity as θ approaches zero, regardless of how v is behaving. The discussion in the preceding paragraph shows that Φ maps B onto $\mathbb{R}^3$. The map Φ is almost injective, but not quite. Since $R_{v,\pi} = R_{-v,\pi}$, antipodal points on the boundary of B (i.e., pairs of points of the form $(u, -u)$ with $\|u\| = \pi$) map to the same element of $\mathsf{SO}(3)$.

This means that $\mathsf{SO}(3)$ can be identified (homeomorphically) with $B/\sim$, where $\sim$ denotes identification of antipodal points on the boundary. It is known that $B/\sim$ is not simply connected. Specifically, consider the loop in $B/\sim$ that begins at some vector u of length π and goes in a straight line through the origin until it reaches $-u$. (Since u and $-u$ are identified, this *is* a loop in $B/\sim$.) It can be shown that this loop cannot be shrunk continuously to a point in $B/\sim$. This, then, shows that $\mathsf{SO}(3)$ is not simply connected. In fact, $B/\sim$ is homeomorphic to the manifold $\mathbb{RP}^3$ (real projective space of dimension 3) which has fundamental group $\mathbb{Z}/2$.

1.6 Homomorphisms and Isomorphisms

Definition 1.15. *Let G and H be matrix Lie groups. A map Φ from G to H is called a* ***Lie group homomorphism*** *if (1) Φ is a group homomorphism and (2) Φ is continuous. If, in addition, Φ is one-to-one and onto and the inverse map Φ^{-1} is continuous, then Φ is called a* ***Lie group isomorphism.***

The condition that Φ be continuous should be regarded as a technicality, in that it is very difficult to give an example of a group homomorphism between two matrix Lie groups which is not continuous. In fact, if $G = \mathbb{R}$ and $H = \mathbb{C}^*$, then any group homomorphism from G to H which is even measurable (a very weak condition) must be continuous. (See Exercise 17 in Chapter 9 of Rudin (1987).)

Note that the inverse of a Lie group isomorphism is continuous (by definition) and a group homomorphism (by elementary group theory), and thus a Lie group isomorphism. If G and H are matrix Lie groups and there exists a Lie group isomorphism from G to H, then G and H are said to be **isomorphic**, and we write $G \cong H$. Two matrix Lie groups which are isomorphic should be thought of as being essentially the same group.

The simplest interesting example of a Lie group homomorphism is the determinant, which is a homomorphism of $\mathsf{GL}(n;\mathbb{C})$ into $\mathbb{C}^*$. Another simple example is the map $\Phi : \mathbb{R} \to \mathsf{SO}(2)$ given by

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

This map is clearly continuous, and calculation (using standard trigonometric identities) shows that it is a homomorphism. (Compare Exercise 6.)

1.6.1 Example: SU(2) and SO(3)

A very important topic for us will be the relationship between the groups $\mathsf{SU}(2)$ and $\mathsf{SO}(3)$. This example is designed to show that $\mathsf{SU}(2)$ and $\mathsf{SO}(3)$ are almost (but not quite!) isomorphic. Specifically, there exists a Lie group homomorphism Φ which maps $\mathsf{SU}(2)$ onto $\mathsf{SO}(3)$ and which is *two*-to-one. We now describe this map.

Consider the space V of all 2×2 complex matrices which are self-adjoint (i.e., $A^* = A$) and have trace zero. This is a three-dimensional *real* vector space with the following basis:

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \; A_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}; \; A_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We may define an inner product (Section B.6 of Appendix B) on V by the formula

$$\langle A, B \rangle = \frac{1}{2} \operatorname{trace}(AB).$$

(Except for the factor of $\frac{1}{2}$, this is simply the restriction to V of the Hilbert–Schmidt inner product described in Section B.6.) Direct computation shows that $\{A_1, A_2, A_3\}$ is an orthonormal basis for V. Having chosen an orthonormal basis for V, we can identify V with $\mathbb{R}^3$.

Now, suppose that U is an element of $\mathsf{SU}(2)$ and A is an element of V, and consider UAU^{-1}. Then (Section B.5), $\operatorname{trace}(UAU^{-1}) = \operatorname{trace}(A) = 0$ and

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

and so UAU^{-1} is again in V. Furthermore, for a fixed U, the map $A \to UAU^{-1}$ is linear in A. Thus for each $U \in \mathsf{SU}(2)$, we can define a linear map Φ_U of V to itself by the formula

$$\Phi_U(A) = UAU^{-1}.$$

Note that $U_1 U_2 A U_2^{-1} U_1^{-1} = (U_1 U_2) A (U_1 U_2)^{-1}$, and so $\Phi_{U_1 U_2} = \Phi_{U_1} \Phi_{U_2}$. Moreover, given $U \in \mathsf{SU}(2)$ and $A, B \in V$, we have

$$\langle \Phi_U(A), \Phi_U(B) \rangle = \frac{1}{2} \operatorname{trace}(UAU^{-1} UBU^{-1}) = \frac{1}{2} \operatorname{trace}(AB) = \langle A, B \rangle .$$

Thus, Φ_U is an orthogonal transformation of V.

Once we identify V with $\mathbb{R}^3$ (using the above orthonormal basis), then we may think of Φ_U as an element of $\mathsf{O}(3)$. Since $\Phi_{U_1 U_2} = \Phi_{U_1} \Phi_{U_2}$, we see that Φ (i.e., the map $U \to \Phi_U$) is a homomorphism of $\mathsf{SU}(2)$ into $\mathsf{O}(3)$. It is easy to see that Φ is continuous and, thus, a Lie group homomorphism. Recall that every element of $\mathsf{O}(3)$ has determinant ± 1. Now, $\mathsf{SU}(2)$ is connected (Exercise

8), Φ is continuous, and Φ_I is equal to I, which has determinant one. It follows that Φ must actually map $\mathsf{SU}(2)$ into the identity component of $\mathsf{O}(3)$, namely $\mathsf{SO}(3)$.

The map $U \to \Phi_U$ is not one-to-one, since for any $U \in \mathsf{SU}(2)$, $\Phi_U = \Phi_{-U}$. (Observe that if U is in $\mathsf{SU}(2)$, then so is $-U$.) It is possible to show that Φ_U is a two-to-one map of $\mathsf{SU}(2)$ onto $\mathsf{SO}(3)$. (The least obvious part of this assertion is that Φ maps *onto* $\mathsf{SO}(3)$. This will be easy to prove once we have introduced the concept of the Lie algebra and proved Theorem 2.21.) The significance of this homomorphism is that $\mathsf{SO}(3)$ is not simply connected, but $\mathsf{SU}(2)$ is. The map Φ allows us to relate problems on the non-simply-connected group $\mathsf{SO}(3)$ to problems on the simply-connected group $\mathsf{SU}(2)$.

1.7 The Polar Decomposition for $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$

In this section, we consider the polar decompositions for $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$. These decompositions can be used to prove the connectedness of $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$ and to show that the fundamental groups of $\mathsf{SL}(n;\mathbb{R})$ and $\mathsf{SL}(n;\mathbb{C})$ are the same as those of $\mathsf{SO}(n)$ and $\mathsf{SU}(n)$, respectively (Appendix E). These decompositions are supposed to be analogous to the unique decomposition of a nonzero complex number z as $z = up$, with $|u| = 1$ and p real and positive.

A real symmetric matrix P is said to be **positive** if $\langle x, Px\rangle > 0$ for all nonzero vectors $x \in \mathbb{R}^n$. (Symmetric means that $P^{tr} = P$.) Equivalently, a symmetric matrix is positive if all of its eigenvalues are positive. Given a symmetric positive matrix P, there exists an orthogonal matrix R such that

$$P = RDR^{-1},$$

where D is diagonal with positive diagonal entries $\lambda_1, \ldots, \lambda_n$. (If we choose an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors for P, then R is the matrix whose columns are $v_1, \ldots, v_n$.) We can then construct a square root of P as

$$P^{1/2} = RD^{1/2}R^{-1},$$

where $D^{1/2}$ is the diagonal matrix whose (positive) diagonal entries are $\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}$. Then, $P^{1/2}$ is also symmetric and positive. It can be shown that $P^{1/2}$ is the *unique* positive symmetric matrix whose square is P (Exercise 21).

We now prove the following result.

Proposition 1.16. *Given A in $\mathsf{SL}(n;\mathbb{R})$, there exists a unique pair (R, P) such that $R \in \mathsf{SO}(n)$, P is real, symmetric, and positive, and $A = RP$. The matrix P satisfies* $\det P = 1$.

Proof. If there were such a pair, then we would have $A^{tr}A = PR^{-1}RP = P^2$. Now, $A^{tr}A$ is symmetric (check!) and positive, since $\langle x, A^{tr}Ax\rangle = \langle Ax, Ax\rangle > 0$, where $Ax \neq 0$ because A is invertible. Let us then *define* P by

$$P = (A^{tr}A)^{1/2},$$

so that P is real, symmetric, and positive. Since we want $A = RP$, we must set $R = AP^{-1} = A((A^{tr}A)^{1/2})^{-1}$. We check that R is orthogonal:

$$\begin{aligned} RR^{tr} &= A((A^{tr}A)^{1/2})^{-1}((A^{tr}A)^{1/2})^{-1}A^{tr} \\ &= A(A^{tr}A)^{-1}A^{tr} = I. \end{aligned}$$

This shows that R is in $\mathsf{O}(n)$. To check that R is in $\mathsf{SO}(n)$, we note that $1 = \det A = \det R \det P$. Since P is positive, we have $\det P > 0$. This means that we cannot have $\det R = -1$, so we must have $\det R = 1$. It follows that $\det P = 1$ as well.

We have now established the existence of a pair (R, P) with the desired properties. To establish the uniqueness of the pair, we recall that if such a pair exists, then we must have $P^2 = A^{tr}A$. However, we have remarked earlier that a real, positive, symmetric matrix has a unique real, positive, symmetric square root, so P is unique. It follows that $R = AP^{-1}$ is also unique. □

If P is a self-adjoint complex matrix (i.e., $P^* = P$), then we say P is **positive** if $\langle x, Px\rangle > 0$ for all nonzero vectors x in $\mathbb{C}^n$. An argument similar to the one above establishes the following polar decomposition for $\mathsf{SL}(n;\mathbb{C})$.

Proposition 1.17. *Given A in $\mathsf{SL}(n;\mathbb{C})$, there exists a unique pair (U, P) with $U \in \mathsf{SU}(n)$, P self-adjoint and positive, and $A = UP$. The matrix P satisfies $\det P = 1$.*

It is left to the reader to work out the appropriate polar decompositions for the groups $\mathsf{GL}(n;\mathbb{R})$, $\mathsf{GL}(n;\mathbb{R})^+$, and $\mathsf{GL}(n;\mathbb{C})$.

1.8 Lie Groups

As explained in this section and in Appendix C, a Lie group is something that is simultaneously a smooth manifold and a group. As the terminology suggests, every matrix Lie group is a Lie group. (This is not at all obvious from the definition of a matrix Lie group, but it is true nevertheless, as we will prove in the next chapter.) The reverse is not true: Not every Lie group is isomorphic to a matrix Lie group. Nevertheless, I have restricted attention in this book to matrix Lie groups for several reasons. First, not everyone who wants to learn about Lie groups is familiar with manifold theory. Second, even for someone familiar with manifolds, the definitions of the Lie algebra and exponential mapping for a general Lie group are substantially more complicated and abstract than in the matrix case. Third, most of the interesting examples of Lie groups are matrix Lie groups. Fourth, the results we will prove for matrix Lie groups (e.g., about the relationship between Lie group homomorphisms and Lie algebra homomorphisms) continue to hold for general Lie groups. Indeed,

the proofs of these results are much the same as in the general case, except that one can get started more quickly in the matrix case. Although in the long run the manifold approach to Lie groups is unquestionably the right one, the matrix approach allows one to get into the meat of Lie group theory with minimal preparation.

This section gives a very brief account of the manifold approach to Lie groups. Appendix C gives more information, and complete accounts can be found in standard textbooks such as those by Bröcker and tom Dieck (1985), Varadarajan (1974), and Warner (1983). Appendix C gives two examples of Lie groups that cannot be represented as matrix Lie groups and also discusses two important constructions (covering groups and quotient groups) which can be performed for general Lie groups but not for matrix Lie groups.

Definition 1.18. *A **Lie group** is a differentiable manifold G which is also a group and such that the group product*

$$G \times G \to G$$

and the inverse map $g \to g^{-1}$ are differentiable.

A manifold is an object that looks locally like a piece of $\mathbb{R}^n$. An example would be a torus, the two-dimensional surface of a "doughnut" in $\mathbb{R}^3$, which looks locally (but not globally) like $\mathbb{R}^2$. For a precise definition, see Appendix C.

Example. As an example, let

$$G = \mathbb{R} \times \mathbb{R} \times S^1 = \{(x, y, u) | x \in \mathbb{R}, y \in \mathbb{R}, u \in S^1 \subset \mathbb{C}\}$$

and define the group product $G \times G \to G$ by

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

Let us first check that this operation makes G into a group. It is not obvious but easily checked that this operation is associative; the product of three elements with either grouping is

$$(x_1 + x_2 + x_3, y_1 + y_2 + y_3, e^{i(x_1y_2 + x_1y_3 + x_2y_3)} u_1 u_2 u_3).$$

There is an identity element in G, namely $e = (0, 0, 1)$ and each element (x, y, u) has an inverse given by $(-x, -y, e^{ixy} u^{-1})$.

Thus, G is, in fact, a group. Furthermore, both the group product and the map that sends each element to its inverse are clearly smooth, and so G is a Lie group. Note that there is nothing about matrices in the way we have defined G; that is, G is not given to us as a matrix group. We may still ask whether G is isomorphic to some matrix Lie group, but even this is not true. As shown in Appendix C, there is no continuous, injective homomorphism of G into any $\mathsf{GL}(n; \mathbb{C})$. Thus, this example shows that not every Lie group is

a matrix Lie group. Nevertheless, G is closely related to a matrix Lie group, namely the Heisenberg group. The reader is invited to try to work out what the relationship is before consulting the appendix.

Now let us think about the question of whether every matrix Lie group is a Lie group. This is certainly not obvious, since nothing in our definition of a matrix Lie group says anything about its being a manifold. (Indeed, the whole point of considering matrix Lie groups is that one can define and study them without having to go through manifold theory first!) Nevertheless, it is true that every matrix Lie group is a Lie group, and it would be a particularly misleading choice of terminology if this were not so.

Theorem 1.19. *Every matrix Lie group is a smooth embedded submanifold of $M_n(\mathbb{C})$ and is thus a Lie group.*

The proof of this theorem makes use of the notion of the Lie algebra of a matrix Lie group and is given in Chapter 2. Let us think first about the case of $\mathsf{GL}(n;\mathbb{C})$. This is an open subset of the space $M_n(\mathbb{C})$ and thus a manifold of (real) dimension $2n^2$. The matrix product is certainly a smooth map of $M_n(\mathbb{C})$ to itself, and the map that sends a matrix to its inverse is smooth on $\mathsf{GL}(n;\mathbb{C})$, by the formula for the inverse in terms of the classical adjoint. Thus, $\mathsf{GL}(n;\mathbb{C})$ itself is a Lie group. If $G \subset \mathsf{GL}(n;\mathbb{C})$ is a matrix Lie group, then we will prove in Chapter 2 that G is a smooth embedded submanifold of $\mathsf{GL}(n;\mathbb{C})$. (See Corollary 2.33 to Theorem 2.27.) The matrix product and inverse will be restrictions of smooth maps to smooth submanifolds and, thus, will be smooth. This will show, then, that G is also a Lie group.

It is customary to call a map Φ between two Lie groups a Lie group homomorphism if Φ is a group homomorphism and Φ is *smooth*, whereas we have (in Definition 1.15) required only that Φ be continuous. However, the following proposition shows that our definition is equivalent to the more standard one.

Proposition 1.20. *Let G and H be Lie groups and let Φ be a group homomorphism from G to H. If Φ is continuous, it is also smooth.*

Thus, group homomorphisms from G to H come in only two varieties: the very bad ones (discontinuous) and the very good ones (smooth). There simply are not any intermediate ones. (See, for example, Exercise 19.) We will prove this in the next chapter (for the case of matrix Lie groups). See Corollary 2.34 to Theorem 2.27.

In light of Theorem 1.19, every matrix Lie group is a (smooth) manifold. As such, a matrix Lie group is automatically *locally* path-connected. It follows that a matrix Lie group is path-connected if and only if it is connected. (See the remarks following Definition 1.7.)

1.9 Exercises

1. Let a be an irrational real number and let G be the following subgroup of $\mathsf{GL}(2;\mathbb{C})$:
$$G = \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{ita} \end{pmatrix} \middle| \, t \in \mathbb{R} \right\}.$$
Show that
$$\overline{G} = \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{is} \end{pmatrix} \middle| \, t, s \in \mathbb{R} \right\},$$
where $\overline{G}$ denotes the closure of the set G inside the space of 2×2 matrices. Assume the following result: The set of numbers of the form $e^{2\pi i n a}$, $n \in \mathbb{Z}$, is dense in S^1.
Note: The group $\overline{G}$ can be thought of as the torus $S^1 \times S^1$, which, in turn, can be thought of as $[0, 2\pi] \times [0, 2\pi]$, with the ends of the intervals identified. The set $G \subset [0, 2\pi] \times [0, 2\pi]$ is called an **irrational line**. Drawing a picture of this set should make it plausible that G is dense in $[0, 2\pi] \times [0, 2\pi]$.
2. *Orthogonal groups.* Let $\langle \cdot, \cdot \rangle$ denote the standard inner product on $\mathbb{R}^n$: $\langle x, y \rangle = \sum_k x_k y_k$. Show that a matrix A preserves this inner product if and only if the column vectors of A are orthonormal.
Show that for any $n \times n$ real matrix B,
$$\langle Bx, y \rangle = \langle x, B^{tr} y \rangle,$$
where $(B^{tr})_{kl} = B_{lk}$. Using this, show that a matrix A preserves the inner product on $\mathbb{R}^n$ if and only if $A^{tr} A = I$.
Note: A similar analysis applies to the complex orthogonal groups $\mathsf{O}(n;\mathbb{C})$ and $\mathsf{SO}(n;\mathbb{C})$.
3. *Unitary groups.* Let $\langle \cdot, \cdot \rangle$ denote the standard inner product on $\mathbb{C}^n$: $\langle x, y \rangle = \sum_k \overline{x_k} y_k$. Following Exercise 2, show that $\langle Ax, Ay \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{C}^n$ if and only if $A^* A = I$ and that this holds if and only if the columns of A are orthonormal. Here, $(A^*)_{kl} = \overline{A_{lk}}$.
4. *Generalized orthogonal groups.* Let $[\cdot, \cdot]_{n,k}$ be the symmetric bilinear form on $\mathbb{R}^{n+k}$ defined in (1.1). Let g be the $(n+k) \times (n+k)$ diagonal matrix with first n diagonal entries equal to one and last k diagonal entries equal to minus one:
$$g = \begin{pmatrix} I_n & 0 \\ 0 & -I_k \end{pmatrix}.$$
Show that for all $x, y \in \mathbb{R}^{n+k}$,
$$[x, y]_{n,k} = \langle x, gy \rangle.$$
Show that a $(n+k) \times (n+k)$ real matrix A is in $\mathsf{O}(n;k)$ if and only if $A^{tr} g A = g$. Show that $\mathsf{O}(n;k)$ and $\mathsf{SO}(n;k)$ are subgroups of $\mathsf{GL}(n+k;\mathbb{R})$ and are matrix Lie groups.

5. *Symplectic groups.* Let $B[x,y]$ be the skew-symmetric bilinear form on $\mathbb{R}^{2n}$ given by $B[x,y] = \sum_{k=1}^{n}(x_k y_{n+k} - x_{n+k}y_k)$. Let J be the $2n \times 2n$ matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Show that for all $x, y \in \mathbb{R}^{2n}$,

$$B[x,y] = \langle x, Jy \rangle$$

Show that a $2n \times 2n$ matrix A is in $\mathsf{Sp}(n;\mathbb{R})$ if and only if $A^{tr}JA = J$. Show that $\mathsf{Sp}(n;\mathbb{R})$ is a subgroup of $\mathsf{GL}(2n;\mathbb{R})$ and a matrix Lie group.
Note: A similar analysis applies to $\mathsf{Sp}(n;\mathbb{C})$.

6. *The groups* $\mathsf{O}(2)$ *and* $\mathsf{SO}(2)$. Show that the matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

is in $\mathsf{SO}(2)$ and that

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

Show that every element A of $\mathsf{O}(2)$ is of one of the two forms:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

(Note that if A is of the first form, then $\det A = 1$, and if A is of the second form, then $\det A = -1$.)
Hint: Recall that for A to be in $\mathsf{O}(2)$, the columns of A must be orthonormal.

7. *The groups* $\mathsf{O}(1;1)$ *and* $\mathsf{SO}(1;1)$. Show that the matrix

$$\begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$$

is in $\mathsf{SO}(1;1)$ and that

$$\begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}\begin{pmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{pmatrix} = \begin{pmatrix} \cosh(t+s) & \sinh(t+s) \\ \sinh(t+s) & \cosh(t+s) \end{pmatrix}.$$

Show that every element of $\mathsf{O}(1;1)$ can be written in one of the four forms:

$$\begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}; \quad \begin{pmatrix} -\cosh t & \sinh t \\ \sinh t & -\cosh t \end{pmatrix};$$
$$\begin{pmatrix} \cosh t & -\sinh t \\ \sinh t & -\cosh t \end{pmatrix}; \quad \begin{pmatrix} -\cosh t & -\sinh t \\ \sinh t & \cosh t \end{pmatrix}.$$

(Note that since $\cosh t$ is always positive, there is no overlap among the four cases. Note also that matrices of the first two forms have determinant one and matrices of the last two forms have determinant minus one.)
Hint: Use condition (1.2).

8. *The group* $\mathsf{SU}(2)$. Show that if α and β are arbitrary complex numbers satisfying $|\alpha|^2 + |\beta|^2 = 1$, then the matrix

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

is in $\mathsf{SU}(2)$. Show that every $A \in \mathsf{SU}(2)$ can be expressed in this form for a unique pair (α, β) satisfying $|\alpha|^2 + |\beta|^2 = 1$. (Thus, $\mathsf{SU}(2)$ can be thought of as the three-dimensional sphere S^3 sitting inside $\mathbb{C}^2 = \mathbb{R}^4$. In particular, this shows that $\mathsf{SU}(2)$ is simply connected.)

9. *The groups* $\mathsf{Sp}(1;\mathbb{R})$, $\mathsf{Sp}(1;\mathbb{C})$, *and* $\mathsf{Sp}(1)$. Show that $\mathsf{Sp}(1;\mathbb{R}) = \mathsf{SL}(2;\mathbb{R})$, $\mathsf{Sp}(1;\mathbb{C}) = \mathsf{SL}(2;\mathbb{C})$, and $\mathsf{Sp}(1) = \mathsf{SU}(2)$.

10. *The Heisenberg group.* Determine the center $Z(H)$ of the Heisenberg group H. Show that the quotient group $H/Z(H)$ is abelian.

11. A subset E of a matrix Lie group G is called **discrete** if for each A in E there is a neighborhood U of A in G such that U contains no point in E except for A. Suppose that G is a connected matrix Lie group and N is a discrete normal subgroup of G. Show that N is contained in the center of G.

12. This problem gives an alternative proof of Proposition 1.9, namely that $\mathsf{GL}(n;\mathbb{C})$ is connected. Suppose A and B are invertible $n \times n$ matrices. Show that there are only finitely many complex numbers λ for which $\det(\lambda A + (1-\lambda)B) = 0$. Show that there exists a continuous path $A(t)$ of the form $A(t) = \lambda(t)A + (1-\lambda(t))B$ connecting A to B and such that $A(t)$ lies in $\mathsf{GL}(n;\mathbb{C})$. Here, $\lambda(t)$ is a continuous path in the plane with $\lambda(0) = 0$ and $\lambda(1) = 1$.

13. *Connectedness of* $\mathsf{SO}(n)$. Show that $\mathsf{SO}(n)$ is connected, using the following outline.
For the case $n = 1$, there is nothing to show, since a 1×1 matrix with determinant one must be $[1]$. Assume, then, that $n \geq 2$. Let e_1 denote the unit vector with entries $1, 0, \ldots, 0$ in $\mathbb{R}^n$. Given any unit vector $v \in \mathbb{R}^n$, show that there exists a continuous path $R(t)$ in $\mathsf{SO}(n)$ with $R(0) = I$ and $R(1)v = e_1$. (Thus, any unit vector can be "continuously rotated" to e_1.)
Now, show that any element R of $\mathsf{SO}(n)$ can be connected to a block-diagonal matrix of the form

$$\begin{pmatrix} 1 & \\ & R_1 \end{pmatrix}$$

with $R_1 \in \mathsf{SO}(n-1)$ and proceed by induction.

14. *The connectedness of* $\mathsf{SL}(n;\mathbb{R})$. Using the polar decomposition of $\mathsf{SL}(n;\mathbb{R})$ (Proposition 1.16) and the connectedness of $\mathsf{SO}(n)$ (Exercise 13), show that $\mathsf{SL}(n;\mathbb{R})$ is connected.
Hint: Recall that if P is a real, symmetric matrix, then there exists a real, orthogonal matrix R_1 such that $P = R_1 D R_1^{-1}$, where D is diagonal.
15. *The connectedness of* $\mathsf{GL}(n;\mathbb{R})^+$. Using the connectedness of $\mathsf{SL}(n;\mathbb{R})$ (Exercise 14) show that $\mathsf{GL}(n;\mathbb{R})^+$ is connected.
16. If R is an element of $\mathsf{SO}(3)$, show that R must have an eigenvector with eigenvalue 1.
Hint: Since $\mathsf{SO}(3) \subset \mathsf{SU}(3)$, every (real or complex) eigenvalue of R must have absolute value 1.
17. Show that the set of translations is a normal subgroup of the Euclidean group $\mathsf{E}(n)$. Show that the quotient group $\mathsf{E}(n)/(\text{translations})$ is isomorphic to $\mathsf{O}(n)$. (Assume Proposition 1.5.)
18. Let a be an irrational real number. Show that the set of numbers of the form $e^{2\pi i n a}$, $n \in \mathbb{Z}$, is dense in S^1. (See Problem 1.)
19. Show that every continuous homomorphism Φ from $\mathbb{R}$ to S^1 is of the form $\Phi(x) = e^{iax}$ for some $a \in \mathbb{R}$. (This shows in particular that every such homomorphism is smooth.)
20. Suppose $G \subset \mathsf{GL}(n_1;\mathbb{C})$ and $H \subset \mathsf{GL}(n_2;\mathbb{C})$ are matrix Lie groups and that $\Phi : G \to H$ is a Lie group homomorphism. Then, the image of G under Φ is a subgroup of H and thus of $\mathsf{GL}(n_2;\mathbb{C})$. Is the image of G under Φ necessarily a matrix Lie group? Prove or give a counter-example.
21. Suppose P is a real, positive, symmetric matrix with determinant one. Show that there is a unique real, positive, symmetric matrix Q whose square is P.
Hint: The existence of Q was discussed in Section 1.7. To prove uniqueness, consider two real, positive, symmetric square roots Q_1 and Q_2 of P and show that the eigenspaces of both Q_1 and Q_2 coincide with the eigenspaces of P.

2

Lie Algebras and the Exponential Mapping

2.1 The Matrix Exponential

The exponential of a matrix plays a crucial role in the theory of Lie groups. The exponential enters into the definition of the Lie algebra of a matrix Lie group (Section 2.5) and is the mechanism for passing information from the Lie algebra to the Lie group. Since many computations are done much more easily at the level of the Lie algebra, the exponential is indispensable in studying (matrix) Lie groups.

Let X be an $n \times n$ real or complex matrix. We wish to define the exponential of X, denoted e^X or $\exp X$, by the usual power series

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}. \tag{2.1}$$

We will follow the convention of using letters such as X and Y for the variable in the matrix exponential.

Proposition 2.1. *For any $n \times n$ real or complex matrix X, the series (2.1) converges. The matrix exponential e^X is a continuous function of X.*

Before proving this, let us review some elementary analysis. Recall that the norm of a vector $x = (x_1, \ldots, x_n)$ in $\mathbb{C}^n$ is defined to be

$$\|x\| = \sqrt{\langle x, x \rangle} = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{1/2}.$$

We now define the norm of a matrix by thinking of the space $M_n(\mathbb{C})$ of all $n \times n$ matrices as $\mathbb{C}^{n^2}$. This means that we define

$$\|X\| = \left(\sum_{k,l=1}^{n} |X_{kl}|^2 \right)^{1/2}. \tag{2.2}$$

This norm satisfies the inequalities

$$\|X+Y\| \leq \|X\| + \|Y\|, \tag{2.3}$$

$$\|XY\| \leq \|X\| \, \|Y\| \tag{2.4}$$

for all $X, Y \in M_n(\mathbb{C})$. The first of these inequalities is the triangle inequality and is a standard result from elementary analysis. The second of these inequalities follows from the Schwarz inequality (Exercise 1). If X_m is a sequence of matrices, then it is easy to see that X_m converges to a matrix X in the sense of Definition 1.3 if and only if $\|X_m - X\| \to 0$ as $m \to \infty$.

The norm (2.2) is called the **Hilbert–Schmidt** norm. There is another commonly used norm on the space of matrices, called the **operator norm**, whose definition is not relevant to us. It is easily shown that convergence in the Hilbert–Schmidt norm is equivalent to convergence in the operator norm. (This is true because we work with linear operators on the *finite-dimensional* space $\mathbb{C}^n$.) Furthermore, the operator norm also satisfies (2.3) and (2.4). Thus, it matters little whether we use the operator norm or the Hilbert–Schmidt norm.

A sequence X_m of matrices is said to be a **Cauchy sequence** if

$$\|X_m - X_l\| \to 0$$

as $m, l \to \infty$. Thinking of the space $M_n(\mathbb{C})$ of matrices as $\mathbb{C}^{n^2}$ and using a standard result from analysis, we have the following.

Proposition 2.2. *If X_m is a Cauchy sequence in $M_n(\mathbb{C})$, then there exists a unique matrix X such that X_m converges to X.*

That is, every Cauchy sequence in $M_n(\mathbb{C})$ converges.

Now, consider an infinite series whose terms are matrices:

$$X_0 + X_1 + X_2 + \cdots. \tag{2.5}$$

If

$$\sum_{m=0}^{\infty} \|X_m\| < \infty,$$

then the series (2.5) is said to **converge absolutely**. If a series converges absolutely, then it is not hard to show that the partial sums of the series form a Cauchy sequence, and, hence, by Proposition 2.2, the series converges. That is, any series which converges absolutely also converges. (The converse is not true; a series of matrices can converge without converging absolutely.)

We now turn to the proof of Proposition 2.1.

Proof. In light of (2.4), we see that

$$\|X^m\| \leq \|X\|^m,$$

and, hence,

$$\sum_{m=0}^{\infty} \left\| \frac{X^m}{m!} \right\| \leq \sum_{m=0}^{\infty} \frac{\|X\|^m}{m!} = e^{\|X\|} < \infty.$$

Thus, the series (2.1) converges absolutely, and so it converges.

To show continuity, note that since X^m is a continuous function of X, the partial sums of (2.1) are continuous. However, it is easy to see that (2.1) converges uniformly on each set of the form $\{\|X\| \leq R\}$, and so the sum is, again, continuous. □

We now list some elementary properties of the matrix exponential.

Proposition 2.3. *Let X and Y be arbitrary $n \times n$ matrices. Then, we have the following:*

1. $e^0 = I$.
2. $\left(e^X\right)^* = e^{X^*}$.
3. e^X *is invertible and* $\left(e^X\right)^{-1} = e^{-X}$.
4. $e^{(\alpha+\beta)X} = e^{\alpha X} e^{\beta X}$ *for all* α *and* β *in* $\mathbb{C}$.
5. *If* $XY = YX$, *then* $e^{X+Y} = e^X e^Y = e^Y e^X$.
6. *If* C *is invertible, then* $e^{CXC^{-1}} = Ce^X C^{-1}$.
7. $\left\|e^X\right\| \leq e^{\|X\|}$.

It is *not* true in general that $e^{X+Y} = e^X e^Y$, although, by Point 4, it is true if X and Y commute. This is a crucial point, which we will consider in detail later. (See the Lie product formula in Section 2.4 and the Baker–Campbell–Hausdorff formula in Chapter 3.)

Proof. Point 1 is obvious and Point 2 follows from taking term-by-term adjoints of the series for e^X. Points 3 and 4 are special cases of Point 5. To verify Point 5, we simply multiply the power series term by term. (It is left to the reader to verify that this is legal.) Thus,

$$e^X e^Y = \left(I + X + \frac{X^2}{2!} + \cdots\right)\left(I + Y + \frac{Y^2}{2!} + \cdots\right).$$

Multiplying this out and collecting terms where the power of X plus the power of Y equals m, we get

$$e^X e^Y = \sum_{m=0}^{\infty} \sum_{k=0}^{m} \frac{X^k}{k!} \frac{Y^{m-k}}{(m-k)!} = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} X^k Y^{m-k}. \tag{2.6}$$

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

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

and, thus, (2.6) becomes

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

To prove Point 6, simply note that

$$\left(CXC^{-1}\right)^m = CX^m C^{-1}$$

and, thus, the two sides of Point 6 are equal term by term.

Point 7 is evident from the proof of Proposition 2.1. □

Proposition 2.4. *Let X be a $n \times n$ complex matrix. Then, e^{tX} is a smooth curve in $M_n(\mathbb{C})$ and*

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

In particular,

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

Proof. Differentiate the power series for e^{tX} term by term. (This is permitted because, for each i and j, $\left(e^{tX}\right)_{ij}$ is given by a convergent power series in t, and it is a standard theorem that one can differentiate power series term by term.) □

2.2 Computing the Exponential of a Matrix

We consider here methods for exponentiating general matrices. A special method for exponentiating 2×2 matrices is described in Exercises 6 and 7.

2.2.1 Case 1: X is diagonalizable

Suppose that X is an $n \times n$ real or complex matrix and that X is diagonalizable over $\mathbb{C}$; that is, there exists an invertible complex matrix C such that $X = CDC^{-1}$, with

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

It is easily verified that e^D is the diagonal matrix with eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_n}$, and so in light of Proposition 2.3, we have

$$e^X = C \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix} C^{-1}.$$

Thus, if we can explicitly diagonalize X, we can explicitly compute e^X. Note that if X is real, then although C may be complex and the λ_k's may be complex, e^X must come out to be real, since each term in the series (2.1) is real.

For example, take

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

Then, the eigenvectors of X are $\begin{pmatrix} 1 \\ i \end{pmatrix}$ and $\begin{pmatrix} i \\ 1 \end{pmatrix}$, with eigenvalues $-ia$ and ia, respectively. Thus, the invertible matrix

$$C = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

maps the basis vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to the eigenvectors of X, and so (check) $C^{-1}XC$ is a diagonal matrix D. Thus, $X = CDC^{-1}$ and

$$\begin{aligned} e^X &= \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} e^{-ia} & 0 \\ 0 & e^{ia} \end{pmatrix} \begin{pmatrix} 1/2 & -i/2 \\ -i/2 & 1/2 \end{pmatrix} \\ &= \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}. \end{aligned} \tag{2.7}$$

Note that explicitly if X (and hence a) is real, then e^X is real. See Exercise 6 for an alternative method of calculation.

2.2.2 Case 2: X is nilpotent

An $n \times n$ matrix X is said to be **nilpotent** if $X^m = 0$ for some positive integer m. Of course, if $X^m = 0$, then $X^l = 0$ for all $l > m$. In this case, the series (2.1), which defines e^X, terminates after the first m terms, and so can be computed explicitly.

For example, let us compute e^X, where

$$X = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}.$$

Note that

$$X^2 = \begin{pmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and that $X^3 = 0$. Thus,

$$e^X = \begin{pmatrix} 1 & a & b + \frac{1}{2}ac \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

2.2.3 Case 3: X arbitrary

A general matrix X may be neither nilpotent nor diagonalizable. However, by Theorem B.6, every matrix X can be written (uniquely) in the form $X = S + N$, with S diagonalizable, N nilpotent, and $SN = NS$. Then, since N and S commute,

$$e^X = e^{S+N} = e^S e^N$$

and e^S and e^N can be computed as in the two previous subsections.

For example, take

$$X = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}.$$

Then,

$$X = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}.$$

The two terms clearly commute (since the first one is a multiple of the identity), and, so,

$$e^X = \begin{pmatrix} e^a & 0 \\ 0 & e^a \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^a & e^a b \\ 0 & e^a \end{pmatrix}.$$

2.3 The Matrix Logarithm

We wish to define a matrix logarithm, which should be an inverse function (to the extent possible) to the matrix exponential. Let us recall the situation for the logarithm of complex numbers, in order to see what is reasonable to expect in the matrix case. Since e^z is never zero, only nonzero numbers can have a logarithm. Every nonzero complex number can be written as e^z for some z, but the z is not unique. There is no continuous way to define the logarithm on the set of all nonzero complex numbers. The situation for matrices is similar. For any $X \in M_n(\mathbb{C})$, e^X is invertible; therefore, only invertible matrices can possibly have a logarithm. We will see (Theorem 2.9) that every invertible matrix can be written as e^X, for some $X \in M_n(\mathbb{C})$. However, the X is not unique and there is no continuous way to define the matrix logarithm on the set of all invertible matrices.

The simplest way to define the matrix logarithm is by a power series. We recall how this works in the complex case.

Lemma 2.5. *The function*

$$\log z = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m} \tag{2.8}$$

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

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

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

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

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

Proof. The usual logarithm for real, positive numbers satisfies

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

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

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

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

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

This series has radius of convergence 1 and defines a complex analytic function on the set $\{|z-1| < 1\}$, which coincides with the usual logarithm for real z in the interval $(0, 2)$. Now, $\exp(\log z) = z$ for $z \in (0, 2)$, and by analyticity, this identity continues to hold on the whole set $\{|z-1| < 1\}$. (That is to say, the functions $z \to \exp(\log z)$ and $z \to z$ are both complex analytic functions and they agree on the interval $(0, 2)$; therefore they must agree on the whole disk $\{|z-1| < 1\}$.)

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

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

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

Definition 2.6. *For any $n \times n$ matrix A, define $\log A$ by*

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m} \tag{2.9}$$

whenever the series converges.

Since the complex-valued series (2.8) has radius of convergence 1 and since $\|(A-I)^m\| \leq \|A-I\|^m$, the matrix-valued series (2.9) will converge

if $\|A - I\| < 1$. However, in contrast to the complex-valued case, the series (2.9) may converge even if $\|A - I\| > 1$, since $\|(A - I)^m\|$ may be strictly smaller than $\|A - I\|^m$. For example, if $A - I$ is nilpotent, then (2.9) terminates and, thus, converges. (See Exercise 8.) Nevertheless, we will mostly content ourselves with considering the case $\|A - I\| < 1$.

Theorem 2.7. *The function*

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

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

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

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

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

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

Proof. Since $\|(A - I)^m\| \leq \|(A - I)\|^m$ and since the series (2.8) has radius of convergence 1, the series (2.9) converges absolutely for all A with $\|A - I\| < 1$. The proof of continuity is essentially the same as for the exponential.

We will now show that $\exp(\log A) = A$ for all A with $\|A - I\| < 1$. We do this by considering two cases.

Case 1: A is diagonalizable.

Suppose that $A = CDC^{-1}$, with D diagonal. Then, $A - I = CDC^{-1} - I = C(D - I)C^{-1}$. It follows that $(A - I)^m$ is of the form

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

where $z_1, \ldots, z_n$ are the eigenvalues of A.

Now, if $\|A - I\| < 1$, then it is not hard to show (Exercise 2) that each eigenvalue z_k of A must satisfy $|z_k - 1| < 1$. Thus,

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

and by Lemma 2.5,

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

Case 2: A is not diagonalizable.

If A is not diagonalizable, then, using Theorem B.7, it is not difficult to construct a sequence A_m of diagonalizable matrices with $A_m \to A$. (See Exercise 5.) If $\|A - I\| < 1$, then $\|A_m - I\| < 1$ for all sufficiently large m. By Case 1, $\exp(\log A_m) = A_m$, and, so, by the continuity of exp and log, $\exp(\log A) = A$.

Thus, we have shown that $\exp(\log A) = A$ for all A with $\|A - I\| < 1$. Now, the same argument as in the complex case shows that if $\|X\| < \log 2$, then $\left\|e^X - I\right\| < 1$. The same two-case argument shows that $\log(\exp X) = X$ for all such X. □

Proposition 2.8. *There exists a constant c such that for all $n \times n$ matrices B with $\|B\| < \frac{1}{2}$,*

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

Proof. Note that

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

so that

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

This is what we want. (It is easily verified that the sum in the last expression is convergent.) □

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

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

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

We conclude this section with a result that, although we will not use it elsewhere, is worth recording. The proof is sketched in Exercises 8 and 9.

Theorem 2.9. *Every invertible $n \times n$ matrix can be expressed as e^X for some $X \in M_n(\mathbb{C})$.*

2.4 Further Properties of the Matrix Exponential

In this section, we give several additional results involving the exponential of a matrix that will be important in our study of Lie algebras.

Theorem 2.10 (Lie Product Formula). *Let X and Y be $n \times n$ complex matrices. Then,*

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

This theorem has a big brother, called the Trotter product formula, which gives the same result in the case where X and Y are suitable unbounded operators on an infinite-dimensional Hilbert space. The Trotter product formula is described, for example, in Reed and Simon (1980), Section VIII.8.

Proof. If we multiply the power series for $e^{\frac{X}{m}}$ and $e^{\frac{Y}{m}}$, all but three of the terms will involve $1/m^2$ or higher powers of $1/m$. Thus,

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

Now, since $e^{\frac{X}{m}} e^{\frac{Y}{m}} \to I$ as $m \to \infty$, $e^{\frac{X}{m}} e^{\frac{Y}{m}}$ is in the domain of the logarithm for all sufficiently large m. By Proposition 2.8,

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

Exponentiating the logarithm then gives

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

and, therefore ,

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

Thus, by the continuity of the exponential, we conclude that

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

which is the Lie product formula. □

Recall (Section B.5) that the trace of a matrix is defined as the sum of its diagonal entries and that similar matrices have the same trace.

Theorem 2.11. *For any $X \in M_n(\mathbb{C})$, we have*

$$\det\left(e^X\right) = e^{\mathrm{trace}(X)}.$$

Proof. There are three cases, as in Section 2.2.

Case 1: X is diagonalizable. Suppose there is a complex invertible matrix C such that

$$X = C \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} C^{-1}.$$

Then,

$$e^X = C \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix} C^{-1}.$$

Thus, $\text{trace}(X) = \sum \lambda_i$ and $\det(e^X) = \prod e^{\lambda_i} = e^{\sum \lambda_i}$.

Case 2: X is nilpotent. If X is nilpotent, then by Theorem B.7, there is an invertible matrix C such that

$$X = C \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} C^{-1}.$$

In that case (it is easy to see), e^X will be upper triangular, with ones on the diagonal:

$$e^X = C \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} C^{-1}.$$

Thus, if X is nilpotent, $\text{trace}(X) = 0$ and $\det(e^X) = 1$.

Case 3: X is arbitrary. As pointed out in Section 2.2, every matrix X can be written as the sum of two commuting matrices S and N, with S diagonalizable (over $\mathbb{C}$) and N nilpotent. Since S and N commute, $e^X = e^S e^N$. So, by the two previous cases,

$$\det\left(e^X\right) = \det\left(e^S\right) \det\left(e^N\right) = e^{\text{trace}(S)} e^{\text{trace}(N)} = e^{\text{trace}(X)},$$

which is what we want. (Note that $\text{trace}(N) = 0$ and $\text{trace}(S) = \text{trace}(X)$.)

□

Definition 2.12. *A function $A : \mathbb{R} \to \mathsf{GL}(n;\mathbb{C})$ is called a **one-parameter subgroup** of $\mathsf{GL}(n;\mathbb{C})$ if*

1. *A is continuous,*
2. *$A(0) = I$,*
3. *$A(t+s) = A(t)A(s)$ for all $t, s \in \mathbb{R}$.*

Theorem 2.13 (One-Parameter Subgroups). *If A is a one-parameter subgroup of $\mathsf{GL}(n;\mathbb{C})$, then there exists a unique $n \times n$ complex matrix X such that*

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

By taking $n = 1$, and noting that $\mathsf{GL}(1;\mathbb{C}) \cong \mathbb{C}^*$, this theorem provides a method of solving Exercise 19 in Chapter 1.

Proof. The uniqueness is immediate, since if there is such an X, then $X = \frac{d}{dt}A(t)\big|_{t=0}$. So, we need only worry about existence.

Let B_ε be the open ball of radius ε about zero in $M_n(\mathbb{C})$; that is, $B_\varepsilon = \{X \in M_n(\mathbb{C}) | \, \|X\| < \varepsilon\}$. Assume that $\varepsilon < \log 2$. Then, we have shown that "exp" takes B_ε injectively into $M_n(\mathbb{C})$, with continuous inverse that we denote "log." Now, let $U = \exp(B_{\varepsilon/2})$, which is an open set in $\mathsf{GL}(n;\mathbb{C})$.

Lemma 2.14. *Every $g \in U$ has a unique square root h in U, given by $h = \exp(\frac{1}{2}\log g)$.*

Proof. Let $X = \log g$. Then, $h = \exp(X/2)$ is a square root of g, since $h^2 = \exp(X) = g$. Suppose $h' \in U$ satisfies $(h')^2 = g$. Let $Y = \log h'$; then, $\exp(Y) = h'$ and $\exp(2Y) = (h')^2 = g = \exp(X)$. We have that $Y \in B_{\varepsilon/2}$ and, thus, $2Y \in B_\varepsilon$, and also that $X \in B_{\varepsilon/2} \subset B_\varepsilon$. Since exp is injective on B_ε and $\exp(2Y) = \exp(X)$, we must have $2Y = X$. Thus, $h' = \exp(Y) = \exp(X/2) = h$. This shows the uniqueness of the square root in U. □

Returning to the proof of Theorem 2.13, the continuity of A guarantees that there exists $t_0 > 0$ such that $A(t) \in U$ for all t with $|t| \leq t_0$. Then, let $X = \frac{1}{t_0}\log(A(t_0))$, so that $t_0 X = \log(A(t_0))$. Then, $t_0 X \in B_{\varepsilon/2}$ and $A(t_0) = \exp(t_0 X)$. Then, $A(t_0/2)$ is in U and $A(t_0/2)^2 = A(t_0)$. By the lemma, $A(t_0)$ has a *unique* square root in U, and that unique square root is $\exp(t_0 X/2)$. So, we must have $A(t_0/2) = \exp(t_0 X/2)$.

Applying this argument repeatedly, we conclude that

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

for all positive integers k. Then, for any integer m, we have $A(mt_0/2^k) = A(t_0/2^k)^m = \exp(mt_0 X/2^k)$. This means that $A(t) = \exp(tX)$ for all real numbers t of the form $t = mt_0/2^k$, and the set of such t's is dense in $\mathbb{R}$. Since both $\exp(tX)$ and $A(t)$ are continuous, it follows that $A(t) = \exp(tX)$ for all real numbers t. □

2.5 The Lie Algebra of a Matrix Lie Group

The Lie algebra is an indispensable tool in studying matrix Lie groups. On the one hand, Lie algebras are simpler than matrix Lie groups, because (as we will see) the Lie algebra is a linear space. Thus, we can understand much about Lie algebras just by doing linear algebra. On the other hand, the Lie algebra of a matrix Lie group contains much information about that group. (See, for example, Theorem 2.27 in Section 2.7, and the Baker–Campbell–Hausdorff Formula (Chapter 3).) Thus, many questions about matrix Lie groups can be answered by considering a similar but easier problem for the Lie algebra.

Definition 2.15. *Let G be a matrix Lie group. The* ***Lie algebra of*** *G, denoted $\mathfrak{g}$, is the set of all matrices X such that e^{tX} is in G for all real numbers t.*

This means that X is in $\mathfrak{g}$ if and only if the one-parameter subgroup generated by X lies in G. Note that even though G is a subgroup of $\mathrm{GL}(n;\mathbb{C})$ (and not necessarily of $\mathrm{GL}(n;\mathbb{R})$), we do *not* require that e^{tX} be in G for all complex numbers t, but only for all *real* numbers t. Also, it is definitely not enough to have just e^X in G. That is, it is easy to give an example of an X and a G such that $e^X \in G$ but such that $e^{tX} \notin G$ for some real values of t (Exercise 10). Such an X is not in the Lie algebra of G.

There is an abstract notion of a Lie algebra (not necessarily associated to any group), which is described in Section 2.8. The results of Section 2.6 will show that $\mathfrak{g}$ is, indeed, a Lie algebra in that sense.

It is customary to use lowercase Gothic (Fraktur) characters such as $\mathfrak{g}$ and $\mathfrak{h}$ to refer to Lie algebras.

We will show in Section 2.7 that every matrix Lie group is an embedded submanifold of $\mathrm{GL}(n;\mathbb{C})$. We will then show that $\mathfrak{g}$ is the tangent space to G at the identity. See Corollary 2.35. This means that $\mathfrak{g}$ can alternatively be defined as the set of all derivatives of smooth curves through the identity in G.

2.5.1 Physicists' Convention

Physicists are accustomed to considering the map $X \to e^{iX}$ instead of $X \to e^X$. Thus, a physicist would think of the Lie algebra of G as the set of all matrices X such that $e^{itX} \in G$ for all real numbers t. In the physics literature, the Lie algebra is frequently referred to as the space of "infinitesimal group elements." The physics literature does not always distinguish clearly between a matrix Lie group and its Lie algebra.

Before examining general properties of the Lie algebra, let us compute the Lie algebras of the matrix Lie groups introduced in the previous chapter.

2.5.2 The general linear groups

If X is any $n \times n$ complex matrix, then by Proposition 2.3, e^{tX} is invertible. Thus, the Lie algebra of $\mathrm{GL}(n;\mathbb{C})$ is the space of all $n \times n$ complex matrices. This Lie algebra is denoted $\mathfrak{gl}(n;\mathbb{C})$.

If X is any $n \times n$ real matrix, then e^{tX} will be invertible and real. On the other hand, if e^{tX} is real for all real numbers t, then $X = \frac{d}{dt} e^{tX}\big|_{t=0}$ will also be real. Thus, the Lie algebra of $\mathrm{GL}(n;\mathbb{R})$ is the space of all $n \times n$ real matrices, denoted $\mathfrak{gl}(n;\mathbb{R})$.

Note that the preceding argument shows that if G is a subgroup of $\mathrm{GL}(n;\mathbb{R})$, then the Lie algebra of G must consist entirely of real matrices. We will use this fact when appropriate in what follows.

2.5.3 The special linear groups

Recall Theorem 2.11: $\det(e^X) = e^{\mathrm{trace}(X)}$. Thus, if $\mathrm{trace}(X) = 0$, then $\det(e^{tX}) = 1$ for all real numbers t. On the other hand, if X is any $n \times n$ matrix such that $\det(e^{tX}) = 1$ for all t, then $e^{t\,\mathrm{trace}(X)} = 1$ for all t. This means that $t\,\mathrm{trace}(X)$ is an integer multiple of $2\pi i$ for all t, which is only possible if $\mathrm{trace}(X) = 0$. Thus, the Lie algebra of $\mathsf{SL}(n;\mathbb{C})$ is the space of all $n \times n$ complex matrices with trace zero, denoted $\mathsf{sl}(n;\mathbb{C})$.

Similarly, the Lie algebra of $\mathsf{SL}(n;\mathbb{R})$ is the space of all $n \times n$ real matrices with trace zero, denoted $\mathsf{sl}(n;\mathbb{R})$.

2.5.4 The unitary groups

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

$$\left(e^{tX}\right)^* = \left(e^{tX}\right)^{-1} = e^{-tX}. \tag{2.10}$$

By Point 2 of Proposition 2.3, $\left(e^{tX}\right)^* = e^{tX^*}$, and so (2.10) becomes

$$e^{tX^*} = e^{-tX}. \tag{2.11}$$

Clearly, a sufficient condition for (2.11) to hold is that $X^* = -X$. On the other hand, if (2.11) holds for all t, then by differentiating at $t = 0$, we see that $X^* = -X$ is necessary.

Thus, the Lie algebra of $\mathsf{U}(n)$ is the space of all $n \times n$ complex matrices X such that $X^* = -X$, denoted $\mathsf{u}(n)$.

By combining the two previous computations, we see that the Lie algebra of $\mathsf{SU}(n)$ is the space of all $n \times n$ complex matrices X such that $X^* = -X$ and $\mathrm{trace}(X) = 0$, denoted $\mathsf{su}(n)$.

2.5.5 The orthogonal groups

The identity component of $\mathsf{O}(n)$ is just $\mathsf{SO}(n)$. Since (Proposition 2.16) the exponential of a matrix in the Lie algebra is automatically in the identity component, the Lie algebra of $\mathsf{O}(n)$ is the same as the Lie algebra of $\mathsf{SO}(n)$.

Now, an $n \times n$ real matrix R is orthogonal if and only if $R^{tr} = R^{-1}$. So, given an $n \times n$ real matrix X, e^{tX} is orthogonal if and only if $(e^{tX})^{tr} = (e^{tX})^{-1}$, or

$$e^{tX^{tr}} = e^{-tX}. \tag{2.12}$$

Clearly, a sufficient condition for this to hold is that $X^{tr} = -X$. If (2.12) holds for all t, then by differentiating at $t = 0$, we must have $X^{tr} = -X$.

Thus, the Lie algebra of $\mathsf{O}(n)$, as well as the Lie algebra of $\mathsf{SO}(n)$, is the space of all $n \times n$ real matrices X with $X^{tr} = -X$, denoted $\mathsf{so}(n)$. Note that the condition $X^{tr} = -X$ forces the diagonal entries of X to be zero, and, so, necessarily the trace of X is zero.

The same argument shows that the Lie algebra of $\mathsf{SO}(n;\mathbb{C})$ is the space of $n \times n$ complex matrices satisfying $X^{tr} = -X$, denoted $\mathsf{so}(n;\mathbb{C})$. This is not the same as $\mathsf{su}(n)$.

2.5.6 The generalized orthogonal groups

A matrix A is in $\mathsf{O}(n;k)$ if and only if $A^{tr}gA = g$, where g is the $(n+k)\times(n+k)$ diagonal matrix with the first n diagonal entries equal to one and the last k diagonal entries equal to minus one. This condition is equivalent to the condition $g^{-1}A^{tr}g = A^{-1}$, or, since explicitly $g^{-1} = g$, $gA^{tr}g = A^{-1}$. Now, if X is an $(n+k) \times (n+k)$ real matrix, then e^{tX} is in $\mathsf{O}(n;k)$ if and only if

$$ge^{tX^{tr}}g = e^{tgX^{tr}g} = e^{-tX}.$$

This condition holds for all real t if and only if $gX^{tr}g = -X$. Thus, the Lie algebra of $\mathsf{O}(n;k)$, which is the same as the Lie algebra of $\mathsf{SO}(n;k)$, consists of all $(n+k) \times (n+k)$ real matrices X with $gX^{tr}g = -X$. This Lie algebra is denoted $\mathsf{so}(n;k)$.

(In general, the group $\mathsf{SO}(n;k)$ will not be connected, in contrast to the group $\mathsf{SO}(n)$. The identity component of $\mathsf{SO}(n;k)$, which is also the identity component of $\mathsf{O}(n;k)$, is denoted $\mathsf{SO}(n;k)_e$. The Lie algebra of $\mathsf{SO}(n;k)_e$ is the same as the Lie algebra of $\mathsf{SO}(n;k)$.)

2.5.7 The symplectic groups

These are denoted $\mathsf{sp}(n;\mathbb{R})$, $\mathsf{sp}(n;\mathbb{C})$, and $\mathsf{sp}(n)$. The calculation of these Lie algebras is similar to that of the generalized orthogonal groups, and I will just record the result here. Let J be the matrix in the definition of the symplectic groups. Then, $\mathsf{sp}(n;\mathbb{R})$ is the space of $2n \times 2n$ real matrices X such that $JX^{tr}J = X$, $\mathsf{sp}(n;\mathbb{C})$ is the space of $2n \times 2n$ complex matrices satisfying the same condition, and $\mathsf{sp}(n) = \mathsf{sp}(n;\mathbb{C}) \cap \mathsf{u}(2n)$. A simple calculation shows that the elements of $\mathsf{sp}(n;\mathbb{C})$ are precisely the $2n \times 2n$ matrices of the form

$$\begin{pmatrix} A & B \\ C & -A^{tr} \end{pmatrix},$$

where A is an arbitrary $n \times n$ matrix and B and C are arbitrary *symmetric* matrices.

2.5.8 The Heisenberg group

Recall that the Heisenberg group H is the group of all 3×3 real matrices A of the form

$$A = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.13}$$

with $a, b, c \in \mathbb{R}$. Recall also that in Section 2.2, Case 2, we computed the exponential of a matrix of the form

$$X = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix} \tag{2.14}$$

and saw that e^X was in H. On the other hand, if X is any matrix such that e^{tX} is of the form (2.13), then all of the entries of $X = \frac{d}{dt} e^{tX}\big|_{t=0}$ which are on or below the diagonal must be zero, so that X is of form (2.14).

Thus, the Lie algebra of the Heisenberg group is the space of all 3×3 real matrices that are strictly upper triangular.

2.5.9 The Euclidean and Poincaré groups

Recall that the Euclidean group $\mathsf{E}(n)$ is (or can be thought of as) the group of $(n+1) \times (n+1)$ real matrices of the form

$$\begin{pmatrix} & & x_1 \\ & R & \vdots \\ & & x_n \\ 0 & \cdots & 0 \quad 1 \end{pmatrix}$$

with $R \in \mathsf{O}(n)$. Now, if X is an $(n+1) \times (n+1)$ real matrix such that e^{tX} is in $\mathsf{E}(n)$ for all t, then $X = \frac{d}{dt} e^{tX}\big|_{t=0}$ must be zero along the bottom row:

$$X = \begin{pmatrix} & & y_1 \\ & Y & \vdots \\ & & y_n \\ 0 & \cdots & 0 \end{pmatrix} \tag{2.15}$$

Our goal, then, is to determine which matrices of the form (2.15) are actually in the Lie algebra of the Euclidean group. A simple computation shows that for $n \geq 1$,

$$\begin{pmatrix} & & y_1 \\ & Y & \vdots \\ & & y_n \\ 0 & \cdots & 0 \end{pmatrix}^n = \begin{pmatrix} & Y^n & Y^{n-1}y \\ 0 & \cdots & 0 \end{pmatrix},$$

where y is the column vector with entries $y_1, \ldots, y_n$. It follows that if X is as in (2.15), then e^{tX} is of the form

$$e^{tX} = \begin{pmatrix} & & * \\ & e^{tY} & \vdots \\ & & * \\ 0 & \cdots & 0 \quad 1 \end{pmatrix}.$$

Now, we have already established that e^{tY} is in $\mathsf{O}(n)$ for all t if and only if $Y^{tr} = -Y$. Thus, we see that the Lie algebra of $\mathsf{E}(n)$ is the space of all $(n+1) \times (n+1)$ real matrices of the form (2.15) with Y satisfying $Y^{tr} = -Y$.

A similar argument shows that the Lie algebra of $\mathsf{P}(n;1)$ is the space of all $(n+2) \times (n+2)$ real matrices of the form

$$\begin{pmatrix} & & y_1 \\ & Y & \vdots \\ & & y_{n+1} \\ 0 & \cdots & 0 \end{pmatrix},$$

with $Y \in \mathsf{so}(n;1)$.

2.6 Properties of the Lie Algebra

We will now establish various basic properties of the Lie algebra of a matrix Lie group. The reader is invited to verify by direct calculation that these general properties hold for the examples computed in the previous section.

Proposition 2.16. *Let G be a matrix Lie group, and X an element of its Lie algebra. Then, e^X is an element of the identity component of G.*

Proof. By definition of the Lie algebra, e^{tX} lies in G for all real t. However, as t varies from 0 to 1, e^{tX} is a continuous path connecting the identity to e^X. □

Proposition 2.17. *Let G be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Let X be an element of $\mathfrak{g}$, and A an element of G. Then, AXA^{-1} is in $\mathfrak{g}$.*

Proof. This is immediate, since, by Proposition 2.3,

$$e^{t(AXA^{-1})} = Ae^{tX}A^{-1},$$

and, thus, $Ae^{tX}A^{-1} \in G$ for all t. □

Theorem 2.18. *Let G be a matrix Lie group, $\mathfrak{g}$ its Lie algebra, and X and Y elements of $\mathfrak{g}$. Then*

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

If one follows the physics convention for the definition of the Lie algebra, then condition 3 should be replaced with the condition $-i\,(XY - YX) \in \mathfrak{g}$. Properties 1 and 2 show that $\mathfrak{g}$ is a real vector space, (i.e., a real subspace of the space of $M_n(\mathbb{C})$). Property 3 shows that $\mathfrak{g}$ is, in fact, a Lie algebra in the abstract sense described in Section 2.8. Note that Property 1 applies only to *real* numbers s (compare Definition 2.20).

Proof. Point 1 is immediate, since $e^{t(sX)} = e^{(ts)X}$, which must be in G if X is in $\mathfrak{g}$. Point 2 is easy to verify *if* X and Y commute, since, in that case, $e^{t(X+Y)} = e^{tX}e^{tY}$. If X and Y do not commute, this argument does not work. However, the Lie product formula states that

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

Because X and Y are in the Lie algebra, $e^{tX/m}$ and $e^{tY/m}$ are in G, as is $\left(e^{tX/m}e^{tY/m}\right)^m$, since G is a group. However, because G is a matrix Lie group, the limit of things in G must be again in G, provided that the limit is invertible. Since $e^{t(X+Y)}$ is automatically invertible, we conclude that it must be in G. This shows that $X + Y$ is in $\mathfrak{g}$.

Now for Point 3. Recall (Proposition 2.4) that $\frac{d}{dt}e^{tX}\big|_{t=0} = X$. It follows that $\frac{d}{dt}e^{tX}Y\big|_{t=0} = XY$, and, hence, by the product rule (Exercise 3),

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

Now, by Proposition 2.17, $e^{tX}Ye^{-tX}$ is in $\mathfrak{g}$ for all t. Furthermore, we have (by Points 1 and 2) established that $\mathfrak{g}$ is a real subspace of $M_n(\mathbb{C})$. This means, in particular, that $\mathfrak{g}$ is a topologically closed subset of $M_n(\mathbb{C})$. It follows that

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

belongs to $\mathfrak{g}$. □

Definition 2.19. *Given two $n \times n$ matrices A and B, the* ***bracket*** *(or* ***commutator****) of A and B, denoted $[A, B]$, is defined to be*

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

According to Theorem 2.18, the Lie algebra of any matrix Lie group is closed under brackets.

It is important to note that even if the elements of G have complex entries, the Lie algebra $\mathfrak{g}$ of G is not necessarily a complex vector space. That is, for X in $\mathfrak{g}$, iX may not be in $\mathfrak{g}$. For example, elements of $\mathsf{SU}(n)$ will, in general, have complex entries (i.e., $\mathsf{SU}(n)$ is not contained in $\mathsf{GL}(n;\mathbb{R})$). Nevertheless, if X is in the Lie algebra $\mathfrak{su}(n)$, then $X^* = -X$ and, so, $(iX)^* = iX$. This means that iX is not in $\mathfrak{su}(n)$ unless X is zero.

Definition 2.20. *A matrix Lie group G is said to be* ***complex*** *if its Lie algebra $\mathfrak{g}$ is a complex subspace of $M_n(\mathbb{C})$ (i.e., if $iX \in \mathfrak{g}$ for all $X \in \mathfrak{g}$).*

Examples of complex groups are $\mathsf{GL}(n;\mathbb{C})$, $\mathsf{SL}(n;\mathbb{C})$, $\mathsf{SO}(n;\mathbb{C})$, and $\mathsf{Sp}(n;\mathbb{C})$. The condition in Definition 2.20 is equivalent to the condition that G be a complex submanifold of $\mathsf{GL}(n;\mathbb{C})$. (See Appendix C.)

We return now to the setting of general, not necessarily complex, matrix Lie groups. The following very important theorem tells us that a Lie group homomorphism between two Lie groups gives rise in a natural way to a map between the corresponding Lie algebras. In particular, this will tell us that two isomorphic Lie groups have "the same" Lie algebras (i.e., the Lie algebras are isomorphic in the sense of Section 2.8). See Exercise 12.

Theorem 2.21. *Let G and H be matrix Lie groups, with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Suppose that $\Phi : G \to H$ is a Lie group homomorphism. Then, there exists a unique real linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ such that*

$$\Phi(e^X) = e^{\phi(X)} \tag{2.16}$$

for all $X \in \mathfrak{g}$. The map ϕ has following additional properties:

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

Suppose that G, H, and K are matrix Lie groups and $\Phi : H \to K$ and $\Psi : G \to H$ are Lie group homomorphisms. Let $\Lambda : G \to K$ be the composition of Φ and Ψ, $\Lambda(A) = \Phi(\Psi(A))$. Let ϕ, ψ, and λ be the associated Lie algebra maps. Then,

$$\lambda(X) = \phi(\psi(X)).$$

In practice, given a Lie group homomorphism Φ, the way one goes about computing ϕ is by using Property 3. Of course, since ϕ is (real) linear, it suffices to compute ϕ on a basis for $\mathfrak{g}$. In the language of differentiable manifolds, Property 3 says that ϕ is the derivative (or differential) of Φ at the identity, which is the standard definition of ϕ. (See also Corollary 2.35 in Section 2.7.)

A linear map with Property 2 is called a **Lie algebra homomorphism**. (See Section 2.8.) This theorem says that every Lie group homomorphism gives rise to a Lie algebra homomorphism. We will see eventually that the converse is true *under certain circumstances*. Specifically, suppose that G and H are Lie groups and that $\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. If G is *simply connected*, then there exists a unique Lie group homomorphism $\Phi : G \to H$ such that Φ and ϕ are related as in Theorem 2.21. (The proof of this deep result is in Chapter 3.) We now proceed with the proof of Theorem 2.21.

Proof. The proof is similar to the proof of Theorem 2.18. Since Φ is a continuous group homomorphism, $\Phi(e^{tX})$ will be a one-parameter subgroup of H, for each $X \in \mathfrak{g}$. Thus, by Theorem 2.13, there is a unique matrix Z such that

$$\Phi\left(e^{tX}\right) = e^{tZ} \tag{2.17}$$

for all $t \in \mathbb{R}$. This Z must lie in $\mathfrak{h}$ since $e^{tZ} = \Phi\left(e^{tX}\right) \in H$.

We now define $\phi(X) = Z$ and check in several steps that ϕ has the required properties.

Step 1: $\Phi(e^X) = e^{\phi(X)}$.
This follows from (2.17) and our definition of ϕ, by putting $t = 1$.

Step 2: $\phi(sX) = s\phi(X)$ *for all* $s \in \mathbb{R}$.
This is immediate, since if $\Phi(e^{tX}) = e^{tZ}$, then $\Phi(e^{tsX}) = e^{tsZ}$.

Step 3: $\phi(X + Y) = \phi(X) + \phi(Y)$.
By Steps 1 and 2,

$$e^{t\phi(X+Y)} = e^{\phi[t(X+Y)]} = \Phi\left(e^{t(X+Y)}\right).$$

By the Lie product formula and the fact that Φ is a continuous homomorphism, we have

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

However, we then have

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

Differentiating this result at $t = 0$ gives the desired result.

Step 4: $\phi\left(AXA^{-1}\right) = \Phi(A)\phi(X)\Phi(A)^{-1}$.
By Steps 1 and 2,

$$\exp t\phi(AXA^{-1}) = \exp\phi(tAXA^{-1}) = \Phi(\exp tAXA^{-1}).$$

Using a property of the exponential and Step 1, this becomes

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

Differentiating this at $t = 0$ gives the desired result.

Step 5: $\phi([X, Y]) = [\phi(X), \phi(Y)]$.
Recall from the proof of Theorem 2.18 that

$$[X, Y] = \frac{d}{dt}e^{tX}Ye^{-tX}\Big|_{t=0}.$$

Hence,

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

where we have used the fact that a derivative commutes with a linear transformation.

Now, by Step 4,

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

Step 6: $\phi(X) = \frac{d}{dt}\Phi(e^{tX})\big|_{t=0}$.
This follows from (2.17) and our definition of ϕ.

Step 7: *ϕ is the unique real linear map such that* $\Phi(e^X) = e^{\phi(X)}$.
Suppose that ψ is another such map. Then,

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

so that

$$\psi(X) = \frac{d}{dt}\Phi(e^{tX})\Big|_{t=0}.$$

Thus, by Step 6, ψ coincides with ϕ.

Step 8: $\lambda = \phi \circ \psi$.
For any $X \in \mathfrak{g}$,

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

Thus, $\lambda(X) = \phi(\psi(X))$. □

Definition 2.22 (The Adjoint Mapping). *Let G be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Then, for each $A \in G$, define a linear map $\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$ by the formula*

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

Proposition 2.23. *Let G be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Let $\mathsf{GL}(\mathfrak{g})$ denote the group of all invertible linear transformations of $\mathfrak{g}$. Then, for each $A \in G$, Ad_A is an invertible linear transformation of $\mathfrak{g}$ with inverse $\mathrm{Ad}_{A^{-1}}$, and the map $A \to \mathrm{Ad}_A$ is a group homomorphism of G into $\mathsf{GL}(\mathfrak{g})$. Furthermore, for each $A \in G$, Ad_A satisfies $\mathrm{Ad}_A([X,Y]) = [\mathrm{Ad}_A(X), \mathrm{Ad}_A(Y)]$ for all $X, Y \in \mathfrak{g}$.*

Proof. Easy. Note that Proposition 2.17 guarantees that $\mathrm{Ad}_A(X)$ is actually in $\mathfrak{g}$ for all $X \in \mathfrak{g}$. □

Since $\mathfrak{g}$ is a real vector space with some dimension k, $\mathsf{GL}(\mathfrak{g})$ is essentially the same as $\mathsf{GL}(k;\mathbb{R})$. Thus, we will regard $\mathsf{GL}(\mathfrak{g})$ as a matrix Lie group. It is easy to show that $\mathrm{Ad} : G \to \mathsf{GL}(\mathfrak{g})$ is continuous, and so is a Lie group homomorphism. By Theorem 2.21, there is an associated real linear map $X \to \mathrm{ad}_X$ from the Lie algebra of G to the Lie algebra of $\mathsf{GL}(\mathfrak{g})$ (i.e., from $\mathfrak{g}$ to $\mathfrak{gl}(\mathfrak{g})$), with the property that

$$e^{\mathrm{ad}_X} = \mathrm{Ad}\left(e^X\right).$$

Here, $\mathfrak{gl}(\mathfrak{g})$ is the Lie algebra of $\mathsf{GL}(\mathfrak{g})$, namely the space of all linear maps of $\mathfrak{g}$ to itself.

Proposition 2.24. *Let G be a matrix Lie group, let $\mathfrak{g}$ be its Lie algebra, and let $\mathrm{Ad} : G \to \mathsf{GL}(\mathfrak{g})$ be the Lie group homomorphism defined above. Let $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ be the associated Lie algebra map. Then, for all $X, Y \in \mathfrak{g}$*

$$\mathrm{ad}_X(Y) = [X, Y]. \tag{2.18}$$

Proof. Recall that by Point 3 of Theorem 2.21, ad can be computed as follows:

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

Thus,

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

which is what we wanted to prove. □

We have proved, as a consequence of Theorem 2.21 and Proposition 2.24, the following result, which we will make use of later.

Proposition 2.25. *For any X in $M_n(\mathbb{C})$, let $\mathrm{ad}_X : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be given by $\mathrm{ad}_X Y = [X, Y]$. Then, for any Y in $M_n(\mathbb{C})$, we have*

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

This result can also be proved by direct calculation—see Exercise 19.

2.7 The Exponential Mapping

Definition 2.26. *If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, then the **exponential mapping for** G is the map*

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

That is, the exponential mapping for G is the matrix exponential restricted to the Lie algebra $\mathfrak{g}$ of G. We have shown (Theorem 2.9) that every matrix in $\mathsf{GL}(n;\mathbb{C})$ is the exponential of some $n \times n$ matrix. Nevertheless, if $G \subset \mathsf{GL}(n;\mathbb{C})$ is a closed subgroup, there may exist A in G such that there is no X *in the Lie algebra* $\mathfrak{g}$ *of* G with $\exp X = A$. Consider, for example, the matrix

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

in $\mathsf{SL}(2;\mathbb{C})$. I claim that there exists no $X \in \mathsf{sl}(2;\mathbb{C})$ with $\exp X = A$. To see this, consider an arbitrary matrix X in $\mathsf{sl}(2;\mathbb{C})$. Since $\operatorname{trace}(X) = 0$, the eigenvalues of X are negatives of each other. There are then two possibilities. First, the eigenvalues of X could both be zero. In that case, $\exp X$ will have 1 as an eigenvalue and, so, $\exp X \neq A$. Second, the eigenvalues of X could be of the form $(\lambda, -\lambda)$, with λ being a nonzero complex number. In that case, X has *distinct* eigenvalues and is, therefore, diagonalizable. It follows that $\exp X$ is also diagonalizable. However, A is not diagonalizable. (The eigenvalues of A are -1 and -1; if it were diagonalizable it would have to be $-I$.) This shows that $\exp X \neq A$. (See also Exercises 26, 27, 29, 30, and 31.)

We see, then, that the exponential mapping for a matrix Lie group G does not necessarily map $\mathfrak{g}$ onto G. Furthermore, the exponential mapping may not be one-to-one on $\mathfrak{g}$. Nevertheless, it provides a crucial mechanism for passing information between the group and the Lie algebra. Indeed, we will see (Corollary 2.29) that the exponential mapping is *locally* one-to-one and onto, a result that will be essential, for example, in Chapter 3.

Theorem 2.27. *For* $0 < \varepsilon < \ln 2$, *let* $U_\varepsilon = \{X \in M_n(\mathbb{C}) \,|\, \|X\| < \varepsilon\}$ *and let* $V_\varepsilon = \exp(U_\varepsilon)$. *Suppose* $G \subset \mathsf{GL}(n;\mathbb{C})$ *is a matrix Lie group with Lie algebra* $\mathfrak{g}$. *Then there exists* $\varepsilon \in (0, \ln 2)$ *such that for all* $A \in V_\varepsilon$, A *is in* G *if and only if* $\log A$ *is in* $\mathfrak{g}$.

The condition $\varepsilon < \ln 2$ guarantees (Theorem 2.7) that for all $X \in V_\varepsilon$, $\log(\exp X)$ is defined and equal to X.

Note that if $X = \log A$ is in $\mathfrak{g}$, then $A = \exp X$ is in G. Thus, the content of the theorem is that for some ε, having A in $V_\varepsilon \cap G$ implies that $\log A$ must be in $\mathfrak{g}$. There are several important consequences of this theorem, described after the proof.

Proof. We begin with a lemma.

Lemma 2.28. *Suppose* B_m *are elements of* G *and that* $B_m \to I$. *Let* $Y_m = \log B_m$, *which is defined for all sufficiently large* m. *Suppose that* Y_m *is nonzero for all* m *and that* $Y_m / \|Y_m\| \to Y \in M_n(\mathbb{C})$. *Then,* $Y \in \mathfrak{g}$.

Proof. To show that $Y \in \mathfrak{g}$, we must show that $\exp tY \in G$ for all $t \in \mathbb{R}$. As $m \to \infty$, $(t/\|Y_m\|)\, Y_m \to tY$. Note that since $B_m \to I$, $Y_m \to 0$, and, so, $\|Y_m\| \to 0$. Thus, we can find integers k_m such that $(k_m \|Y_m\|) \to t$. Then,

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

However, $\exp(k_m Y_m) = \exp(Y_m)^{k_m} = (B_m)^{k_m} \in G$ and G is closed, and we conclude that $\exp(tY) \in G$. □

Let us think of $M_n(\mathbb{C})$ as $\mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$. Then, $\mathfrak{g}$ is a subspace of $\mathbb{R}^{2n^2}$. Let D denote the orthogonal complement of $\mathfrak{g}$ with respect to the usual inner product on $\mathbb{R}^{2n^2}$. Consider the map $\Phi : \mathfrak{g} \oplus D \to \mathsf{GL}(n;\mathbb{C})$ given by

$$\Phi(X,Y) = e^X e^Y.$$

Of course, we can identify $\mathfrak{g} \oplus D$ with $\mathbb{R}^{2n^2}$. Moreover, $\mathsf{GL}(n;\mathbb{C})$ is an open subset of $M_n(\mathbb{C}) \cong \mathbb{R}^{2n^2}$. Thus, we can regard Φ as a map from $\mathbb{R}^{2n^2}$ to itself.

Now, using the properties of the matrix exponential, we see that

$$\left.\frac{d}{dt}\Phi(tX,0)\right|_{t=0} = X,$$
$$\left.\frac{d}{dt}\Phi(0,tY)\right|_{t=0} = Y.$$

This shows that the derivative of Φ at the point $0 \in \mathbb{R}^{2n^2}$ is the identity. (Recall that the derivative at a point of a function from $\mathbb{R}^{2n^2}$ to itself is a linear map of $\mathbb{R}^{2n^2}$ to itself, in this case the identity map.) In particular, the derivative of Φ at 0 is invertible. Thus, the inverse function theorem says that Φ has a continuous local inverse, defined in a neighborhood of I.

Now, as we have remarked, what we need to prove is that for some ε, $A \in V_\varepsilon \cap G$ implies $\log A \in \mathfrak{g}$. Suppose this is not the case. Then we can find a sequence A_m in G such that $A_m \to I$ as $m \to \infty$ and such that for all m, $\log A_m \notin \mathfrak{g}$. Using the local inverse of the map Φ, we can write A_m (for all sufficiently large m) as

$$A_m = e^{X_m} e^{Y_m}, \quad X_m \in \mathfrak{g},\ Y_m \in D,$$

in such a way that X_m and Y_m tend to zero as m tends to infinity. We must have $Y_m \neq 0$, since otherwise we would have $\log A_m = X_m \in \mathfrak{g}$.

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

Corollary 2.29. *If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, there exists a neighborhood U of 0 in $\mathfrak{g}$ and a neighborhood V of I in G such that the exponential mapping takes U homeomorphically onto V.*

Proof. Let ε be such that Theorem 2.27 holds and set $U = U_\varepsilon \cap \mathfrak{g}$ and $V = V_\varepsilon \cap G$. The theorem implies that exp takes U onto V. Furthermore, exp is a homeomorphism of U onto V, since there is a continuous inverse map, namely, the restriction of the matrix logarithm to V. □

Definition 2.30. *If U and V are as in Corollary 2.29, then the inverse map $\exp^{-1} : V \to \mathfrak{g}$ is called the* ***logarithm*** *for G.*

Corollary 2.31. *If G is a connected matrix Lie group, then every element A of G can be written in the form*

$$A = e^{X_1} e^{X_2} \cdots e^{X_m} \tag{2.19}$$

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

Even if G is connected, it is definitely *not* the case in general that every element of G can be written as single exponential, $A = \exp X$ (with $X \in \mathfrak{g}$), as the example given earlier in this section shows.

Proof. Since G is connected, we can find a continuous path $A(t)$ in G with $A(0) = I$ and $A(1) = A$. Let V be a neighborhood of I in G as in Corollary 2.29, so that every element of V is the exponential of an element of $\mathfrak{g}$. A standard argument using the compactness of the interval $[0, 1]$ shows that we can pick a sequence of numbers $t_0, \ldots, t_m$ with $0 = t_0 < t_1 < \cdots < t_m = 1$ such that

$$A_{t_{k-1}}^{-1} A_{t_k} \in V$$

for all $k = 1, \ldots, m$. Then,

$$A = (A_{t_0}^{-1} A_{t_1})(A_{t_1}^{-1} A_{t_2}) \cdots (A_{t_{m-1}}^{-1} A_{t_m}).$$

If we choose $X_k \in \mathfrak{g}$ with $\exp X_k = A_{t_{k-1}}^{-1} A_{t_k}$ $(k = 1, \ldots, m)$, we have

$$A = e^{X_1} \cdots e^{X_m}.$$

□

Corollary 2.32. *Suppose G is a connected matrix Lie group, H is a matrix Lie group, and Φ_1 and Φ_2 are Lie group homomorphisms of G into H. Let ϕ_1 and ϕ_2 be the associated Lie algebra homomorphisms. If $\phi_1 = \phi_2$, then $\Phi_1 = \Phi_2$.*

Proof. Let g be any element of G. Since G is connected, Corollary 2.31 tells us that g can be written as $g = e^{X_1} e^{X_2} \cdots e^{X_n}$, with $X_i \in \mathfrak{g}$. Then,

$$\begin{aligned}
\Phi_1(g) &= \Phi_1(e^{X_1}) \cdots \Phi_1(e^{X_n}) \\
&= e^{\phi_1(X_1)} \cdots e^{\phi_1(X_n)} \\
&= e^{\phi_2(X_1)} \cdots e^{\phi_2(X_n)} \\
&= \Phi_2(e^{X_1}) \cdots \Phi_2(e^{X_n}) \\
&= \Phi_2(g).
\end{aligned}$$

□

We are now in a position to obtain Theorem 1.19 of Chapter 1 as a consequence of Theorem 2.27.

Corollary 2.33. *Every matrix Lie group G is a smooth embedded submanifold of $M_n(\mathbb{C})$ and, hence, a Lie group.*

Proof. Let $\varepsilon \in (0, \ln 2)$ be such that Theorem 2.27 holds. Then for any $A_0 \in G$, consider the neighborhood $A_0 V_\varepsilon$ of A_0 in $M_n(\mathbb{C})$. Note that $A \in A_0 V_\varepsilon$ if and only if $A_0^{-1} A \in V_\varepsilon$. Define a local coordinate system on $A_0 V_\varepsilon$ by writing each $A \in A_0 V_\varepsilon$ as $A = A_0 \exp X$, for $X \in U_\varepsilon \subset M_n(\mathbb{C})$. It follows from Theorem 2.27 that (for $A \in A_0 V_\varepsilon$) $A \in G$ if and only if $X \in \mathfrak{g}$. This means that in this local coordinate system defined near A_0, G looks like the subspace $\mathfrak{g}$ of $M_n(\mathbb{C})$. Since we can find such local coordinates near any point A_0 in G, G is an embedded submanifold of $M_n(\mathbb{C})$. This shows, as discussed in Section C.2.6, that G is a Lie group. □

Corollary 2.33 implies that a matrix Lie group G is necessarily *locally* path-connected. It follows that G is connected (in the usual topological sense) if and only if it is path-connected. Thus our definition of connectedness in Section 1.7 (which was actually *path*-connectedness) is equivalent to the usual topological definition.

Corollary 2.34. *Every continuous homomorphism between two matrix Lie groups is smooth.*

Proof. Given $A \in G$, we write nearby elements $B \in G$ (as in the proof of Corollary 2.33) as $B = A \exp X$, $X \in \mathfrak{g}$. Then,

$$\Phi(B) = \Phi(A)\Phi(\exp X) = \Phi(A)\exp(\phi(X)).$$

This says that in exponential coordinates near A, Φ is a composition of the linear map ϕ, the exponential mapping, and multiplication on the left by $\Phi(A)$, all of which are smooth. This shows that Φ is smooth near any point $A \in G$. □

Corollary 2.35. *Suppose $G \subset \mathrm{GL}(n;\mathbb{C})$ is a matrix Lie group with Lie algebra $\mathfrak{g}$. Then, a matrix X is in $\mathfrak{g}$ if and only if there exists a smooth curve γ in $M_n(\mathbb{C})$ such that 1) $\gamma(t)$ lies in G for all t; 2) $\gamma(0) = I$; 3) $d\gamma/dt|_{t=0} = X$. Thus, $\mathfrak{g}$ is the tangent space at the identity to G.*

See Proposition C.3 for a description of the tangent space of an embedded submanifold in terms of derivatives of smooth curves.

Proof. If X is in $\mathfrak{g}$, then we may take $\gamma(t) = \exp(tX)$ and then $\gamma(0) = I$ and $d\gamma/dt|_{t=0} = X$. In the other direction, suppose that $\gamma(t)$ is a smooth curve in G with $\gamma(0) = I$. Then, by Theorem 2.27, $\log(\gamma(t))$ is in $\mathfrak{g}$ for all sufficiently small t. Now, $\mathfrak{g}$ is a real subspace of $M_n(\mathbb{C})$ and, therefore, also a topologically closed subset of $M_n(\mathbb{C})$. Thus, the quantity

$$\left.\frac{d\log(\gamma(t))}{dt}\right|_{t=0} = \lim_{h\to 0}\frac{\log(\gamma(h))-0}{h}$$

is again in $\mathfrak{g}$. However,

$$\log(\gamma(t)) = (\gamma(t)-I) - \frac{(\gamma(t)-I)^2}{2} + \frac{(\gamma(t)-I)^3}{3} + \cdots .$$

If we differentiate this term by term (it is not hard to see that this is permitted) and apply the product rule, all terms but the first will give zero. (For example, the derivative of the second term is $-\frac{1}{2}[(d\gamma/dt)(\gamma(t)-I)+(\gamma(t)-I)(d\gamma/dt)]$, which is equal to zero at $t=0$.) Thus, we obtain that

$$\left.\frac{d\log(\gamma(t))}{dt}\right|_{t=0} = \left.\frac{d\gamma}{dt}\right|_{t=0} \in \mathfrak{g}.$$

□

2.8 Lie Algebras

We now consider the abstract notion of a Lie algebra, not necessarily given to us as the Lie algebra of a matrix Lie group. Proposition 2.37 shows that the Lie algebra of a matrix Lie group is indeed a Lie algebra in the abstract sense.

Definition 2.36. *A **finite-dimensional real or complex Lie algebra** is a finite-dimensional real or complex vector space $\mathfrak{g}$, together with a map $[\cdot,\cdot]$ from $\mathfrak{g}\times\mathfrak{g}$ into $\mathfrak{g}$, with the following properties:*

1. *$[\cdot,\cdot]$ is bilinear.*
2. *$[X,Y] = -[Y,X]$ for all $X, Y \in \mathfrak{g}$.*
3. *$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ for all $X, Y, Z \in \mathfrak{g}$.*

Condition 2 is called "skew symmetry." Condition 3 is called the **Jacobi identity**. Note also that Condition 2 implies that $[X,X] = 0$ for all $X \in \mathfrak{g}$. We will deal only with finite-dimensional Lie algebras and will from now on interpret "Lie algebra" as "finite-dimensional Lie algebra."

It should be emphasized here that $\mathfrak{g}$ can be *any* vector space (not necessarily a space of matrices) and that the "bracket" operation $[\cdot,\cdot]$ can be *any* bilinear, skew-symmetric map that satisfies the Jacobi identity. In particular, $[X,Y]$ is not necessarily equal to $XY-YX$; indeed, the expression $XY-YX$ does not even make sense in general, since $\mathfrak{g}$ does not necessarily have a product operation defined on it. For example, let $\mathfrak{g} = \mathbb{R}^3$ and define $[x,y]$ to be $x\times y$, where $\times$ is the cross product (vector product). This operation is, clearly, bilinear and skew-symmetric, and it can be checked that is satisfies the Jacobi identity. There is, so far as I can see, no product operation "xy" on $\mathbb{R}^3$ such that $x\times y = xy - yx$.

Although the bracket operation in a Lie algebra does not have to be given to us as $[X,Y] = XY - YX$, it is possible to construct Lie algebras in this way. That is to say, if $\mathcal{A}$ is an associative algebra and we define $[\cdot,\cdot] : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ by $[X,Y] = XY - YX$, then this operation does, indeed, make $\mathcal{A}$ into a Lie algebra. This operation is clearly bilinear and skew-symmetric, and it is a simple computation to check, using the associativity of $\mathcal{A}$, the Jacobi identity. For any Lie algebra, the Jacobi identity means that the bracket operation *behaves as if* it were $XY - YX$, even if it is not actually defined this way. Indeed, it can be shown that every Lie algebra $\mathfrak{g}$ can be embedded into some associative algebra $\mathcal{A}$ in such a way that the bracket on $\mathfrak{g}$ corresponds to the operation $XY - YX$ in $\mathcal{A}$.

If $\mathfrak{g}$ is a Lie algebra, we can think of the bracket operation as making $\mathfrak{g}$ into an algebra in the general sense. This algebra, however, is not associative. The Jacobi identity is to be thought of as a substitute for associativity.

Proposition 2.37. *The space $M_n(\mathbb{R})$ of all $n \times n$ real matrices is a real Lie algebra with respect to the bracket operation $[A,B] = AB - BA$. The space $M_n(\mathbb{C})$ of all $n \times n$ complex matrices is a complex Lie algebra with respect to the same bracket operation.*

Let V be a finite-dimensional real or complex vector space, and let $\mathfrak{gl}(V)$ denote the space of linear maps of V into itself. Then, $\mathfrak{gl}(V)$ becomes a real or complex Lie algebra with the bracket operation $[A,B] = AB - BA$.

Proof. The only nontrivial point is the Jacobi identity. The only way to prove this is to write everything out and see, and this is best left to the reader. Note that each double bracket generates 4 terms, for a total of 12. Each of the six orderings of $\{X,Y,Z\}$ occurs twice, once with a plus sign and once with a minus sign. Note that the associativity of the matrix product is essential to the proof. □

Definition 2.38. *A **subalgebra** of a real or complex Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h}$ of $\mathfrak{g}$ such that $[H_1,H_2] \in \mathfrak{h}$ for all H_1 and $H_2 \in \mathfrak{h}$. If $\mathfrak{g}$ is a complex Lie algebra and $\mathfrak{h}$ is a real subspace of $\mathfrak{g}$ which is closed under brackets, then $\mathfrak{h}$ is said to be a **real subalgebra** of $\mathfrak{g}$.*

*If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, then a linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ is called a **Lie algebra homomorphism** if $\phi([X,Y]) = [\phi(X), \phi(Y)]$ for all $X, Y \in \mathfrak{g}$. If, in addition, ϕ is one-to-one and onto, then ϕ is called a **Lie algebra isomorphism**. A Lie algebra isomorphism of a Lie algebra with itself is called a **Lie algebra automorphism**.*

A subalgebra of a Lie algebra is, again, a Lie algebra. A real subalgebra of a complex Lie algebra is a real Lie algebra. The inverse of a Lie algebra isomorphism is, again, a Lie algebra isomorphism.

Proposition 2.39. *The Lie algebra $\mathfrak{g}$ of a matrix Lie group G is a real Lie algebra.*

Proof. By Theorem 2.18, $\mathfrak{g}$ is a real subalgebra of the space $M_n(\mathbb{C})$ of all complex matrices and is, thus, a real Lie algebra. □

Theorem 2.40 (Ado). *Every finite-dimensional real Lie algebra is isomorphic to a subalgebra of* $\mathfrak{gl}(n;\mathbb{R})$. *Every finite-dimensional complex Lie algebra is isomorphic to a complex subalgebra of* $\mathfrak{gl}(n;\mathbb{C})$.

This deep theorem is proved, for example, in Varadarajan (1974). The proof is beyond the scope of this book and requires a careful examination of the structure of complex Lie algebras. The theorem tells us that every Lie algebra is (isomorphic to) a Lie algebra of matrices. This is in contrast to the situation for Lie groups, where most, but not all, Lie groups are matrix Lie groups—see Section C.3.

We now introduce the abstract Lie algebra version of the map "ad," which we introduced earlier for the Lie algebra of a matrix Lie group.

Definition 2.41. *Let* $\mathfrak{g}$ *be a Lie algebra. For* $X \in \mathfrak{g}$, *define a linear map* $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ *by*

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

*Thus, "*ad*" (i.e., the map* $X \to \mathrm{ad}_X$*) can be viewed as a linear map from* $\mathfrak{g}$ *into* $\mathfrak{gl}(\mathfrak{g})$, *where* $\mathfrak{gl}(\mathfrak{g})$ *denotes the space of linear operators from* $\mathfrak{g}$ *to* $\mathfrak{g}$.

Since $\mathrm{ad}_X(Y)$ is just $[X, Y]$, it might seem foolish to introduce the additional "ad" notation. However, thinking of $[X, Y]$ as a linear map in Y for each fixed X gives a somewhat different perspective. In any case, the "ad" notation is extremely useful in some situations. For example, instead of writing

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

we can now write

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

This sort of notation will be essential in Chapter 3.

Proposition 2.42. *If* $\mathfrak{g}$ *is a Lie algebra, then*

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

that is, ad: $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ *is a Lie algebra homomorphism.*

Proof. Observe that

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

whereas

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

So, we want to show that

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

or, equivalently,

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

which is exactly the Jacobi identity. □

2.8.1 Structure constants

Let $\mathfrak{g}$ be a finite-dimensional real or complex Lie algebra, and let $X_1, \ldots, X_n$ be a basis for $\mathfrak{g}$ (as a vector space). Then, for each i and j, $[X_i, X_j]$ can be written uniquely in the form

$$[X_i, X_j] = \sum_{k=1}^{n} c_{ijk} X_k.$$

The constants c_{ijk} are called the **structure constants** of $\mathfrak{g}$ (with respect to the chosen basis). Clearly, the structure constants determine the bracket operation on $\mathfrak{g}$. In some of the literature, the structure constants play an important role, although we will not have much necessity to use them in this book. (They appear mainly in Appendix D, where the quantities ε_{ijk} are the structure constants for the Lie algebra $\mathsf{so}(3)$.) In the physics literature, the structure constants are defined as $[X_i, X_j] = \sqrt{-1} \sum_k c_{ijk} X_k$, reflecting the factor of $\sqrt{-1}$ difference between the physics definition of the Lie algebra and our own.

The structure constants satisfy the following two conditions:

$$c_{ijk} + c_{jik} = 0,$$

$$\sum_m (c_{ijm} c_{mkl} + c_{jkm} c_{mil} + c_{kim} c_{mjl}) = 0$$

for all i, j, k, l. The first of these conditions comes from the skew symmetry of the bracket, and the second comes from the Jacobi identity. (The reader is invited to verify these conditions for himself.)

2.8.2 Direct sums

If $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are Lie algebras, we can define the **direct sum** of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ as follows. We consider the direct sum of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ in the vector space sense, and we define a bracket operation on $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ by

$$[(X_1, X_2), (Y_1, Y_2)] = ([X_1, Y_1], [X_2, Y_2]).$$

It is straightforward to verify that this operation satisfies the Jacobi identity and makes $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ into a Lie algebra. If $G_1 \subset \mathsf{GL}(n_1; \mathbb{C})$ and $G_2 \subset \mathsf{GL}(n_2; \mathbb{C})$ are matrix Lie groups and $G_1 \times G_2$ is their direct product (regarded as a subgroup of $\mathsf{GL}(n_1 + n_2; \mathbb{C})$ in the obvious way), then it is easily verified that the Lie algebra of $G_1 \times G_2$ is isomorphic to $\mathfrak{g}_1 \oplus \mathfrak{g}_2$.

2.9 The Complexification of a Real Lie Algebra

Definition 2.43. *If V is a finite-dimensional real vector space, then the **complexification** of V, denoted $V_{\mathbb{C}}$, is the space of formal linear combinations*

$$v_1 + iv_2,$$

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

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

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

Proposition 2.44. *Let* $\mathfrak{g}$ *be a finite-dimensional real Lie algebra and* $\mathfrak{g}_{\mathbb{C}}$ *its complexification (as a real vector space). Then, the bracket operation on* $\mathfrak{g}$ *has a unique extension to* $\mathfrak{g}_{\mathbb{C}}$ *which makes* $\mathfrak{g}_{\mathbb{C}}$ *into a complex Lie algebra. The complex Lie algebra* $\mathfrak{g}_{\mathbb{C}}$ *is called the* ***complexification*** *of the real Lie algebra* $\mathfrak{g}$.

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

$$[X_1 + iX_2, Y_1 + iY_2] = ([X_1, Y_1] - [X_2, Y_2]) + i([X_1, Y_2] + [X_2, Y_1]). \quad (2.20)$$

To show existence, we must now check that (2.20) is really bilinear and skew symmetric and that it satisfies the Jacobi identity. It is clear that (2.20) is *real* bilinear, and skew-symmetric. The skew symmetry means that if (2.20) is complex linear in the first factor, it is also complex linear in the second factor. Thus, we need only show that

$$[i(X_1 + iX_2), Y_1 + iY_2] = i[X_1 + iX_2, Y_1 + iY_2]. \quad (2.21)$$

The left-hand side of (2.21) is

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

whereas the right-hand side of (2.21) is

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

and, indeed, these are equal.

It remains to check the Jacobi identity. Of course, the Jacobi identity holds if X, Y, and Z are in $\mathfrak{g}$. However, observe that the expression on the left-hand side of the Jacobi identity is (complex!) linear in X for fixed Y and Z. It follows that the Jacobi identity holds if X is in $\mathfrak{g}_{\mathbb{C}}$, and Y and Z are in $\mathfrak{g}$. The same argument then shows that we can extend to Y in $\mathfrak{g}_{\mathbb{C}}$, and then to Z in $\mathfrak{g}_{\mathbb{C}}$. Thus, the Jacobi identity holds in $\mathfrak{g}_{\mathbb{C}}$. □

Proposition 2.45. *The Lie algebras* $\mathsf{gl}(n;\mathbb{C})$, $\mathsf{sl}(n;\mathbb{C})$, $\mathsf{so}(n;\mathbb{C})$, *and* $\mathsf{sp}(n;\mathbb{C})$ *are complex Lie algebras. In addition, we have the following isomorphisms of complex Lie algebras:*

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

Proof. From the computations in the previous section, we see easily that the specified Lie algebras are, in fact, complex subalgebras of $\mathsf{gl}(n;\mathbb{C})$ and hence are complex Lie algebras.

Now, $\mathsf{gl}(n;\mathbb{C})$ is the space of all $n \times n$ complex matrices, whereas $\mathsf{gl}(n;\mathbb{R})$ is the space of all $n \times n$ real matrices. Clearly, then, every $X \in \mathsf{gl}(n;\mathbb{C})$ can be written uniquely in the form $X_1 + iX_2$, with $X_1, X_2 \in \mathsf{gl}(n;\mathbb{R})$. This gives us a complex vector space isomorphism of $\mathsf{gl}(n;\mathbb{R})_{\mathbb{C}}$ with $\mathsf{gl}(n;\mathbb{C})$, and it is a triviality to check that this is a Lie algebra isomorphism.

On the other hand, $\mathsf{u}(n)$ is the space of all $n \times n$ complex skew-self-adjoint matrices. However, if X is any $n \times n$ complex matrix, then

$$X = \frac{X - X^*}{2} + i\frac{X + X^*}{2i},$$

where $(X - X^*)/2$ and $(X + X^*)/2i$ are both skew. Thus, X can be written as a skew matrix plus i times a skew matrix, and it is easy to see that this decomposition is unique. Thus, every X in $\mathsf{gl}(n;\mathbb{C})$ can be written uniquely as $X_1 + iX_2$, with X_1 and X_2 in $\mathsf{u}(n)$. It follows that $\mathsf{u}(n)_{\mathbb{C}} \cong \mathsf{gl}(n;\mathbb{C})$. If X has trace zero, then so do X_1 and X_2, which shows that $\mathsf{su}(n)_{\mathbb{C}} \cong \mathsf{sl}(n;\mathbb{C})$.

The verification of the remaining isomorphisms is similar and is left as an exercise to the reader. □

Note that $\mathsf{u}(n)_{\mathbb{C}} \cong \mathsf{gl}(n;\mathbb{R})_{\mathbb{C}} \cong \mathsf{gl}(n;\mathbb{C})$. However, $\mathsf{u}(n)$ is *not* isomorphic to $\mathsf{gl}(n;\mathbb{R})$, except when $n = 1$. The real Lie algebras $\mathsf{u}(n)$ and $\mathsf{gl}(n;\mathbb{R})$ are called **real forms** of the complex Lie algebra $\mathsf{gl}(n;\mathbb{C})$. A given complex Lie algebra may have several nonisomorphic real forms. See Exercise 17.

Physicists do not always clearly distinguish between a matrix Lie group and its (real) Lie algebra, or between a real Lie algebra and its complexification. Thus, for example, some references in the physics literature to $\mathsf{SU}(2)$ actually refer to the complexified Lie algebra, $\mathsf{sl}(2;\mathbb{C})$.

2.10 Exercises

1. The Schwarz inequality from elementary analysis tells us that for all $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ in $\mathbb{C}^n$, we have

$$|u_1v_1 + \cdots + u_nv_n|^2 \leq \left(\sum_{k=1}^{n} |u_k|^2\right)\left(\sum_{k=1}^{n} |v_k|^2\right).$$

Use this to verify that $\|XY\| \leq \|X\|\,\|Y\|$ for all $X, Y \in M_n(\mathbb{C})$, where the norm $\|X\|$ of a matrix X is defined by (2.2).

2. Show that for $X \in M_n(\mathbb{C})$ and any orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathbb{C}^n$, $\|X\|^2 = \sum_{j,k=1}^{n} |\langle u_j, Xu_k\rangle|^2$, where $\|X\|$ is defined by (2.2). Now show that if v is an eigenvector for X with eigenvalue λ, then $|\lambda| \leq \|X\|$.
3. *The product rule.* Recall that a matrix-valued function $A(t)$ is said to be smooth if each $A_{ij}(t)$ is smooth. The derivative of such a function is defined as

$$\left(\frac{dA}{dt}\right)_{ij} = \frac{dA_{ij}}{dt}$$

or, equivalently,

$$\frac{d}{dt}A(t) = \lim_{h\to 0} \frac{A(t+h) - A(t)}{h}.$$

Let $A(t)$ and $B(t)$ be two such functions. Prove that $A(t)B(t)$ is again smooth and that

$$\frac{d}{dt}\left[A(t)B(t)\right] = \frac{dA}{dt}B(t) + A(t)\frac{dB}{dt}.$$

4. Show that for all $X \in M_n(\mathbb{C})$,

$$\lim_{m\to\infty} \left[I + \frac{X}{m}\right]^m = e^X.$$

5. Using Theorem B.7, show that every $n \times n$ complex matrix A is the limit of a sequence of diagonalizable matrices.
Hint: If the characteristic polynomial of A has n distinct roots, then A is diagonalizable.
6. Show that every 2×2 matrix X with trace zero satisfies

$$X^2 = -\det(X) I.$$

If X is 2×2 with trace zero, show by direct calculation using the power series for the exponential that

$$e^X = \cos\left(\sqrt{\det X}\right) I + \frac{\sin\sqrt{\det X}}{\sqrt{\det X}} X. \tag{2.22}$$

Use this to give an alternative derivation of the result in (2.7).
Notes: Since the functions $\cos\theta$ and $\sin\theta/\theta$ are even functions of θ, the value of (2.22) is independent of the choice of the square root of $\det X$. The value of the coefficient of X in (2.22) is to be interpreted as 1 when $\det X = 0$, in accordance with the limit $\lim_{\theta\to 0} \sin\theta/\theta = 1$.

7. Use the result of Exercise 6 to compute the exponential of the matrix

$$X = \begin{pmatrix} 4 & 3 \\ -1 & 2 \end{pmatrix}.$$

Hint: Write X as the sum of a multiple of the identity and a matrix with trace zero.

8. A matrix A is said to be **unipotent** if $A - I$ is nilpotent (i.e., if A is of the form $A = I + N$, with N nilpotent). Note that $\log A$ is defined whenever A is unipotent, because the series in Definition 2.6 terminates.
(a) Show that if A is unipotent, then $\log A$ is nilpotent.
(b) Show that if X is nilpotent, then e^X is unipotent.
(c) Show that if A is unipotent, then $\exp(\log A) = A$ and that if X is nilpotent, then $\log(\exp X) = X$.
Hint: Let $A(t) = I + t(A - I)$. Show that $\exp(\log(A(t)))$ depends polynomially on t and that $\exp(\log(A(t))) = A(t)$ for all sufficiently small t.

9. Show that every invertible $n \times n$ matrix A can be written as $A = e^X$ for some $X \in M_n(\mathbb{C})$.
Hint: Theorem B.5 implies that A is similar to a block-diagonal matrix in which each block is of the form $\lambda I + N_\lambda$, with N_λ being nilpotent. Use this result and Exercise 8.

10. Give an example of a matrix Lie group G and a matrix X such that $e^X \in G$, but $X \notin \mathfrak{g}$.

11. Suppose G is a matrix Lie group in $\mathrm{GL}(n;\mathbb{C})$ and let $\mathfrak{g}$ be its Lie algebra. Suppose that A is in G and that $\|A - I\| < 1$, so that the power series for $\log A$ is convergent. Is it necessarily the case that $\log A$ is in $\mathfrak{g}$? Prove or give a counterexample.

12. Show that two isomorphic matrix Lie groups have isomorphic Lie algebras.

13. *The Lie algebra* $\mathrm{so}(3;1)$. Write out explicitly the general form of a 4×4 real matrix in $\mathrm{so}(3;1)$.

14. Verify directly that Proposition 2.17 and Theorem 2.18 hold for the Lie algebra of $\mathrm{SU}(n)$.

15. *The Lie algebra* $\mathrm{su}(2)$. Show that the following matrices form a basis for the real Lie algebra $\mathrm{su}(2)$:

$$E_1 = \tfrac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}; \; E_2 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \; E_3 = \tfrac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Compute $[E_1, E_2]$, $[E_2, E_3]$, and $[E_3, E_1]$. Show that there is an invertible linear map $\phi : \mathrm{su}(2) \to \mathbb{R}^3$ such that $\phi([X,Y]) = \phi(X) \times \phi(Y)$ for all $X, Y \in \mathrm{su}(2)$, where $\times$ denotes the cross product (vector product) on $\mathbb{R}^3$.

16. *The Lie algebras* $\mathrm{su}(2)$ *and* $\mathrm{so}(3)$. Show that the real Lie algebras $\mathrm{su}(2)$ and $\mathrm{so}(3)$ are isomorphic.
Note: Nevertheless, the corresponding *groups* $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$ are not isomorphic. (Rather, $\mathrm{SO}(3)$ is isomorphic to $\mathrm{SU}(2)/\{I, -I\}$.)

17. *The Lie algebras* $\mathsf{su}(2)$ *and* $\mathsf{sl}(2;\mathbb{R})$. Show that $\mathsf{su}(2)$ and $\mathsf{sl}(2;\mathbb{R})$ are not isomorphic Lie algebras, even though $\mathsf{su}(2)_{\mathbb{C}} \cong \mathsf{sl}(2;\mathbb{R})_{\mathbb{C}} \cong \mathsf{sl}(2;\mathbb{C})$.
Hint: Using Exercise 15, show that $\mathsf{su}(2)$ has no two-dimensional subalgebras.
18. Let G be a matrix Lie group and let $\mathfrak{g}$ be its Lie algebra. For each $A \in G$, show that Ad_A is a Lie algebra automorphism of $\mathfrak{g}$.
19. ("Ad" and "ad") Let X and Y be $n \times n$ matrices. Show by induction that

$$(\mathrm{ad}_X)^m(Y) = \sum_{k=0}^{m} \binom{m}{k} X^k Y (-X)^{m-k},$$

where

$$(\mathrm{ad}_X)^m(Y) = \underbrace{[X, \cdots [X, [X, Y]] \cdots]}_{m}.$$

Now, show by direct computation that

$$e^{\mathrm{ad}_X}(Y) = \mathrm{Ad}_{e^X}(Y) = e^X Y e^{-X}.$$

Assume that it is legal to multiply power series term by term. (This result was obtained indirectly in Proposition 2.25.)
Hint: Recall that Pascal's Triangle gives a relationship between numbers of the form $\binom{m+1}{k}$ and numbers of the form $\binom{m}{k}$.
20. If $\mathfrak{g}$ is a Lie algebra, then a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is called an **ideal** if $[X, H] \in \mathfrak{h}$ for all $X \in \mathfrak{g}$ and $H \in \mathfrak{h}$. If $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie algebra homomorphism, show that $\ker \phi$ is an ideal in $\mathfrak{g}_1$.
21. Classify up to isomorphism all one-dimensional and two-dimensional real Lie algebras. (There is one isomorphism class of one-dimensional algebras and two isomorphism classes of two-dimensional algebras.)
22. Show that for any Lie algebra $\mathfrak{g}$ and any X in $\mathfrak{g}$, ad_X is a derivation of $\mathfrak{g}$; that is,

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

for all Y and Z in $\mathfrak{g}$.
23. *The complexification of a real Lie algebra.* Let $\mathfrak{g}$ be a real Lie algebra, $\mathfrak{g}_{\mathbb{C}}$ its complexification, and $\mathfrak{h}$ an arbitrary complex Lie algebra. Show that every real Lie algebra homomorphism of $\mathfrak{g}$ into $\mathfrak{h}$ extends uniquely to a complex Lie algebra homomorphism of $\mathfrak{g}_{\mathbb{C}}$ into $\mathfrak{h}$. (This is the **universal property** of the complexification of a real Lie algebra. This property can be used as an alternative definition of the complexification.)
24. If $\mathfrak{g}$ is a Lie algebra, the **center** of $\mathfrak{g}$ is the set of all $Z \in \mathfrak{g}$ such that $[X, Z] = 0$ for all $X \in \mathfrak{g}$. Show that the center of $\mathfrak{g}$ is an ideal (as defined in Exercise 20).
25. Suppose that G is a connected, commutative matrix Lie group with Lie algebra $\mathfrak{g}$. Show that the exponential mapping for G maps $\mathfrak{g}$ onto G.

26. *The exponential mapping for the Heisenberg group.* Show that the exponential mapping from the Lie algebra of the Heisenberg group to the Heisenberg group is one-to-one and onto.
27. *The exponential mapping for* $\mathsf{U}(n)$. Show that the exponential mapping from $\mathsf{u}(n)$ to $\mathsf{U}(n)$ is onto, but not one-to-one. (Note that this shows that $\mathsf{U}(n)$ is connected.)
    Hint: Every unitary matrix has an orthonormal basis of eigenvectors.
28. Consider the space $\mathsf{gl}(n;\mathbb{C})$ of all $n \times n$ complex matrices. As usual, for $X \in \mathsf{gl}(n;\mathbb{C})$, define $\mathrm{ad}_X : \mathsf{gl}(n;\mathbb{C}) \to \mathsf{gl}(n;\mathbb{C})$ by $\mathrm{ad}_X(Y) = [X,Y]$. Suppose that X is a diagonalizable matrix. Show, then, that ad_X is diagonalizable as an operator on $\mathsf{gl}(n;\mathbb{C})$.
    Hint: Consider first the case where X is actually diagonal.
    Note: The problem of diagonalizing ad_X is an important one that we will encounter again in Chapter 6, when we consider semisimple Lie algebras.
29. Show explicitly that $\exp : \mathsf{so}(3) \to \mathsf{SO}(3)$ is onto.
    Hint: Using Exercise 16 from Chapter 1, show that in a suitable orthonormal basis, R is of the form

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}.$$

30. *The exponential mapping for* $\mathsf{SL}(2;\mathbb{R})$. Show that the image of the exponential mapping for $\mathsf{SL}(2;\mathbb{R})$ consists of precisely those matrices $A \in \mathsf{SL}(2;\mathbb{R})$ such that $\mathrm{trace}\,(A) > -2$, together with the matrix $-I$ (which has trace -2). To do this, consider the possibilities for the eigenvalues of a matrix in the Lie algebra $\mathsf{sl}(2;\mathbb{R})$ and in the group $\mathsf{SL}(2;\mathbb{R})$. In the Lie algebra, show that the eigenvalues are of the form $(\lambda, -\lambda)$ or $(i\lambda, -i\lambda)$, with λ real. In the group, show that the eigenvalues are of the form $(a, 1/a)$ or $(-a, -1/a)$, with a real and positive, or of the form $(e^{i\theta}, e^{-i\theta})$, with θ real. The case of a repeated eigenvalue ($(0,0)$ in the Lie algebra and $(1,1)$ or $(-1,-1)$ in the group) will have to be treated separately using the Jordan canonical form (Section B.4).
    Show that the image of the exponential mapping is not dense in $\mathsf{SL}(2;\mathbb{R})$.
31. Determine the image of the exponential mapping for $\mathsf{SL}(2;\mathbb{C})$. Is the image of the exponential mapping dense in $\mathsf{SL}(2;\mathbb{C})$?

3

The Baker–Campbell–Hausdorff Formula

In this chapter, we will, as usual, restrict our attention mainly to matrix Lie groups. Nevertheless, the proofs of the main results are the same for general Lie groups, provided one has already established the basic results about the Lie algebra and the exponential mapping for general Lie groups.

3.1 The Baker–Campbell–Hausdorff Formula for the Heisenberg Group

A crucial result of this chapter will be the following: Let G and H be matrix Lie groups, with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, and suppose that G is simply connected. Then, if $\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, there exists a unique Lie group homomorphism $\Phi : G \to H$ such that $\Phi(\exp X) = \exp(\phi(X))$ for all X in $\mathfrak{g}$. (This is Theorem 3.7 in Section 3.6.) This result is extremely important because it implies that if G is simply connected, then there is a natural one-to-one correspondence between the representations of G and the representations of its Lie algebra $\mathfrak{g}$ (as explained in Chapter 4). In practice, it is much easier to determine the representations of the Lie algebra than to determine directly the representations of the corresponding group.

This result (relating Lie algebra homomorphisms and Lie group homomorphisms) is deep. The "modern" proof (e.g., Varadarajan (1974), Theorem 2.7.5) makes use of the Frobenius theorem, which is both hard to understand and hard to prove (Varadarajan (1974), Section 1.3). Our proof will, instead, use the Baker–Campbell–Hausdorff formula, which is more easily stated and more easily motivated than the Frobenius theorem, but still deep.

The idea is the following. The desired group homomorphism $\Phi : G \to H$ must satisfy

$$\Phi\left(e^X\right) = e^{\phi(X)}. \tag{3.1}$$

We would like, then, to *define* Φ by this relation. This approach has two serious difficulties. First, a given element of G may not be expressible as e^X

(with X in $\mathfrak{g}$), and even if it is, the X may not be unique. Second, it is very far from clear why the Φ in (3.1) (even to the extent it is well defined) should be a group homomorphism.

It is the second issue which the Baker–Campbell–Hausdorff formula addresses. (The first issue will be addressed using the simple connectedness of G.) Specifically, (one form of) the Baker–Campbell–Hausdorff formula says that if X and Y are sufficiently small, then

$$\log(e^X e^Y) = X + Y + \tfrac{1}{2}[X,Y] + \tfrac{1}{12}[X,[X,Y]] - \tfrac{1}{12}[Y,[X,Y]] + \cdots. \quad (3.2)$$

It is not supposed to be evident at the moment what "$\cdots$" refers to. The only important point is that all of the terms in (3.2) are given in terms of X and Y, brackets of X and Y, brackets of brackets involving X and Y, etc. Then, because ϕ is a Lie algebra homomorphism,

$$\begin{aligned}\phi\left(\log(e^X e^Y)\right) &= \phi(X) + \phi(Y) + \tfrac{1}{2}[\phi(X),\phi(Y)] \\ &\quad + \tfrac{1}{12}[\phi(X),[\phi(X),\phi(Y)]] - \tfrac{1}{12}[\phi(Y),[\phi(X),\phi(Y)]] + \cdots \\ &= \log\left(e^{\phi(X)}e^{\phi(Y)}\right). \end{aligned} \quad (3.3)$$

The relation (3.3) is extremely significant. For, of course, we have

$$e^X e^Y = e^{\log(e^X e^Y)},$$

and so by (3.1),

$$\Phi\left(e^X e^Y\right) = e^{\phi(\log(e^X e^Y))}.$$

Thus, (3.3) tells us that

$$\Phi\left(e^X e^Y\right) = e^{\log\left(e^{\phi(X)}e^{\phi(Y)}\right)} = e^{\phi(X)}e^{\phi(Y)} = \Phi(e^X)\Phi(e^Y).$$

Thus, the Baker–Campbell–Hausdorff formula shows that on elements of the form e^X, with X small, Φ is a group homomorphism. (See Corollary 3.4.)

Another way of looking at this is to say that the Baker–Campbell–Hausdorff formula shows that all the information about the group product, at least near the identity, is "encoded" in the Lie algebra. Thus, if ϕ is a Lie algebra homomorphism (which by definition preserves the Lie algebra structure) and if we define Φ near the identity by (3.1), then we can expect Φ to preserve the group structure (i.e., to be a group homomorphism).

In this section, we will look at how all of this works out in the very special case of the Heisenberg group. In the next section, we will consider the general situation.

Theorem 3.1. *Suppose X and Y are $n \times n$ complex matrices, and that X and Y commute with their commutator. That is, suppose that*

$$[X,[X,Y]] = [Y,[X,Y]] = 0.$$

Then,

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

This is the special case of (3.2) in which the series terminates after the $[X, Y]$ term.

Proof. Consider X and Y in $M_n(\mathbb{C})$. We will prove that,

$$e^{tX}e^{tY} = \exp\left(tX + tY + \frac{t^2}{2}[X, Y]\right),$$

which reduces to the desired result in the case $t = 1$. Since, by assumption, $[X, Y]$ commutes with X and Y, the above relation is equivalent to

$$e^{tX}e^{tY}e^{-\frac{t^2}{2}[X,Y]} = e^{t(X+Y)}. \tag{3.4}$$

Let us denote by $A(t)$ the left-hand side of (3.4) and by $B(t)$ the right-hand side. Our strategy will be to show that $A(t)$ and $B(t)$ satisfy the same differential equation, with the same initial conditions. We can see immediately that

$$\frac{dB}{dt} = B(t)(X + Y).$$

On the other hand, differentiating $A(t)$ by means of the product rule gives

$$\begin{aligned}\frac{dA}{dt} &= e^{tX}Xe^{tY}e^{-\frac{t^2}{2}[X,Y]} + e^{tX}e^{tY}Ye^{-\frac{t^2}{2}[X,Y]} \\ &\quad + e^{tX}e^{tY}e^{-\frac{t^2}{2}[X,Y]}\left(-t[X, Y]\right).\end{aligned} \tag{3.5}$$

(The correctness of the last term may be verified by differentiating term by term.)

Now, since Y commutes with $[X, Y]$, it also commute with $e^{-\frac{t^2}{2}[X,Y]}$. Thus, the second term on the right in (3.5) can be rewritten as

$$e^{tX}e^{tY}e^{-\frac{t^2}{2}[X,Y]}Y.$$

The first term on the right in (3.5) is more complicated, since X does not necessarily commute with e^{tY}. However,

$$\begin{aligned}Xe^{tY} &= e^{tY}e^{-tY}Xe^{tY} \\ &= e^{tY}\mathrm{Ad}_{e^{-tY}}(X) \\ &= e^{tY}e^{-t\mathrm{ad}_Y}(X).\end{aligned}$$

However, since $[Y, [Y, X]] = -[Y, [X, Y]] = 0$,

$$e^{-t\mathrm{ad}_Y}(X) = X - t[Y, X] = X + t[X, Y],$$

with all higher terms being zero. Using the fact that everything commutes with $e^{-\frac{t^2}{2}[X,Y]}$ gives

$$e^{tX} X e^{tY} e^{-\frac{t^2}{2}[X,Y]} = e^{tX} e^{tY} e^{-\frac{t^2}{2}[X,Y]} (X + t[X,Y]).$$

Making these substitutions into (3.5) gives

$$\begin{aligned}\frac{dA}{dt} &= e^{tX} e^{tY} e^{-\frac{t^2}{2}[X,Y]} (X + t[X,Y]) + e^{tX} e^{tY} e^{-\frac{t^2}{2}[X,Y]} Y \\ &\quad + e^{tX} e^{tY} e^{-\frac{t^2}{2}[X,Y]} (-t[X,Y]) \\ &= e^{tX} e^{tY} e^{-\frac{t^2}{2}[X,Y]} (X + Y) \\ &= A(t)(X + Y).\end{aligned}$$

Thus, $A(t)$ and $B(t)$ satisfy the same differential equation. Moreover, $A(0) = B(0) = I$. Thus, by standard uniqueness results for ordinary differential equations, $A(t) = B(t)$ for all t. Putting $t = 1$ gives the theorem. □

Theorem 3.2. *Let H denote the Heisenberg group and $\mathfrak{h}$ its Lie algebra. Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$ and let $\phi : \mathfrak{h} \to \mathfrak{g}$ be a Lie algebra homomorphism. Then, there exists a unique Lie group homomorphism $\Phi : H \to G$ such that*

$$\Phi\left(e^X\right) = e^{\phi(X)}$$

for all $X \in \mathfrak{h}$.

Proof. Recall (Exercise 26 in Chapter 2) that the Heisenberg group has the very special property that its exponential mapping is one-to-one and onto. Let "log" denote the inverse of this map. Define $\Phi : H \to G$ by the formula

$$\Phi(A) = e^{\phi(\log A)}.$$

We will show that Φ is a Lie group homomorphism.

If X and Y are in the Lie algebra of the Heisenberg group (3×3 strictly upper triangular matrices), then $[X,Y]$ is of the form

$$\begin{pmatrix} 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

such a matrix commutes with both X and Y. Thus, X and Y commute with their commutator. Since ϕ is a Lie algebra homomorphism, $\phi(X)$ and $\phi(Y)$ will also commute with their commutator:

$$\begin{aligned}[\phi(X), [\phi(X), \phi(Y)]] &= \phi([X, [X,Y]]) = 0, \\ [\phi(Y), [\phi(X), \phi(Y)]] &= \phi([Y, [X,Y]]) = 0.\end{aligned}$$

We want to show that Φ is a homomorphism (i.e., that $\Phi(AB) = \Phi(A)\Phi(B)$). To show this, note that A can be written as e^X for a unique $X \in \mathfrak{h}$ and B can be written as e^Y for a unique $Y \in \mathfrak{h}$. Thus, by Theorem 3.1,

$$\Phi(AB) = \Phi\left(e^X e^Y\right) = \Phi\left(e^{X+Y+\frac{1}{2}[X,Y]}\right).$$

Using the definition of Φ and the fact that ϕ is a Lie algebra homomorphism, we see that

$$\Phi(AB) = \exp\left(\phi(X) + \phi(Y) + \frac{1}{2}[\phi(X), \phi(Y)]\right).$$

Finally, using Theorem 3.1 again (applied to the elements $\phi(X)$ and $\phi(Y)$), we have

$$\Phi(AB) = e^{\phi(X)} e^{\phi(Y)} = \Phi(A)\Phi(B).$$

Thus, Φ is a group homomorphism. It is easy to check that Φ is continuous (by checking that log, exp, and ϕ are all continuous), and, so, Φ is a Lie group homomorphism. Moreover, Φ by definition has the right relationship to ϕ. Furthermore, since the exponential mapping is one-to-one and onto, there can be at most one Φ with $\Phi\left(e^X\right) = e^{\phi(X)}$. □

3.2 The General Baker–Campbell–Hausdorff Formula

The importance of the Baker–Campbell–Hausdorff formula lies not in the details of the formula, but in the fact that there is a formula and in the fact that it gives $\log(e^X e^Y)$ in terms of brackets of X and Y, brackets of brackets, and so forth. This tells us something very important, namely that (at least for elements of the form e^X, X small) the group product for a matrix Lie group G is *completely expressible in terms of the Lie algebra.* (This is because $\log(e^X e^Y)$ and, hence, also $e^X e^Y$ itself, can be computed in Lie-algebraic terms by (3.2).)

We will actually state and prove an integral form of the Baker–Campbell–Hausdorff formula, rather than the series form (3.2). However, the integral form is sufficient to obtain the desired result (3.3). (See Corollary 3.4.) The series form of the Baker–Campbell–Hausdorff formula is stated precisely and proved in Varadarajan (1974), Section 2.15. See also Section 3.5.

Consider the function

$$g(z) = \frac{\log z}{1 - \frac{1}{z}}.$$

This function is defined and analytic in the disk $\{|z-1| < 1\}$, and, thus, for z in this set, $g(z)$ can be expressed as

$$g(z) = \sum_{m=0}^{\infty} a_m (z-1)^m,$$

for some set of constants $\{a_m\}$. This series has radius of convergence one.

Now, suppose V is a finite-dimensional complex vector space. Choose an arbitrary basis for V, so that V can be identified with $\mathbb{C}^n$ and, thus, the norm

of a linear operator on V can be defined. Then, for any operator A on V with $\|A - I\| < 1$, we can define

$$g(A) = \sum_{m=0}^{\infty} a_m (A - I)^m.$$

We are now ready to state the integral form of the Baker–Campbell–Hausdorff formula.

Theorem 3.3 (Baker–Campbell–Hausdorff). *For all $n \times n$ complex matrices X and Y with $\|X\|$ and $\|Y\|$ sufficiently small,*

$$\log\left(e^X e^Y\right) = X + \int_0^1 g(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y})(Y)\, dt. \tag{3.6}$$

The proof of this theorem is given in Section 3.4 of this chapter. Note that $e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}$ and, hence, also $g(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y})$ are linear operators on the space $\mathrm{gl}(n;\mathbb{C})$ of all $n \times n$ complex matrices. In (3.6), this operator is being applied to the matrix Y. The fact that X and Y are assumed small guarantees that $e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}$ is close to the identity operator on $\mathrm{gl}(n;\mathbb{C})$ for $0 \leq t \leq 1$. This ensures that $g(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y})$ is well defined.

If X and Y commute, then we expect to have $\log(e^X e^Y) = \log(e^{X+Y}) = X + Y$. Exercise 5 shows that the Baker–Campbell–Hausdorff formula indeed gives $X + Y$ in that case.

Formula (3.6) is admittedly horrible looking. However, we are interested not in the details of the formula but in the fact that it expresses $\log(e^X e^Y)$ (and hence $e^X e^Y$) in terms of the Lie-algebraic quantities ad_X and ad_Y.

Since the goal of the Baker–Campbell–Hausdorff theorem is to compute $\log(e^X e^Y)$, one may well ask, "Why do we not simply expand both exponentials and the logarithm in power series and multiply everything out?" Indeed, one can do this, and if one does it for the first several terms, one will get the same answer as the Baker–Campbell–Hausdorff formula. However, there is a serious problem with this approach, namely: How does one know that the terms in such an expansion are expressible in terms of commutators? Consider, for example, the quadratic term. It is clear that this will be a linear combination of X^2, Y^2, XY, and YX. However, to be expressible in terms of commutators, it must actually be a constant times $(XY - YX)$. Of course, for the quadratic term, one can just multiply it out and see, and, indeed, one gets $\frac{1}{2}(XY - YX) = \frac{1}{2}[X, Y]$. However, it is far from clear how to prove that a similar result occurs for all the higher terms. (See Exercise 6.) Although it is possible (but not easy) to prove directly that all terms in the expansion of $\log(e^X e^Y)$ are expressible in terms of commutators (Proposition 1 in Section V.5 of Jacobson (1962)), this is not the approach we will take.

We now state an important corollary of the Baker–Campbell–Hausdorff theorem.

Corollary 3.4. *Let G be a matrix Lie group and $\mathfrak{g}$ its Lie algebra. Suppose that $\phi : \mathfrak{g} \to \mathfrak{gl}(n;\mathbb{C})$ is a Lie algebra homomorphism. Then, for all sufficiently small X and Y in $\mathfrak{g}$, $\log\left(e^X e^Y\right)$ is in $\mathfrak{g}$, and*

$$\phi\left[\log(e^X e^Y)\right] = \log\left(e^{\phi(X)} e^{\phi(Y)}\right). \tag{3.7}$$

Proof. The proof uses the same reasoning as in (3.3). Note that if X and Y lie in some Lie algebra $\mathfrak{g}$, then ad_X and ad_Y will leave $\mathfrak{g}$ invariant, and, therefore, so will $g(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y})(Y)$. Thus, whenever formula (3.6) holds, $\log(e^X e^Y)$ will lie in $\mathfrak{g}$. It remains only to verify (3.7). The idea is that if ϕ is a Lie algebra homomorphism, then it will take a big, messy expression involving "ad" and X and Y, and turn it into the same expression with X and Y replaced by $\phi(X)$ and $\phi(Y)$.

More precisely, since ϕ is a Lie algebra homomorphism,

$$\phi[Y, X] = [\phi(Y), \phi(X)]$$

or

$$\phi(\mathrm{ad}_Y(X)) = \mathrm{ad}_{\phi(Y)}(\phi(X)).$$

More generally,

$$\phi((\mathrm{ad}_Y)^n(X)) = \left(\mathrm{ad}_{\phi(Y)}\right)^n(\phi(X)).$$

This being the case,

$$\begin{aligned}
\phi\left(e^{t\mathrm{ad}_Y}(X)\right) &= \sum_{m=0}^{\infty} \frac{t^m}{m!} \phi((\mathrm{ad}_Y)^m(X)) \\
&= \sum_{m=0}^{\infty} \frac{t^m}{m!} \left(\mathrm{ad}_{\phi(Y)}\right)^m(\phi(X)) \\
&= e^{t\mathrm{ad}_{\phi(Y)}}(\phi(X)).
\end{aligned}$$

Similarly,

$$\phi\left(\left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}\right)(Y)\right) = e^{\mathrm{ad}_{\phi(X)}} e^{t\mathrm{ad}_{\phi(Y)}}(\phi(Y)).$$

Assume now that X and Y are small enough that the Baker–Campbell–Hausdorff formula applies to X and Y and to $\phi(X)$ and $\phi(Y)$. Then, using the linearity of the integral and reasoning similar to the above, we have

$$\begin{aligned}
\phi[\log\left(e^X e^Y\right)] &= \phi(X) + \int_0^1 \sum_{m=0}^{\infty} a_m \phi\left[\left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y} - I\right)^m(Y)\right] dt \\
&= \phi(X) + \int_0^1 \sum_{m=0}^{\infty} a_m \left(e^{\mathrm{ad}_{\phi(X)}} e^{t\mathrm{ad}_{\phi(Y)}} - I\right)^m(\phi(Y))\, dt \\
&= \log\left(e^{\phi(X)} e^{\phi(Y)}\right).
\end{aligned}$$

This is what we wanted to show. □

3.3 The Derivative of the Exponential Mapping

Before coming to the proof of the Baker–Campbell–Hausdorff formula itself, we will obtain a result concerning derivatives of the exponential mapping. This result is valuable in its own right and will play a central role in our proof of the Baker–Campbell–Hausdorff formula.

Observe that if X and Y commute, then

$$e^{X+tY} = e^X e^{tY}$$

and so

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

In general, X and Y do not commute, and

$$\left.\frac{d}{dt} e^{X+tY}\right|_{t=0} \neq e^X Y.$$

(However, see Exercise 4.) This, as it turns out, is an important point. In particular, note that in the language of multivariate calculus,

$$\left.\frac{d}{dt} e^{X+tY}\right|_{t=0} = \begin{cases} \text{directional derivative of "exp" at } X, \\ \text{in the direction of } Y \end{cases}. \tag{3.8}$$

Thus, computing the left-hand side of (3.8) is the same as computing all of the directional derivatives of the (matrix-valued) function "exp." We expect the directional derivative to be a linear function of Y, for each fixed X.

Now, the function

$$\frac{1-e^{-z}}{z} = \frac{1-(1-z+\frac{z^2}{2!}-\cdots)}{z}$$

is an entire analytic function of z, even at $z=0$, and is given by the power series

$$\frac{1-e^{-z}}{z} = \sum_{k=0}^{\infty} (-1)^k \frac{z^k}{(k+1)!} = 1 - \frac{z}{2!} + \frac{z^2}{3!} - \cdots.$$

This series (which has infinite radius of convergence) makes sense when z is replaced by a linear operator A on some finite-dimensional vector space.

Theorem 3.5 (Derivative of Exponential). *Let X and Y be $n \times n$ complex matrices. Then,*

$$\begin{aligned} \left.\frac{d}{dt} e^{X+tY}\right|_{t=0} &= e^X \left\{ \frac{I - e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}(Y) \right\} \\ &= e^X \left\{ Y - \frac{[X,Y]}{2!} + \frac{[X,[X,Y]]}{3!} - \cdots \right\}. \end{aligned} \tag{3.9}$$

More generally, if $X(t)$ is a smooth matrix-valued function, then

$$\frac{d}{dt}e^{X(t)} = e^{X(t)}\left\{\frac{I - e^{-\mathrm{ad}_{X(t)}}}{\mathrm{ad}_{X(t)}}\left(\frac{dX}{dt}\right)\right\}. \tag{3.10}$$

Note that the directional derivative in (3.9) is indeed linear in Y for each fixed X. Note also that (3.9) is just a special case of (3.10), by taking $X(t) = X + tY$ and evaluating at $t = 0$.

Furthermore, observe that if X and Y commute, then only the first term in the series (3.9) survives. In that case, we obtain $\frac{d}{dt}e^{X+tY}\big|_{t=0} = e^X Y$, as expected.

The formula for the derivative of the exponential mapping is well known. The proof here follows that of Tuynman [The Derivation of the Exponential Map of Matrices, *Amer. Math. Monthly* **102** (1995), 818-819].

Proof. I prove only form (3.9); then, (3.10) follows by elementary calculus. For any $n \times n$ matrices X and Y, set

$$\Delta(X,Y) = \frac{d}{dt}e^{X+tY}\bigg|_{t=0}.$$

I leave it as an exercise (Exercise 3) to show that $\exp : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is a continuously differentiable map. This implies that $\Delta(X,Y)$ is jointly continuous in X and Y and that it is linear in Y for each fixed X (by a basic property of continuously differentiable functions of several variables).

Now, for every positive integer m, we have

$$e^{X+tY} = \left[\exp\left(\frac{X}{m} + t\frac{Y}{m}\right)\right]^m. \tag{3.11}$$

Thus, applying the product rule (extended to m factors), we will get m terms, in each of which $m-1$ of the factors in (3.11) are simply evaluated at $t = 0$ and the remaining factor is differentiated at $t = 0$. So, we get

$$\begin{aligned}
\frac{d}{dt}e^{X+tY}\bigg|_{t=0} &= \sum_{k=0}^{m-1}\exp\left(\frac{X}{m}\right)^{m-k-1}\left[\frac{d}{dt}\exp\left(\frac{X}{m}+t\frac{Y}{m}\right)\bigg|_{t=0}\right]\exp\left(\frac{X}{m}\right)^k \\
&= \exp\left(\frac{m-1}{m}X\right)\sum_{k=0}^{m-1}\exp\left(\frac{X}{m}\right)^{-k}\Delta\left(\frac{X}{m},\frac{Y}{m}\right)\exp\left(\frac{X}{m}\right)^k \\
&= \exp\left(\frac{m-1}{m}X\right)\frac{1}{m}\sum_{k=0}^{m-1}\exp\left(-\frac{\mathrm{ad}_X}{m}\right)^k\left(\Delta\left(\frac{X}{m},Y\right)\right). \qquad (3.12)
\end{aligned}$$

In the third equality, we have used the linearity of $\Delta(X,Y)$ in Y and the relationship between Ad and ad.

The left-hand side of (3.12) is equal to the right-hand side for each fixed m and thus the left-hand side is equal to the limit as $m \to \infty$ of the right-hand

side. Let us consider what happens as $m \to \infty$ in the last line of (3.12). The factor in front tends to $\exp(X)$. Since $\Delta(X,Y)$ is jointly continuous in X and Y, the expression $\Delta(X/m,Y)$ tends to $\Delta(0,Y)$, where it is easily verified that $\Delta(0,Y) = Y$. Thus, it remains only to analyze the behavior of

$$\frac{1}{m}\sum_{k=0}^{m-1} \exp\left(-\frac{\mathrm{ad}_X}{m}\right)^k.$$

This is taken care of in the following lemma.

Lemma 3.6. *For any $n \times n$ matrix X, we have*

$$\lim_{m\to\infty} \frac{1}{m}\sum_{k=0}^{m-1} \exp\left(-\frac{\mathrm{ad}_X}{m}\right)^k = \frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}. \tag{3.13}$$

Once this lemma is established, we take the limit as $m \to \infty$ everywhere and we are done. (Note that the quantities in (3.13) are linear operators on a finite-dimensional vector space, namely $M_n(\mathbb{C})$, thus essentially $n^2 \times n^2$ matrices. The operation of multiplying a $n^2 \times n^2$ matrix by a n^2-component vector is jointly continuous in the two variables. Thus, we are justified in separately evaluating the limit in (3.13) and the limit in $\Delta(X/m,Y)$.) We now turn to the proof of Lemma 3.6.

Proof. Let us first reason at a formal level (i.e., pretending that ad_X is a nonzero number instead of an operator). Then, using the usual formula for the sum of a finite geometric series would give

$$\frac{1}{m}\sum_{k=0}^{m-1} \exp\left(-\frac{\mathrm{ad}_X}{m}\right)^k = \frac{1}{m}\frac{1-\exp(-\mathrm{ad}_X)}{1-\exp(-\mathrm{ad}_X/m)} \xrightarrow{m\to\infty} \frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}.$$

To give a rigorous argument, we write $\exp(-\mathrm{ad}_X/m)^k$ as $\exp(-k\mathrm{ad}_X/m)$ and compute

$$\begin{aligned}\frac{1}{m}\sum_{k=0}^{m-1} \exp\left(-\frac{k\mathrm{ad}_X}{m}\right) &= \sum_{i=0}^{\infty}\frac{1}{m}\sum_{k=0}^{m-1}\frac{1}{i!}\left(-\frac{k\mathrm{ad}_X}{m}\right)^i \\ &= \sum_{i=0}^{\infty}\left[\frac{1}{m}\sum_{k=0}^{m-1}\left(\frac{k}{m}\right)^i\right]\frac{(-1)^i}{i!}(\mathrm{ad}_X)^i.\end{aligned}$$

(We have interchanged the finite sum over k with the infinite sum over i.) Now, we may recognize the quantity in square brackets in the last expression as the Riemann sum approximation to the integral $\int_0^1 x^i dx$, where the value of the integral is $1/(i+1)$. So, as m tends to infinity, the quantity in square brackets tends to $1/(i+1)$. Furthermore, since the function x^i is increasing on the interval $[0,1]$, the value of the expression in square brackets will be less than the value of the integral, for each m.

Now, each term in the series is a linear operator on $M_n(\mathbb{C})$, which we can think of as an $n^2 \times n^2$ matrix. The norm of each term (as $n^2 \times n^2$ matrices) is bounded by

$$\frac{1}{i+1}\frac{1}{i!}\|\mathrm{ad}_X\|^i. \tag{3.14}$$

Now, each entry in an $n^2 \times n^2$ matrix is smaller (in absolute value) than the norm of the matrix, as is easily verified. Thus since the sum of the quantities in (3.14) is finite, we can apply the dominated convergence theorem to each entry of the matrix-valued sum to justify interchanging the limit $m \to \infty$ with the infinite sum over i. This gives

$$\lim_{m\to\infty}\frac{1}{m}\sum_{k=0}^{m-1}\exp(-k\mathrm{ad}_X/m) = \sum_{i=0}^{\infty}\frac{(-1)^i(\mathrm{ad}_X)^i}{(i+1)!} = \frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}.$$

□

This concludes the proof of Theorem 3.5. □

3.4 Proof of the Baker–Campbell–Hausdorff Formula

We now turn to the proof of the Baker–Campbell–Hausdorff formula itself. Define

$$Z(t) = \log\left(e^X e^{tY}\right)$$

If X and Y are sufficiently small, then $Z(t)$ is defined for $0 \le t \le 1$. It is left as an exercise to verify that $Z(t)$ is smooth. Our goal is to compute $Z(1)$.

By definition,

$$e^{Z(t)} = e^X e^{tY}$$

so that

$$e^{-Z(t)}\frac{d}{dt}e^{Z(t)} = \left(e^X e^{tY}\right)^{-1} e^X e^{tY} Y = Y.$$

On the other hand, by Theorem 3.5,

$$e^{-Z(t)}\frac{d}{dt}e^{Z(t)} = \left\{\frac{I-e^{-\mathrm{ad}_{Z(t)}}}{\mathrm{ad}_{Z(t)}}\right\}\left(\frac{dZ}{dt}\right).$$

Hence,

$$\left\{\frac{I-e^{-\mathrm{ad}_{Z(t)}}}{\mathrm{ad}_{Z(t)}}\right\}\left(\frac{dZ}{dt}\right) = Y.$$

If X and Y are small enough, then $Z(t)$ will also be small, so that $[I - e^{-\mathrm{ad}_{Z(t)}}]/\mathrm{ad}_{Z(t)}$ will be close to the identity and thus invertible. So,

$$\frac{dZ}{dt} = \left\{\frac{I-e^{-\mathrm{ad}_{Z(t)}}}{\mathrm{ad}_{Z(t)}}\right\}^{-1}(Y). \tag{3.15}$$

Recall that $e^{Z(t)} = e^X e^{tY}$. Applying the homomorphism "Ad" gives

$$\mathrm{Ad}_{e^{Z(t)}} = \mathrm{Ad}_{e^X} \mathrm{Ad}_{e^{tY}}.$$

By the relationship between "Ad" and "ad" (Proposition 2.25), this becomes

$$e^{\mathrm{ad}_{Z(t)}} = e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}$$

or

$$\mathrm{ad}_{Z(t)} = \log\left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}\right).$$

Plugging this into (3.15) gives

$$\frac{dZ}{dt} = \left\{ \frac{I - \left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}\right)^{-1}}{\log\left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}\right)} \right\}^{-1} (Y). \tag{3.16}$$

Now, observe that

$$g(z) = \left\{ \frac{1 - z^{-1}}{\log z} \right\}^{-1}$$

so, formally, (3.16) is the same as

$$\frac{dZ}{dt} = g\left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}\right)(Y). \tag{3.17}$$

It is not hard to show that this formal argument is actually correct.

Now we are done, for if we note that $Z(0) = X$ and integrate (3.17), we get

$$Z(1) = X + \int_0^1 g(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y})(Y)\, dt,$$

which is the Baker–Campbell–Hausdorff formula.

3.5 The Series Form of the Baker–Campbell–Hausdorff Formula

Let us see how to get the first few terms of the series form of Baker–Campbell–Hausdorff from the integral form. Recall the function

$$\begin{aligned} g(z) &= \frac{z \log z}{z-1} \\ &= \frac{[1+(z-1)]\left[(z-1) - \frac{(z-1)^2}{2} + \frac{(z-1)^3}{3} - \cdots\right]}{(z-1)} \\ &= [1+(z-1)]\left[1 - \frac{z-1}{2} + \frac{(z-1)^2}{3} + \cdots\right]. \end{aligned}$$

Multiplying this out and combining terms gives

$$g(z) = 1 + \frac{1}{2}(z-1) - \frac{1}{6}(z-1)^2 + \cdots.$$

The closed-form expression for g is

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

Meanwhile,

$$\begin{aligned} & e^{\mathrm{ad}_X} e^{t\,\mathrm{ad}_Y} - I \\ &= \left(I + \mathrm{ad}_X + \frac{(\mathrm{ad}_X)^2}{2} + \cdots\right)\left(I + t\,\mathrm{ad}_Y + \frac{t^2(\mathrm{ad}_Y)^2}{2} + \cdots\right) - I \\ &= \mathrm{ad}_X + t\,\mathrm{ad}_Y + t\,\mathrm{ad}_X\,\mathrm{ad}_Y + \frac{(\mathrm{ad}_X)^2}{2} + \frac{t^2(\mathrm{ad}_Y)^2}{2} + \cdots. \end{aligned}$$

The crucial observation here is that $e^{\mathrm{ad}_X} e^{t\,\mathrm{ad}_Y} - I$ has no zero-order term, just first order and higher in ad_X and ad_Y. Thus, $\left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y} - I\right)^m$ will contribute only terms of degree m or higher in ad_X and/or ad_Y.

We have, then, up to degree 2 in ad_Y and ad_Y,

$$\begin{aligned} g\left(e^{\mathrm{ad}_X} e^{t\,\mathrm{ad}_Y}\right) &= I + \frac{1}{2}\left[\mathrm{ad}_X + t\,\mathrm{ad}_Y + t\,\mathrm{ad}_X\,\mathrm{ad}_Y + \frac{(\mathrm{ad}_X)^2}{2} + \frac{t^2(\mathrm{ad}_Y)^2}{2} + \cdots\right. \\ &\quad \left. - \frac{1}{6}\left[\mathrm{ad}_X + t\,\mathrm{ad}_Y + \cdots\right]^2 + \cdots\right. \\ &= I + \frac{1}{2}\mathrm{ad}_X + \frac{t}{2}\mathrm{ad}_Y + \frac{t}{2}\mathrm{ad}_X\,\mathrm{ad}_Y + \frac{(\mathrm{ad}_X)^2}{4} + \frac{t^2(\mathrm{ad}_Y)^2}{4} \\ &\quad - \frac{1}{6}\left[(\mathrm{ad}_X)^2 + t^2(\mathrm{ad}_Y)^2 + t\,\mathrm{ad}_X\,\mathrm{ad}_Y + t\,\mathrm{ad}_Y\,\mathrm{ad}_X\right] \\ &\quad + \text{ higher-order terms.} \end{aligned}$$

We now apply $g\left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}\right)$ to Y and integrate. So (neglecting higher-order terms) using Baker–Campbell–Hausdorff and noting that any term with ad_Y acting first is zero:

$$\begin{aligned} & \log\left(e^X e^Y\right) \\ &= X + \int_0^1 \left[Y + \frac{1}{2}[X,Y] + \frac{1}{4}[X,[X,Y]] - \frac{1}{6}[X,[X,Y]] - \frac{t}{6}[Y,[X,Y]]\right] dt \\ &= X + Y + \frac{1}{2}[X,Y] + \left(\frac{1}{4} - \frac{1}{6}\right)[X,[X,Y]] - \frac{1}{6}\int_0^1 t\,dt\,[Y,[X,Y]]. \end{aligned}$$

Thus, if we do the algebra, we end up with

$$\log(e^X e^Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \text{ higher order terms.}$$

This is the expression in (3.2).

3.6 Group Versus Lie Algebra Homomorphisms

Recall Theorem 2.21, which says that given matrix Lie groups G and H and a Lie group homomorphism $\Phi : G \to H$, there is a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$ such that $\Phi(\exp X) = \exp \phi(X)$ for all $X \in \mathfrak{g}$. In this section, we prove a converse to this result in the case that G is *simply connected.*

Theorem 3.7. *Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. If G is simply connected, then there exists a unique Lie group homomorphism $\Phi : G \to H$ such that $\Phi(\exp X) = \exp(\phi(X))$ for all $X \in \mathfrak{g}$.*

This has the following corollary.

Corollary 3.8. *Suppose G and H are simply-connected matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. If $\mathfrak{g}$ is isomorphic to $\mathfrak{h}$, then G is isomorphic to H.*

Proof. Let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra isomorphism. By Theorem 3.7, there exists an associated Lie group homomorphism $\Phi : G \to H$. Since $\phi^{-1} : \mathfrak{h} \to \mathfrak{g}$ is also a Lie algebra homomorphism, there is a corresponding Lie group homomorphism $\Psi : H \to G$. We want to show that Φ and Ψ are inverses of each other.

However, the Lie algebra map associated with the composition is the composition of the Lie algebra maps (Point 3 of Theorem 2.21), which is the identity. So, by Corollary 2.32, $\Phi \circ \Psi = I_H$. Similarly, $\Psi \circ \Phi = I_G$. □

We now proceed with the proof of Theorem 3.7.

Proof. Step 1: Define Φ in a neighborhood of the identity.

Corollary 2.29 says that the exponential mapping for G has a local inverse which maps a neighborhood V of the identity into the Lie algebra $\mathfrak{g}$. If we make V small enough, then we can also assume that for all $A, B \in V$, we have $\log A$ and $\log B$ small enough that the Baker–Campbell–Hausdorff theorem applies to them. We fix one such neighborhood V for the remainder of the proof.

On this neighborhood V, we can *define* $\Phi : V \to H$ by

$$\Phi(A) = \exp\{\phi(\log A)\};$$

that is, on V, we have

$$\Phi = \exp \circ \phi \circ \log .$$

(Note that if there is to be a homomorphism Φ as in the theorem, then Φ must be given by this formula.)

Step 2: Define Φ along a path.

Recall that part of what it means for G to be simply connected is that it is connected. Recall also that when we say G is connected, we really mean that G is path-connected. (By now, we know that G is an embedded submanifold of $\mathsf{GL}(n;\mathbb{C})$. This means that G is locally path-connected and, thus, that G is connected if and only it is path-connected.) Thus, for any $A \in G$, there exists a path $A(t) \in G$ with $A(0) = I$ and $A(1) = A$. A standard argument using the compactness of the interval $[0,1]$ shows that there exists numbers $0 = t_0 < t_1 < t_2 \cdots < t_m = 1$ such that for all s and t satisfying $t_i \leq s \leq t \leq t_{i+1}$ (for some i), we have

$$A(t)A(s)^{-1} \in V. \tag{3.18}$$

In particular, since $t_0 = 0$ and $A(0) = I$, we have $A(t_1) \in V$. We now write $A = A(1)$ in the form

$$A = \left[A(1)A(t_{m-1})^{-1}\right]\left[A(t_{m-1})A(t_{m-2})^{-1}\right]\cdots[A(t_2)A(t_1)^{-1}]A(t_1).$$

Since Φ is supposed to be a homomorphism, it is reasonable to "define" $\Phi(A)$ by

$$\Phi(A) = \Phi(A(1)A(t_{m-1})^{-1})\cdots\Phi(A(t_2)A(t_1)^{-1})\Phi(A(t_1)), \tag{3.19}$$

where each factor on the right is defined as in Step 1.

Step 3: Prove independence of the partition.

For this definition of $\Phi(A)$ to be valid, we must show that the value of $\Phi(A)$ is independent of the choice of the path and independent of the choice of partition $(t_0, \ldots, t_m)$ for a given path. We address independence of the partition first. It is in this step (and only in this step) that we use the Baker–Campbell–Hausdorff theorem. To establish independence of partition, we first show that passing from a particular partition to a refinement of that partition does not change the result. (A refinement of a partition is one which contains all the points of the original partition, together with some other ones.) Note that if a given partition satisfies the condition (3.18), then any refinement of that partition also satisfies this condition.

Suppose, now, that we insert an extra partition point s between t_i and t_{i+1}. Then, the factor $\Phi(A(t_{i+1})A(t_i)^{-1})$ in (3.19) will be replaced by

$$\Phi(A(t_{i+1})A(s)^{-1})\Phi(A(s)A(t_i)^{-1}).$$

Since s is between t_i and t_{i+1}, the condition (3.18) on the original partition guarantees that $A(t_{i+1})A(s)^{-1}$ and $A(s)A(t_i)^{-1}$, in addition to $A(t_{i+1})A(t_i)^{-1}$, are all in V. Now, it follows from Corollary 3.4 to the Baker–Campbell–Hausdorff formula that Φ, as defined in Step 1, is a "local homomorphism";

that is, $\Phi(AB) = \Phi(A)\Phi(B)$ for all A and B sufficiently close to the identity. (When applying the corollary, write A as e^X and B as e^Y.) This means that

$$\Phi(A(t_{i+1})A(t_i)^{-1}) = \Phi(A(t_{i+1})A(s)^{-1})\Phi(A(s)A(t_i)^{-1})$$

and, thus, the value of $\Phi(A)$ is unchanged by the addition of the extra partition point. By repeating this argument, we see that the value of $\Phi(A)$ does not change by the addition of any finite number of points to the partition.

Now, given any two partitions, they have a common refinement, namely their union. The above argument shows that the value of $\Phi(A)$ computed from the first partition is the same as for the common refinement, which is the same as for the second partition. This shows independence of the partition.

Step 4: Prove independence of the path.

Having proved that the value of $\Phi(A)$ is independent of the partition for a fixed path, we now need to prove that $\Phi(A)$ is independent of the choice of path. It is in this step that we use the simple connectedness of G. Suppose $A_0(t)$ and $A_1(t)$ are two paths joining the identity to some $A \in G$. Then, since G is simply connected, a standard topological argument shows that A_0 and A_1 are homotopic with endpoints fixed. This means that there exists a continuous map $A : [0,1] \times [0,1] \to G$ with

$$A(0,t) = A_0(t), \quad A(1,t) = A_1(t)$$

for all $t \in [0,1]$ and also

$$A(s,0) = I, \quad A(s,1) = A$$

for all $s \in [0,1]$.

The compactness of $[0,1] \times [0,1]$ guarantees that there exists an integer N such that for all (s,t) and (s',t') in $[0,1] \times [0,1]$ with $|s - s'| < 2/N$ and $|t - t'| < 2/N$, we have $A(s,t)A(s',t')^{-1} \in V$. We now employ a standard topological trick to deform A_0 "a little bit at a time" into A_1. This means that we define a sequence $B_{k,l}$ of paths, with $k = 0, \ldots, N-1$ and $l = 0, \ldots, N$. We define these paths so that $B_{k,l}(t)$ coincides with $A((k+1)/N, t)$ for t between 0 and $(l-1)/N$, and $B_{k,l}(t)$ coincides with $A(k/N, t)$ for t between l/N and 1. For t between $(l-1)/N$ and l/N, we define $B_{k,l}(t)$ to coincide with the values of $A(\cdot,\cdot)$ on the path that goes "diagonally" in the (s,t)-plane, as indicated in Figure 3.1. (I could write the formula for $B_{k,l}$ in this interval, but the picture is clearer than the formula.) When computing $B_{k,0}$, there are no t-values between 0 and $(l-1)/N$, so $B_{k,0}(t) = A(k/n, t)$ for all $t \in [0,1]$. In particular, $B_{0,0}(t) = A_0(t)$.

We think of deforming the path A_0 into A_1 in steps. First, we deform $A_0 = B_{0,0}$ into $B_{0,1}$ and then into $B_{0,2}$, $B_{0,3}$, and so on until we reach $B_{0,N}$, which we then deform into $B_{1,0}$ and then into $B_{1,1}, \ldots, B_{1,N}$. We continue this process until we reach $B_{N-1,N}$, which we finally deform into A_1.We want to show that the value of $\Phi(A)$ computed along each of these paths is the

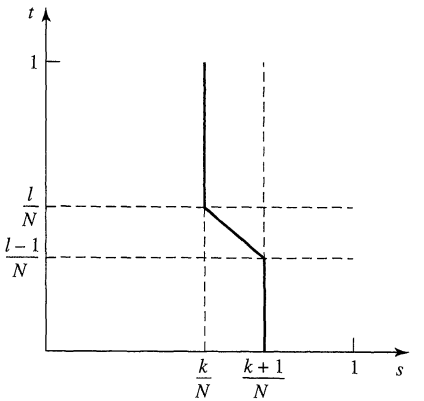


Fig. 3.1. The path $B_{k,l}$

same as the value of $\Phi(A)$ computed along the next one. Now, we note that for $k < l$, $B_{k,l}(t)$ and $B_{k,l+1}(t)$ are the same except for t's in the interval $[(l-1)/N, (l+1)/N]$. We then exploit the independence of the partition that we have just verified. We may choose any partition we like, provided that the condition (3.18) is satisfied. So, for both $B_{k,l}$ and $B_{k,l+1}$, we choose the partition points to be

$$0, \frac{1}{N}, \cdots, \frac{l-1}{N}, \frac{l+1}{N}, \frac{l+2}{N}, \cdots, 1.$$

The way we have chosen N guarantees that this is a valid partition. (Check!)

Now, note (from (3.19)) that the value of $\Phi(A)$ depends only on the values of the path at the partition points. We have chosen our partition in such a way that the values of $B_{k,l}$ and $B_{k,l+1}$ are identical at all the partition points, and, therefore, the value of $\Phi(A)$ is the same for these two paths. A similar argument shows that the value of $\Phi(A)$ computed along $B_{k,N}$ is the same as along $B_{k+1,0}$. (Note that $B_{k,N}(1) = B_{k+1,0}(1) = A$.) Thus, the value of $\Phi(A)$ is the same for each path from $A_0 = B_{0,0}$ all the way to $B_{N-1,N}$ and then (by the same argument) the same as A_1. This shows independence of the path.

Step 5: Prove that Φ is a homomorphism and is properly related to ϕ.

The proof that Φ is a homomorphism is fairly straightforward and is left to the reader. See Exercise 10. It then remains only to verify that Φ has the proper relationship to ϕ. However, since Φ is defined near the identity to be $\Phi = \exp \circ \phi \circ \log$, we see that

$$\frac{d}{dt}\Phi\left(e^{tX}\right)\Big|_{t=0} = \frac{d}{dt}e^{t\phi(X)}\Big|_{t=0} = \phi(X).$$

Thus, ϕ is the Lie algebra homomorphism associated to the Lie group homomorphism Φ.

This completes the proof of Theorem 3.7. □

We will return to the issue of the relationship between Lie group and Lie algebra homomorphisms in Section 4.9 of the next chapter.

3.7 Covering Groups

Theorem 3.7 says that if G is simply connected, then every Lie algebra homomorphism for G can be exponentiated to give a Lie group homomorphism. If G is not simply connected, this will not (in general) be true. It is thus reasonable to look for another group $\tilde{G}$ that has the same Lie algebra as G but such that $\tilde{G}$ is simply connected. Such a group is called the universal covering group (or just the universal cover) of G.

Definition 3.9. *Let G be a connected Lie group. Then, a **universal covering group** (or **universal cover**) of G is a simply-connected Lie group H together with a Lie group homomorphism $\Phi : H \to G$ such that the associated Lie algebra homomorphism $\phi : \mathfrak{h} \to \mathfrak{g}$ is a Lie algebra isomorphism. The homomorphism Φ is called the **covering homomorphism** (or **projection map**).*

Here neither G nor H is assumed to be a *matrix* Lie group. As discussed later, the universal cover of a matrix Lie group may not be a matrix Lie group. For every connected Lie group, a universal cover exists and is unique up to "canonical isomorphism," as explained in the following theorem.

Theorem 3.10. *For any connected Lie group, a universal cover exists. If G is a connected Lie group and (H_1, Φ_1) and (H_2, Φ_2) are universal covers of G, then there exists a Lie group isomorphism $\Psi : H_1 \to H_2$ such that $\Phi_2 \circ \Psi = \Phi_1$.*

Appendix C gives a sketch of the proof of this result. The uniqueness part of the result is a consequence of Theorem 3.7 (Exercise 14).

Since the universal cover of a connected Lie group G is unique (up to canonical isomorphism), it is reasonable to speak of *the* universal cover $(\tilde{G}, \Phi)$ of G. Furthermore, if $\tilde{G}$ is a simply-connected Lie group and $\phi : \tilde{\mathfrak{g}} \to \mathfrak{g}$ is a Lie algebra isomorphism, then by Theorem 3.7 (which actually applies to all Lie groups, not just matrix Lie groups), we can construct an associated Lie group homomorphism $\Phi : \tilde{G} \to G$. Then $(\tilde{G}, \Phi)$ is a universal cover of G. Since ϕ is an isomorphism, we can use ϕ to identify $\tilde{\mathfrak{g}}$ with $\mathfrak{g}$. Thus, in slightly less formal terms, we may define the notion of universal cover as follows: *The universal cover of a Lie group G is the unique simply-connected group $\tilde{G}$ such that the Lie algebra of $\tilde{G}$ is equal to the Lie algebra of G.* (Implicit in this form of the definition is that we have chosen some particular isomorphism $\phi : \tilde{\mathfrak{g}} \to \mathfrak{g}$ to identify $\tilde{\mathfrak{g}}$ with $\mathfrak{g}$.) If we adopt this form of the definition, then the covering homomorphism is defined as *the unique Lie group homomorphism*

$\Phi : \tilde{G} \to G$ *such that the associated Lie algebra map* $\phi : \mathfrak{g} \to \mathfrak{g}$ *is the identity.* (The existence of Φ is by Theorem 3.7.)

The study of universal covering groups is one of the places where we pay a price for our decision to consider only *matrix* Lie groups: the universal cover of a matrix Lie group may not be a matrix Lie group. That is, even if G is a matrix Lie group, the universal cover $\tilde{G}$ of G may not be isomorphic to any matrix Lie group. For example, the universal cover of $\mathsf{SL}(n;\mathbb{R})$ $(n \geq 2)$ is not a matrix Lie group. See Section C.3.

One can also consider covering groups that are not universal covers. A **covering group** of a connected Lie group G is a connected Lie group H (not necessarily simply connected) together with a Lie group homomorphism $\Phi : H \to G$ such that the associated Lie algebra homomorphism $\phi : \mathfrak{h} \to \mathfrak{g}$ is an isomorphism. There may be several nonisomorphic covering groups of a given group G, and these different covers may have different fundamental groups.

Let us now consider some examples of universal covers.

Example 1: $G = S^1$. In this case the universal cover is $\mathbb{R}$ and the covering homomorphism is the map $\Phi : \mathbb{R} \to S^1$ given by $\theta \to e^{i\theta}$.

Example 2: $G = \mathsf{SO}(3)$. In this case, the universal cover is $\mathsf{SU}(2)$ and the covering homomorphism is the map Φ described in Section 1.6. (See also Section 4.9.)

Example 3: $G = \mathsf{U}(n)$. In this case, the universal cover is $\mathbb{R} \times \mathsf{SU}(n)$ and the covering homomorphism is the map $\Phi : \mathbb{R} \times \mathsf{SU}(n) \to \mathsf{U}(n)$ given by

$$\Phi(\theta, U) = e^{i\theta} U. \tag{3.20}$$

Note that since both $\mathbb{R}$ and $\mathsf{SU}(n)$ are simply connected (Appendix E), $\mathbb{R} \times \mathsf{SU}(n)$ is simply connected. It is straightforward to check (Exercise 15) that the Lie algebra map associated to Φ is indeed a Lie algebra isomorphism in this case.

Example 4 $G = \mathsf{SO}(n)$. For $n \geq 3$, the universal cover of $\mathsf{SO}(n)$ is a double cover (i.e., the projection map Φ is two-to-one). This reflects that the fundamental group of $\mathsf{SO}(n)$ $(n \geq 3)$ has two elements. The universal cover of $\mathsf{SO}(n)$ is called $\mathsf{Spin}(n)$ and may be constructed as a certain group of invertible elements in the **Clifford algebra** over $\mathbb{R}^n$. See Bröcker and tom Dieck (1985), Chapter I, Section 6, especially Propositions I.6.17 and I.6.19. In particular, $\mathsf{Spin}(n)$ *is* a matrix Lie group. The cases $n = 3$ and $n = 4$ are special. For $n = 3$, we have (Example 2) $\mathsf{Spin}(3) \cong \mathsf{SU}(2)$ and for $n = 4$ we have $\mathsf{Spin}(4) \cong \mathsf{SU}(2) \times \mathsf{SU}(2)$.

In all of these examples, the universal cover turns out to be, again, a matrix Lie group. More generally, it is possible to show that the universal cover of a compact matrix Lie group is always, again, a matrix Lie group (not necessarily compact).

3.8 Subgroups and Subalgebras

Suppose that G is a matrix Lie group, that H is another matrix Lie group, and that $H \subset G$. Then, certainly, the Lie algebra $\mathfrak{h}$ of H will be a subalgebra of the Lie algebra $\mathfrak{g}$ of G. Does this go the other way around? That is, given a matrix Lie group G with Lie algebra $\mathfrak{g}$ and a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, is there a matrix Lie group H whose Lie algebra is $\mathfrak{h}$?

In the case of the Heisenberg group, the answer is yes. This holds because for the Heisenberg group, the exponential mapping is one-to-one and onto and the Baker–Campbell–Hausdorff formula takes a particularly simple form. (See Exercise 16.)

In general, however, there may not be any matrix Lie group H corresponding to a given subalgebra $\mathfrak{h}$. For example, let $G = \mathsf{GL}(2;\mathbb{C})$ and let

$$\mathfrak{h} = \left\{ \begin{pmatrix} it & 0 \\ 0 & ita \end{pmatrix} \middle| \, t \in \mathbb{R} \right\}, \tag{3.21}$$

where a is irrational. This is a one-dimensional real subalgebra of $\mathfrak{g} = \mathfrak{gl}(2;\mathbb{C})$. If there were going to be a matrix Lie group H with Lie algebra $\mathfrak{h}$, then H would contain the set of all exponentials of elements of $\mathfrak{h}$, namely

$$H_0 = \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{ita} \end{pmatrix} \middle| \, t \in \mathbb{R} \right\}. \tag{3.22}$$

To be a matrix Lie group, H would have to be closed in $\mathsf{GL}(2;\mathbb{C})$, and so it would contain the closure of H_0, which (Exercise 1 in Chapter 1) is the set

$$H_1 = \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{is} \end{pmatrix} \middle| \, s, t \in \mathbb{R} \right\}.$$

However, then, the Lie algebra of H would have to contain the Lie algebra of H_1, which is two dimensional!

Fortunately, all is not lost. We can still get a subgroup H for each subalgebra $\mathfrak{h}$ if we weaken the condition that H be a matrix Lie group. In the above example, the subgroup we want is H_0, even though H_0 is not a matrix Lie group.

Definition 3.11. *If H is **any** subgroup of $\mathsf{GL}(n;\mathbb{C})$, define the Lie algebra $\mathfrak{h}$ of H to be the set of all matrices X such that*

$$e^{tX} \in H$$

for all real t.

Definition 3.12. *If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, then $H \subset G$ is a **connected Lie subgroup** of G if the following conditions are satisfied:*

1. *H is a subgroup of G.*

2. *The Lie algebra* $\mathfrak{h}$ *of* H *is a subspace of* $\mathfrak{g}$.
3. *Every element of* H *can be written in the form* $e^{X_1}e^{X_2}\cdots e^{X_m}$, *with* $X_1, \ldots, X_m \in \mathfrak{h}$.

Connected Lie subgroups are also called **analytic subgroups**. The group H_0 in (3.22) is a connected Lie subgroup of $\mathrm{GL}(2;\mathbb{C})$ whose Lie algebra is the algebra $\mathfrak{h}$ in (3.21). The word "connected" in the phrase "connected Lie subgroup" is justified by the following easy result.

Proposition 3.13. *If* G *is a matrix Lie group and* H *is a connected Lie subgroup of* G, *then* H *is path-connected. That is, any two points in* H *can be connected by a continuous path lying in* H.

Proof. As usual, it suffices to show that any element of H can be connected to the identity by a continuous path lying in H. If $h \in H$ then we write

$$h = e^{X_1}e^{X_2}\cdots e^{X_m}, \quad X_k \in \mathfrak{h},$$

as in Condition 3 in the definition. We consider the path $h(t)$ given by

$$h(t) = he^{-tX_m} = e^{X_1}e^{X_2}\cdots e^{(1-t)X_m}.$$

This path is continuous and lies in H, since (by the definition of $\mathfrak{h}$) e^{-tX_m} lies in H for all t. As t varies from 0 to 1, $h(t)$ connects the element h to the element $e^{X_1}e^{X_2}\cdots e^{X_{m-1}}$ of H. By applying this process m times, we can connect h to the identity. □

Proposition 3.14. *If* G *is a matrix Lie group with Lie algebra* $\mathfrak{g}$ *and* H *is a connected Lie subgroup of* G, *then the Lie algebra* $\mathfrak{h}$ *of* H *is a subalgebra of* $\mathfrak{g}$.

Proof. If $A \in H$ and $Y \in \mathfrak{h}$, then $\exp(tAYA^{-1}) = A\exp(tY)A^{-1}$ belongs to H for all real t. Thus, AYA^{-1} is, again, in $\mathfrak{h}$. Then, as in the proof of Point 3 of Theorem 2.18, if X and Y are in $\mathfrak{h}$ we have $e^{tX}Ye^{-tX}$ in $\mathfrak{h}$ for all t. Therefore, since $\mathfrak{h}$ is a vector space and (thus) a topologically closed subset of $M_n(\mathbb{C})$, we have

$$\begin{aligned} XY - YX &= \frac{d}{dt}e^{tX}Ye^{-tX}\bigg|_{t=0} \\ &= \lim_{h\to 0}\frac{e^{hX}Ye^{-hX} - Y}{h} \in \mathfrak{h}. \end{aligned}$$

□

We are now ready to state the main result of this section, which is our second major application of the Baker–Campbell–Hausdorff formula.

Theorem 3.15. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Let $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{g}$. Then, there exists a unique connected Lie subgroup H of G such that the Lie algebra of H is $\mathfrak{h}$. The subgroup H consists precisely of elements of the form*

$$e^{X_1} e^{X_2} \cdots e^{X_m}$$

with $X_1, \ldots, X_m \in \mathfrak{h}$.

The proof of this result is given at the end of this section.

Given a matrix Lie group G and a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, the associated connected Lie subgroup H *might* be a matrix Lie group. This will happen precisely if H is a closed subset of G. There are various conditions under which it can be proved that H is closed. For example, if $G = \mathrm{GL}(n;\mathbb{C})$ and $\mathfrak{h}$ is semisimple (Chapter 6), then H is automatically closed, and hence a matrix Lie group. (See Helgason (1978), Chapter II, Exercises and Further Results, D.)

If the Baker–Campbell–Hausdorff formula worked globally instead of only locally, the proof of this theorem would be easy. If the Baker–Campbell–Hausdorff formula converged for all X and Y, we could just define H to be the image of $\mathfrak{h}$ under the exponential mapping. In that case, the Baker–Campbell–Hausdorff formula would show that this image is a subgroup, since, then, we would have $e^{H_1} e^{H_2} = e^Z$, with $Z = H_1 + H_2 + \frac{1}{2}[H_1, H_2] + \cdots \in \mathfrak{h}$, provided that $H_1, H_2 \in \mathfrak{h}$ and that $\mathfrak{h}$ is a subalgebra. Unfortunately, the Baker–Campbell–Hausdorff formula is not convergent in general, and, in general, the image of $\mathfrak{h}$ under the exponential mapping is not a subgroup.

Proposition 3.16. *Suppose that G is a matrix Lie group with Lie algebra $\mathfrak{g}$ and suppose that $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$. Suppose that F is a connected matrix Lie group with Lie algebra $\mathfrak{f}$ and that $\Phi : F \to G$ is a Lie group homomorphism with the property that $\phi(\mathfrak{f}) = \mathfrak{h}$. (Here, ϕ is the Lie algebra homomorphism associated to Φ.) Then, the connected Lie subgroup of G with Lie algebra $\mathfrak{h}$ is equal to $\Phi(F)$ (the image of F under Φ).*

Proof. Let H be the connected Lie subgroup of G with Lie algebra $\mathfrak{h}$. Since F is connected, every element A of F can be written as $A = \exp X_1 \cdots \exp X_m$, $X_k \in \mathfrak{f}$. So, every element of $\Phi(F)$ can be written as $\exp \phi(X_1) \cdots \exp \phi(X_m)$, where, by assumption, $\phi(X_k) \in \mathfrak{h}$. This shows that $\Phi(F) \subset H$. Conversely, every element B of H can be written as $B = \exp Y_1 \cdots \exp Y_m$, with $Y_k \in \mathfrak{h} = \phi(\mathfrak{f})$. Choosing X_k's in $\mathfrak{f}$ with $\phi(X_k) = Y_k$, we have that $B = \Phi(\exp X_1) \cdots \Phi(\exp X_m)$. This shows that $H \subset \Phi(F)$. □

If H is a connected Lie subgroup of a matrix Lie group G, then the topology that H inherits as a subset of G may be quite pathological (e.g., not locally connected). However, we can define a different topology on H that is much nicer. For any $A \in H$ and any $\varepsilon > 0$, define

$$U_{A,\varepsilon} = \{ Ae^X \,\big|\, X \in \mathfrak{h} \text{ and } \|X\| < \varepsilon \}.$$

Now define a topology on H as follows: A set $U \subset H$ is open if for each $A \in U$ there exists $\varepsilon > 0$ such that $U_{A,\varepsilon} \subset U$. In this topology, two elements A and B of H are "close" if we can express B as $B = A \exp X$ with $X \in \mathfrak{h}$ and $\|X\|$ small. This topology is finer than the topology H inherits from G; that is, if $A, B \in H$ are close in the usual sense in G, then they are close in this new topology on H, but not vice versa.

If H is a connected Lie subgroup of G, then it can be shown that in this new topology, H is a topological manifold. Furthermore, H can be made into a smooth manifold by using the sets $U_{A,\varepsilon}$ as our basic coordinate neighborhoods and using the quantity X in the expression $A \exp X$ as our local coordinate. The product and inverse maps on H are smooth with respect to this smooth manifold structure, and so H can in this way be made into a Lie group.

We summarize these conclusions in the following theorem. It is not hard to prove this result by elaborating on the discussion in the previous two paragraphs. Compare the section "Lie Subgroups" in Chapter 3 of Warner (1983).

Theorem 3.17. *Suppose that G is a matrix Lie group and H a connected Lie subgroup of G. Then H can be given the structure of a Lie group in such a way that the inclusion of H into G is a Lie group homomorphism.*

Once H has been made into a Lie group, it has a Lie algebra in the sense of Appendix C. This Lie algebra is naturally isomorphic to the subalgebra $\mathfrak{h}$ we began with. Thus, Theorem 3.17 and Ado's Theorem (Theorem 2.40, which we have not proved) imply the following result.

Theorem 3.18. *Every finite-dimensional real Lie algebra is isomorphic to the Lie algebra of some Lie group.*

We now turn to the proof of Theorem 3.15.

Proof. Since G is assumed to be a matrix Lie group, we may as well assume that $G = \mathsf{GL}(n;\mathbb{C})$ so that $\mathfrak{g} = \mathsf{gl}(n;\mathbb{C})$. (After all, if G is a closed subgroup of $\mathsf{GL}(n;\mathbb{C})$ and H is a connected Lie subgroup of $\mathsf{GL}(n;\mathbb{C})$ whose Lie algebra $\mathfrak{h}$ is contained in H, then H is also a connected Lie subgroup of G.) As in the proof of Theorem 2.27, we think of $\mathsf{gl}(n;\mathbb{C})$ as $\mathbb{R}^{2n^2}$ and we decompose $\mathsf{gl}(n;\mathbb{C})$ as the direct sum of $\mathfrak{h}$ and D, where D is the orthogonal complement of $\mathfrak{h}$ with respect to the usual inner product on $\mathbb{R}^{2n^2}$. Then, as shown in the proof of Theorem 2.27, there exists neighborhoods U and V of the origin in $\mathfrak{h}$ and D and a neighborhood W of I in $\mathsf{GL}(n;\mathbb{C})$ such that each $A \in W$ can be written uniquely as

$$A = e^X e^Y \tag{3.23}$$

with $X \in U$, $Y \in V$, and such that X and Y depend continuously on A. Now, define

$$E = \left\{ Y \in V \,\middle|\, e^Y \in H \right\}.$$

Lemma 3.19. *The set E is at most countable.*

Let us assume this result for the moment and continue with the proof of the theorem. Define H to be the set of elements $A \in \mathsf{GL}(n;\mathbb{C})$ that can be expressed in the form $\exp X_1 \cdots \exp X_m$ for some finite collection $X_1, \ldots, X_m$ of elements of $\mathfrak{h}$. This set is, by definition, closed under multiplication. It is closed under inverses since the inverse of $\exp X$ is $\exp(-X)$. So, H is a subgroup of $\mathsf{GL}(n;\mathbb{C})$. Furthermore, H satisfies, by its definition, Condition 3 in the definition of connected Lie subgroups. Thus, it remains only to show that the Lie algebra of H is $\mathfrak{h}$.

Let $\mathfrak{h}'$ be the Lie algebra of H. Clearly, $\mathfrak{h}' \supset \mathfrak{h}$, so it remains to show that $\mathfrak{h}' \subset \mathfrak{h}$. Suppose Z is an element of $\mathfrak{h}'$. Then, as in (3.23), we may write, for all sufficiently small t,

$$e^{tZ} = e^{X(t)} e^{Y(t)},$$

where $X(t) \in U \subset \mathfrak{h}$ and $Y(t) \in V \subset D$ and where $X(t)$ and $Y(t)$ are continuous functions of t. Now, both $\exp(tZ)$ and $\exp X(t)$ belong to H, and since H is a subgroup, we conclude that $\exp Y(t)$ must also belong to H. This means that $Y(t)$ belongs to the set E for all sufficiently small t. If $Y(t)$ were not constant, then it would take on uncountably many values, which would mean that E is uncountable, violating Lemma 3.19. So, $Y(t)$ must be constant, and since $Y(0) = 0$, this means that $Y(t)$ is identically equal to zero. Thus, for small t, we have $\exp(tZ) = \exp X(t)$ and, therefore, $tZ = X(t) \in \mathfrak{h}$. This means $Z \in \mathfrak{h}$ and we conclude that $\mathfrak{h}' \subset \mathfrak{h}$.

So, it remains only to prove Lemma 3.19. (This proof is the only place we use that $\mathfrak{h}$ is a sub*algebra* of $\mathfrak{gl}(n;\mathbb{C})$ and not just a subspace.) Before doing this, we prove another lemma.

Lemma 3.20. *Pick a basis for $\mathfrak{h}$ and call an element R of $\mathfrak{h}$* ***rational*** *if its coefficients with respect to this basis are rational. Then, for every $\delta > 0$ and every $A \in H$, there exist rational elements $R_1, \ldots, R_m$ of $\mathfrak{h}$ such that*

$$A = e^{R_1} e^{R_2} \cdots e^{R_m} e^{X},$$

where X is in $\mathfrak{h}$ and $\|X\| < \delta$.

Proof. Choose $\varepsilon > 0$ small enough that the Baker–Campbell–Hausdorff formula applies for all X and Y with $\|X\| < \varepsilon$ and $\|Y\| < \varepsilon$. Let $C(X,Y)$ denote the quantity on right-hand side of the Baker–Campbell-Hausdorff formula, so that $C(X,Y)$ satisfies

$$e^X e^Y = e^{C(X,Y)}$$

whenever $\|X\|, \|Y\| < \varepsilon$. It is not hard to see that the function $C(X,Y)$ is continuous.

Now, choose $\varepsilon' > 0$ small enough that $\|C(X,Y)\| < \varepsilon$ for all X and Y with $\|X\| < \varepsilon'$ and $\|Y\| < \varepsilon'$. Since $\exp X = (\exp(X/n))^n$, every element A of H can be written in the form

$$A = e^{X_1} \cdots e^{X_m} \tag{3.24}$$

for some sequence $X_1, \ldots, X_m$ in $\mathfrak{h}$ with $\|X_k\| < \varepsilon'$, $k = 1, \ldots, m$.

Now, *because* $\mathfrak{h}$ *is a subalgebra of* $\mathfrak{gl}(n;\mathbb{C})$, $C(X_1, X_2)$ will be, again, an element of $\mathfrak{h}$, since the operators ad_{X_1} and ad_{X_2} on the right-hand side of the Baker–Campbell–Hausdorff formula preserve $\mathfrak{h}$. Choose a rational element R_1 of $\mathfrak{h}$ that is very close to $C(X_1, X_2)$ and that satisfies $\|R_1\| < \varepsilon$. (This is possible because X_1 and X_2 have norm less than ε' and, thus, $C(X_1, X_2)$ has norm less than ε.) Then we have

$$\begin{aligned} e^{X_1} e^{X_2} &= e^{C(X_1, X_2)} \\ &= e^{R_1} e^{-R_1} e^{C(X_1, X_2)} \\ &= e^{R_1} e^{\tilde{X}_2}, \end{aligned}$$

where $\tilde{X}_2 = C(-R_1, C(X_1, X_2))$. Now, $C(\cdot, \cdot)$ is continuous and

$$C(-C(X_1, X_2), C(X_1, X_2)) = -C(X_1, X_2) + C(X_1, X_2) = 0,$$

since $C(X_1, X_2)$ commutes with itself. Thus, if we choose R_1 sufficiently close to $C(X_1, X_2)$, we will have $\|\tilde{X}_2\| < \varepsilon'$.

We see, then, that (3.24) may be rewritten as

$$A = e^{R_1} e^{\tilde{X}_2} e^{X_3} \cdots e^{X_m}.$$

where R_1 is rational and $\tilde{X}_2$ (like X_2) has norm less than ε'. Applying the same argument to $\tilde{X}_2$ and X_3 we obtain

$$A = e^{R_1} e^{R_2} e^{\tilde{X}_3} e^{X_4} \cdots e^{X_m}.$$

Continuing on in the same way we eventually obtain

$$A = e^{R_1} e^{R_2} \cdots e^{R_{m-1}} e^{\tilde{X}_m}$$

with $R_1, \ldots, R_{m-1}$ rational. If, at the very last stage, we choose R_{m-1} so that $\|\tilde{X}_{m-1}\| < \delta$, we have expressed A in the desired form. □

We now supply the proof of Lemma 3.19.

Proof. Fix δ so that for all X and Y with $\|X\|, \|Y\| < \delta$ the quantity $C(X, Y)$ (the right-hand side of the Baker–Campbell–Hausdorff formula) is well defined and contained in U. Then, I claim that for each sequence $R_1, \ldots, R_m$ of rational elements in $\mathfrak{h}$, there is at most one $X \in \mathfrak{h}$ with $\|X\| < \delta$ such that the element

$$e^{R_1} e^{R_2} \cdots e^{R_m} e^{X} \tag{3.25}$$

belongs to $\exp V$. After all, if we have

$$e^{R_1} e^{R_2} \cdots e^{R_m} e^{X_1} = e^{Y_1}, \tag{3.26}$$

$$e^{R_1} e^{R_2} \cdots e^{R_m} e^{X_2} = e^{Y_2} \tag{3.27}$$

with $Y_1, Y_2 \in V$, then

$$e^{Y_2} = e^{Y_1} e^{-X_1} e^{X_2} = e^{Y_1} e^{C(-X_1, X_2)}$$

with $C(-X_1, X_2) \in U$. However, each element of $\exp V \exp U$ has a *unique* representation as $e^Y e^X$ with $X \in U$ and $Y \in V$, so we must have $Y_2 = Y_1$ and and so (by (3.26) and (3.27)) $e^{X_1} = e^{X_2}$, which implies that $X_1 = X_2$, since exp is injective on U.

By Lemma 3.20, every element of H can be expressed in the form (3.25) with $\|X\| < \delta$. Now, there are only countably many rational elements in $\mathfrak{h}$ and thus only countably many expressions of the form $e^{R_1} \cdots e^{R_m}$, each of which produces at most one element of the form (3.25) that belongs to $\exp V$. Thus, the set E in Lemma 3.19 is at most countable. □

This completes the proof of Theorem 3.15. □

3.9 Exercises

1. The **center** of a Lie algebra $\mathfrak{g}$ is defined to be the set of all $X \in \mathfrak{g}$ such that $[X, Y] = 0$ for all $Y \in \mathfrak{g}$. Now, consider the Heisenberg group

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

with Lie algebra

$$\mathfrak{h} = \left\{ \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix} \middle| \, \alpha, \beta, \gamma \in \mathbb{R} \right\}.$$

Determine the center $Z(\mathfrak{h})$ of $\mathfrak{h}$. For any $X, Y \in \mathfrak{h}$, show that $[X, Y] \in Z(\mathfrak{h})$. Note that this implies, in particular, that both X and Y commute with their commutator.
Show by direct computation that for any $X, Y \in \mathfrak{h}$,

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

2. Let X be a linear transformation on a finite-dimensional real or complex vector space. Show that

$$\frac{I - e^{-X}}{X}$$

is invertible if and only none of the eigenvalues of X (over $\mathbb{C}$) is of the form $2\pi i n$, with n a nonzero integer.

Remark: This exercise, combined with the formula in Theorem 3.5, gives the following result (in the language of differentiable manifolds): The exponential mapping $\exp : \mathfrak{g} \to G$ is a local diffeomorphism near $X \in \mathfrak{g}$ if and only $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ has no eigenvalue of the form $2\pi i n$, with n a nonzero integer.

3. Show that for any $n \times n$ matrices X and Y,

$$\left\| \frac{d}{dt} (X + tY)^m \Big|_{t=0} \right\| \leq m \|X\|^{m-1} \|Y\| .$$

Using this, show that the map $\exp : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is continuously differentiable.
Hint: Since we know that the series for the exponential mapping converges uniformly on sets of the form $\{X \mid \|X\| < R\}$, it suffices to show that the series of term-by-term directional derivatives also converges uniformly on such sets. (Compare Theorem 7.17 in Rudin (1976).)

4. Show that for any X and Y in $M_n(\mathbb{C})$, *even if X and Y do not commute*,

$$\frac{d}{dt} \mathrm{trace} \left(e^{X+tY} \right) \Big|_{t=0} = \mathrm{trace} \left(e^X Y \right) .$$

5. Verify that the right-hand side of the Baker–Campbell–Hausdorff formula (3.6) reduces to $X + Y$ in the case that X and Y commute.
6. Compute $\log \left(e^X e^Y \right)$ through third order in X and Y by using the power series for the exponential and the logarithm. Show this gives the same answer as the Baker–Campbell–Hausdorff formula.
7. Using the techniques in Section 3.5, compute the series form of the Baker–Campbell–Hausdorff formula up through fourth-order brackets. (We have already computed up through third-order brackets.)
8. Suppose that X and Y are upper triangular matrices with zeros on the diagonal. Show that the power series for $\log(\exp X \exp Y)$ is convergent. What happens to the series form of the Baker–Campbell–Hausdorff formula in this case?
9. Give an example of matrices X and Y in $\mathfrak{sl}(2; \mathbb{R})$ such that there does not exist any Z in $\mathfrak{sl}(2; \mathbb{R})$ with $\exp X \exp Y = \exp Z$. Use Exercise 30 of Chapter 2. What does this say about the result of applying the Baker–Campbell–Hausdorff formula to X and Y?
10. Complete Step 5 in the proof of Theorem 3.7 by showing that Φ as defined in Steps 1 through 4 is a homomorphism. Given $A, B \in G$, choose a path $A(t)$ connecting I to A and a path $B(t)$ connecting I to B. Then, define a path C by setting $C(t) = A(2t)$ for $0 \leq t \leq 1/2$ and setting $C(t) = A \cdot B(2t - 1)$ for $1/2 \leq t \leq 1$. (Thus, C connects I to AB.) If $t_0, \ldots, t_m$ is a valid partition for $A(t)$ and $s_0, \ldots, s_M$ is a valid partition for $B(t)$, show that

$$\frac{t_0}{2}, \cdots, \frac{t_m}{2}, \frac{1}{2} + \frac{s_0}{2}, \cdots, \frac{1}{2} + \frac{s_M}{2}$$

is a valid partition for $C(t)$. Now, compute $\Phi(A)$, $\Phi(B)$, and $\Phi(AB)$ using these paths and partitions and show that $\Phi(AB) = \Phi(A)\Phi(B)$.

11. If $\tilde{G}$ is a universal cover of a connected group G with projection map Φ, show that Φ maps $\tilde{G}$ onto G.
12. Suppose that G is a connected matrix Lie group and that $\tilde{G}$ is the universal cover of G. Show that G is isomorphic to $\tilde{G}/N$, where N is a discrete subgroup of the center of $\tilde{G}$. Use Exercise 11 from Chapter 1.
13. Suppose that G is a matrix Lie group with Lie algebra $\mathfrak{g}$. Suppose that $\phi : \mathsf{sl}(n;\mathbb{R}) \to \mathfrak{g}$ is a Lie algebra homomorphism. Show that there exists a Lie group homomorphism $\Phi : \mathsf{SL}(n;\mathbb{R}) \to G$ such that $\Phi(\exp X) = \exp \phi(X)$ for all $X \in \mathsf{sl}(n;\mathbb{R})$. This is true even though $\mathsf{SL}(n;\mathbb{R})$ is *not* simply connected.
    Hint: Use the fact that $\mathsf{SL}(n;\mathbb{C})$ is simply connected.
    Note: The result of this problem is false if G is assumed merely to be a Lie group and not a matrix Lie group.
14. Prove the uniqueness portion of Theorem 3.10. Use the fact that Theorem 3.7 (and basic results from Chapter 2) continue to hold for all (not necessarily matrix) Lie groups.
15. Show that the Lie algebra homomorphism associated to the group homomorphism Φ in (3.20) is a Lie algebra isomorphism. (Here the Lie algebra of $\mathbb{R}$ is identified simply with $\mathbb{R}$.)
16. Let $\mathfrak{a}$ be a subalgebra of the Lie algebra of the Heisenberg group. Show that $\exp(\mathfrak{a})$ is a connected Lie subgroup of the Heisenberg group.
17. Show that every connected Lie subgroup of $\mathsf{SU}(2)$ is closed. Show that this is not the case for $\mathsf{SU}(3)$.
18. Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$, let $\mathfrak{h}$ be a subalgebra of $\mathfrak{g}$, and let H be the unique connected Lie subgroup of G with Lie algebra $\mathfrak{h}$. Suppose that there exists a compact simply-connected matrix Lie group K such that the Lie algebra of K is isomorphic to $\mathfrak{h}$. Show that H is closed. Is H necessarily isomorphic to K?

4

Basic Representation Theory

4.1 Representations

Definition 4.1. *Let G be a matrix Lie group. Then, a **finite-dimensional complex representation** of G is a Lie group homomorphism*

$$\Pi : G \to \mathsf{GL}(n;\mathbb{C})$$

($n \geq 1$) or, more generally, a Lie group homomorphism

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

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

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

*If Π or π is a one-to-one homomorphism, then the representation is called **faithful**.*

One should think of a representation as a linear **action** of a group or Lie algebra on a vector space (since, say, to every $g \in G$, there is associated an operator $\Pi(g)$, which acts on the vector space V). In fact, we will use terminology such as "Let Π be a representation of G acting on the space V." Even if $\mathfrak{g}$ is a real Lie algebra, we will consider mainly complex representations of $\mathfrak{g}$. After making a few more definitions, we will discuss the question of why one should be interested in studying representations.

Definition 4.2. *Let Π be a finite-dimensional real or complex representation of a matrix Lie group G, acting on a space V. A subspace W of V is called*

invariant *if* $\Pi(A)w \in W$ *for all* $w \in W$ *and all* $A \in G$. *An invariant subspace* W *is called* ***nontrivial*** *if* $W \neq \{0\}$ *and* $W \neq V$. *A representation with no nontrivial invariant subspaces is called* ***irreducible***.

The terms ***invariant***, ***nontrivial***, *and* ***irreducible*** *are defined analogously for representations of Lie algebras.*

Definition 4.3. *Let* G *be a matrix Lie group, let* Π *be a representation of* G *acting on the space* V, *and let* Σ *be a representation of* G *acting on the space* W. *A linear map* $\phi : V \to W$ *is called an* ***intertwining map*** *of representations if*

$$\phi(\Pi(A)v) = \Sigma(A)\phi(v)$$

for all $A \in G$ *and all* $v \in V$. *The analogous property defines intertwining maps of representations of a Lie algebra.*

If ϕ *is an intertwining map of representations and, in addition,* ϕ *is invertible, then* ϕ *is said to be an* ***equivalence*** *of representations. If there exists an isomorphism between* V *and* W, *then the representations are said to be* ***equivalent***.

Two equivalent representations should be regarded as being "the same" representation. A typical problem in representation theory is to determine, up to equivalence, all of the irreducible representations of a particular group or Lie algebra. In Section 5.4, we will determine all the finite-dimensional complex irreducible representations of the Lie algebra $\mathsf{su}(2)$.

Proposition 4.4. *Let* G *be a matrix Lie group with Lie algebra* $\mathfrak{g}$ *and let* Π *be a (finite-dimensional real or complex) representation of* G, *acting on the space* V. *Then, there is a unique representation* π *of* $\mathfrak{g}$ *acting on the same space such that*

$$\Pi(e^X) = e^{\pi(X)}$$

for all $X \in \mathfrak{g}$. *The representation* π *can be computed as*

$$\pi(X) = \frac{d}{dt}\Pi\left(e^{tX}\right)\Big|_{t=0}$$

and satisfies

$$\pi\left(AXA^{-1}\right) = \Pi(A)\pi(X)\Pi(A)^{-1}$$

for all $X \in \mathfrak{g}$ *and all* $A \in G$.

Proof. Theorem 2.21 states that for each Lie group homomorphism $\Phi : G \to H$, there is an associated Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$. Take $H = \mathsf{GL}(V)$ and $\Phi = \Pi$. Since the Lie algebra of $\mathsf{GL}(V)$ is $\mathfrak{gl}(V)$ (since the exponential of any operator is invertible), the associated Lie algebra homomorphism $\phi = \pi$ maps from $\mathfrak{g}$ to $\mathfrak{gl}(V)$ and, so, constitutes a representation of $\mathfrak{g}$.

The properties of π follow from the properties of ϕ given in Theorem 6 □

the motivation for the term "representation." (Of course, we still call Π a representation even if it is not faithful.)

Despite the origin of the term, the point of representation theory is *not* (at least in this book) to represent a group as a group of matrices. After all, all of our groups are already matrix groups! Although it might seem redundant to study representations of a group which is already represented as a group of matrices, this is precisely what we are going to do.

The reason for this is that a representation can be thought of (as we have already noted) as an action of our group on some vector space. Such actions (representations) arise naturally in many branches of both mathematics and physics, and it is important to understand them.

A typical example would be a differential equation in three-dimensional space which has rotational symmetry. If the equation has rotational symmetry, then the space of solutions will be invariant under rotations. Thus, the space of solutions will constitute a representation of the rotation group $\mathrm{SO}(3)$. If one knows what all of the representations of $\mathrm{SO}(3)$ are, this can help immensely in narrowing down what the space of solutions can be. (As we will see, $\mathrm{SO}(3)$ has many other representations besides the obvious one in which $\mathrm{SO}(3)$ acts on $\mathbb{R}^3$.)

In fact, one of the chief applications of representation theory is to exploit symmetry. If a system has symmetry, then the set of symmetries will form a group, and understanding the representations of the symmetry group allows one to use that symmetry to simplify the problem.

In addition, studying the representations of a group G (or of a Lie algebra $\mathfrak{g}$) can give information about the group (or Lie algebra) itself. For example, if G is a *finite* group, then associated to G is something called the **group algebra**. The structure of this group algebra can be described very nicely in terms of the irreducible representations of G.

In this book, we will be interested primarily in computing the finite-dimensional irreducible complex representations of matrix Lie groups. As we shall see, this problem can be reduced almost completely to the problem of computing the finite-dimensional irreducible complex representations of the associated Lie algebra. In this chapter, we will discuss the theory at an elementary level and will consider in detail the examples of $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$. In Chapter 5, we will study the representations of $\mathrm{SU}(3)$, which is similar to but more involved than that of $\mathrm{SU}(2)$. In Chapter 7, we will look at the general theory of representations of semisimple groups.

4.3 Examples of Representations

4.3.1 The standard representation

A matrix Lie group G is, by definition, a subset of some $\mathrm{GL}(n;\mathbb{C})$. The inclusion map of G into $\mathrm{GL}(n;\mathbb{C})$ (i.e., $\Pi(A) = A$) is a representation of G,

Proposition 4.5.

1. *Let G be a connected matrix Lie group with Lie algebra $\mathfrak{g}$. Let Π be a representation of G and π the associated representation of $\mathfrak{g}$. Then, Π is irreducible if and only if π is irreducible.*
2. *Let G be a connected matrix Lie group, let Π_1 and Π_2 be representations of G, and let π_1 and π_2 be the associated Lie algebra representations. Then, π_1 and π_2 are equivalent if and only if Π_1 and Π_2 are equivalent.*

Proof. For Point 1, suppose first that Π is irreducible. We then want to show that π is irreducible. So, let W be a subspace of V that is invariant under $\pi(X)$ for all $X \in \mathfrak{g}$. We want to show that W is either $\{0\}$ or V. Now, suppose A is an element of G. Since G is assumed connected, Corollary 2.31 tells us that A can be written as $A = e^{X_1} \cdots e^{X_m}$ for some $X_1, \ldots, X_m$ in $\mathfrak{g}$. Since W is invariant under $\pi(X_i)$ it will also be invariant under $\exp(\pi(X_i)) = I + \pi(X_i) + \pi(X_i)^2/2 + \cdots$ and, hence, under

$$\begin{aligned}\Pi(A) &= \Pi\left(e^{X_1} \cdots e^{X_m}\right) = \Pi\left(e^{X_1}\right) \cdots \Pi\left(e^{X_m}\right) \\ &= e^{\pi(X_1)} \cdots e^{\pi(X_m)}.\end{aligned}$$

Since Π is irreducible and W is invariant under each $\Pi(A)$, W must be either $\{0\}$ or V. This shows that π is irreducible.

In the other direction, assume that π is irreducible and that W is an invariant subspace for Π. Then, W is invariant under $\Pi(\exp tX)$ for all $X \in \mathfrak{g}$ and, hence, under

$$\pi(X) = \frac{d}{dt}\Pi\left(e^{tX}\right)\Big|_{t=0}$$

Thus, since π is irreducible, W is $\{0\}$ or V, and we conclude that Π is irreducible. This establishes Point 1 of the proposition.

Point 2 of the proposition is similar and is left as an exercise to the reader (Exercise 1). □

Proposition 4.6. *Let $\mathfrak{g}$ be a real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ its complexification. Then, every finite-dimensional complex representation π of $\mathfrak{g}$ has a unique extension to a complex-linear representation of $\mathfrak{g}_{\mathbb{C}}$, also denoted π and given by*

$$\pi(X + iY) = \pi(X) + i\pi(Y)$$

for all $X, Y \in \mathfrak{g}$. Furthermore, π is irreducible as a representation of $\mathfrak{g}_{\mathbb{C}}$ if and only if it is irreducible as a representation of $\mathfrak{g}$.

Proof. The existence and uniqueness of the extension are trivial and follow from Exercise 23 of Chapter 2.

Concerning irreducibility, let us make sure that we are clear about what the statement means. Suppose that π is a complex representation of the *real* Lie algebra $\mathfrak{g}$, acting on the *complex* vector space V. Then, saying that π

is irreducible means that there is no nontrivial invariant *complex* subspace $W \subset V$. That is, even though $\mathfrak{g}$ is a real Lie algebra, when considering complex representations of $\mathfrak{g}$, we are interested only in complex invariant subspaces.

Now, suppose that π is irreducible as a representation of $\mathfrak{g}$. If W is a complex subspace of V which is invariant under $\mathfrak{g}_{\mathbb{C}}$, then, certainly, W is invariant under $\mathfrak{g} \subset \mathfrak{g}_{\mathbb{C}}$. Therefore, $W = \{0\}$ or $W = V$. Thus, π is irreducible as a representation of $\mathfrak{g}_{\mathbb{C}}$.

On the other hand, suppose that π is irreducible as a representation of $\mathfrak{g}_{\mathbb{C}}$ and suppose that W is a complex subspace of V which is invariant under $\mathfrak{g}$. Then, W will also be invariant under $\pi(X + iY) = \pi(X) + i\pi(Y)$, for all $X, Y \in \mathfrak{g}$. Since every element of $\mathfrak{g}_{\mathbb{C}}$ can be written as $X + iY$, we conclude that, in fact, W is invariant under $\mathfrak{g}_{\mathbb{C}}$. Thus, $W = \{0\}$ or $W = V$ and π is irreducible as a representation of $\mathfrak{g}$. □

Definition 4.7. *Let G be a matrix Lie group, let $\mathcal{H}$ be a Hilbert space, and let $U(\mathcal{H})$ denote the group of unitary operators on $\mathcal{H}$. Then, a homomorphism $\Pi : G \to U(\mathcal{H})$ is called a* ***unitary representation*** *of G if Π satisfies the following continuity condition: If $A_n, A \in G$ and $A_n \to A$, then*

$$\Pi(A_n)v \to \Pi(A)v$$

for all $v \in \mathcal{H}$. A unitary representation with no nontrivial closed invariant subspaces is called ***irreducible.***

This continuity condition is called **strong continuity**. One could require the even stronger condition that $\|\Pi(A_n) - \Pi(A)\| \to 0$, but this turns out to be too stringent a requirement. (That is, most of the interesting unitary representations of G will not have this stronger continuity condition.) In practice, any homomorphism of G into $U(\mathcal{H})$ that one can write down explicitly will be strongly continuous.

Note here that $\mathcal{H}$ is not assumed to be finite dimensional. Although we will deal in this book almost exclusively with finite-dimensional representations, it is good to be aware of the concept of infinite-dimensional unitary representations. If $\mathcal{H}$ is infinite dimensional, there are many technical issues that we will not be able to delve into in this book. For example, the correct notion of a Lie algebra representation associated to an infinite-dimensional unitary representation is quite subtle and we will not address this issue at all. Nevertheless, see Exercise 8 for a calculation of such a Lie algebra representation (in which all technical difficulties are swept under the carpet).

4.2 Why Study Representations?

If a representation Π is a faithful representation of a matrix Lie group G, then $\{\Pi(A) \,|\, A \in G\}$ is a group of matrices that is isomorphic to the original group G. Thus, Π allows us to *represent* G as a group of matrices. This is

called the **standard representation** of G. If G happens to be contained in $\mathsf{GL}(n;\mathbb{R}) \subset \mathsf{GL}(n;\mathbb{C})$, then we can think of the standard representation as a real representation if we prefer. Thus, for example, the standard representation of $\mathsf{SO}(3)$ is the one in which $\mathsf{SO}(3)$ acts in the usual way on $\mathbb{R}^3$ and the standard representation of $\mathsf{SU}(2)$ is the one in which $\mathsf{SU}(2)$ acts on $\mathbb{C}^2$ in the usual way. If G is a subgroup of $\mathsf{GL}(n;\mathbb{R})$ or $\mathsf{GL}(n;\mathbb{C})$, then its Lie algebra $\mathfrak{g}$ will be a subalgebra of $\mathfrak{gl}(n;\mathbb{R})$ or $\mathfrak{gl}(n;\mathbb{C})$. The inclusion of $\mathfrak{g}$ into $\mathfrak{gl}(n;\mathbb{R})$ or $\mathfrak{gl}(n;\mathbb{C})$ is a representation of $\mathfrak{g}$, called the **standard representation**.

4.3.2 The trivial representation

Consider the one-dimensional complex vector space $\mathbb{C}$. Given any matrix Lie group G, we can define the **trivial representation** of G, $\Pi : G \to \mathsf{GL}(1;\mathbb{C})$, by the formula

$$\Pi(A) = I$$

for all $A \in G$. Of course, this is an irreducible representation, since $\mathbb{C}$ has no nontrivial subspaces, let alone nontrivial invariant subspaces. If $\mathfrak{g}$ is a Lie algebra, we can also define the **trivial representation** of $\mathfrak{g}$, $\pi : \mathfrak{g} \to \mathfrak{gl}(1;\mathbb{C})$, by

$$\pi(X) = 0$$

for all $X \in \mathfrak{g}$. This is an irreducible representation.

4.3.3 The adjoint representation

Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. We have already defined the adjoint mapping

$$\mathrm{Ad} : G \to \mathsf{GL}(\mathfrak{g})$$

by the formula

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

Recall that "Ad" is a Lie group homomorphism. Since Ad is a Lie group homomorphism into a group of invertible operators, we see that, in fact, Ad is a representation of G, acting on the space $\mathfrak{g}$. Thus, we can now give Ad its proper name, the **adjoint representation** of G. The adjoint representation is a real representation of G. (If $\mathfrak{g}$ happens to be a complex subspace of $M_n(\mathbb{C})$, then we can think of the adjoint representation as a complex representation.)

Similarly, if $\mathfrak{g}$ is a Lie algebra, we have

$$\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}),$$

defined by the formula

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

We know that "ad" is a Lie algebra homomorphism and is, therefore, a representation of $\mathfrak{g}$, called the **adjoint representation**. In the case that $\mathfrak{g}$ is

the Lie algebra of some matrix Lie group G, we have already established (Chapter 2, Proposition 2.24 and Exercise 19) that Ad and ad are related by $\exp(\mathrm{ad}_X) = \mathrm{Ad}_{e^X}$.

Note that in the case of $\mathsf{SO}(3)$, the standard representation and the adjoint representation are both three dimensional real representations. In fact, these two representations are equivalent (Exercise 3).

4.3.4 Some representations of SU(2)

Consider the space V_m of homogeneous polynomials in two complex variables with total degree m ($m \geq 0$); that is, V_m is the space of functions of the form

$$f(z_1, z_2) = a_0 z_1^m + a_1 z_1^{m-1} z_2 + a_2 z_1^{m-2} z_2^2 + \cdots + a_m z_2^m \tag{4.1}$$

with $z_1, z_2 \in \mathbb{C}$ and the a_i's arbitrary complex constants. The space V_m is an $(m+1)$-dimensional complex vector space.

Now, by definition, an element U of $\mathsf{SU}(2)$ is a linear transformation of $\mathbb{C}^2$. Let z denote the pair $z = (z_1, z_2)$ in $\mathbb{C}^2$. Then, we may define a linear transformation $\Pi_m(U)$ on the space V_m by the formula

$$[\Pi_m(U)f](z) = f(U^{-1}z). \tag{4.2}$$

Explicitly, if f is as in (4.1), then

$$[\Pi_m(U)f](z_1, z_2) = \sum_{k=0}^{m} a_k \left(U_{11}^{-1} z_1 + U_{12}^{-1} z_2\right)^{m-k} \left(U_{21}^{-1} z_1 + U_{22}^{-1} z_2\right)^k.$$

By expanding out the right-hand side of this formula, we see that $\Pi_m(U)f$ is again a homogeneous polynomial of degree m. Thus, $\Pi_m(U)$ actually maps V_m into V_m.

Now, compute

$$\begin{aligned}\Pi_m(U_1)\left[\Pi_m(U_2)f\right](z) &= \left[\Pi_m(U_2)f\right](U_1^{-1}z) = f\left(U_2^{-1}U_1^{-1}z\right) \\ &= \Pi_m(U_1U_2)f(z).\end{aligned}$$

Thus, Π_m is a (finite-dimensional complex) representation of $\mathsf{SU}(2)$. The inverse in (4.2) is necessary in order to make Π_m a representation. We will see eventually that each of the representations Π_m of $\mathsf{SU}(2)$ is irreducible and that every finite-dimensional irreducible representation of $\mathsf{SU}(2)$ is equivalent to one (and only one) of the Π_m's. (Of course, no two of the Π_m's are equivalent, since they do not even have the same dimension.)

Let us now compute the corresponding Lie algebra representation π_m. According to Proposition 4.4, π_m can be computed as

$$\pi_m(X) = \frac{d}{dt}\Pi_m\left(e^{tX}\right)\Big|_{t=0}.$$

So,

$$(\pi_m(X)f)(z) = \frac{d}{dt} f\left(e^{-tX}z\right)\Big|_{t=0}.$$

Now, let $z(t)$ be the curve in $\mathbb{C}^2$ defined as $z(t) = e^{-tX}z$, so that $z(0) = z$. Of course, $z(t)$ can be written as $z(t) = (z_1(t), z_2(t))$, with $z_i(t) \in \mathbb{C}$. By the chain rule,

$$\pi_m(X)f = \frac{\partial f}{\partial z_1}\frac{dz_1}{dt}\Big|_{t=0} + \frac{\partial f}{\partial z_2}\frac{dz_2}{dt}\Big|_{t=0}.$$

However, $dz/dt|_{t=0} = -Xz$, so we obtain the following formula for $\pi_m(X)$:

$$\pi_m(X)f = -\frac{\partial f}{\partial z_1}(X_{11}z_1 + X_{12}z_2) - \frac{\partial f}{\partial z_2}(X_{21}z_1 + X_{22}z_2). \tag{4.3}$$

Now, according to Proposition 4.6, every finite-dimensional complex representation of the Lie algebra $\mathsf{su}(2)$ extends uniquely to a complex-linear representation of the complexification of $\mathsf{su}(2)$. However, the complexification of $\mathsf{su}(2)$ is (isomorphic to) $\mathsf{sl}(2;\mathbb{C})$ (Proposition 2.45). The representation π_m of $\mathsf{su}(2)$ given by (4.3) thus extends to a representation of $\mathsf{sl}(2;\mathbb{C})$, which we will also call π_m and which (as is easily verified) is also given by (4.3).

So, for example, consider the element

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

in the Lie algebra $\mathsf{sl}(2;\mathbb{C})$. Applying formula (4.3) gives

$$(\pi_m(H)f)(z) = -\frac{\partial f}{\partial z_1}z_1 + \frac{\partial f}{\partial z_2}z_2.$$

Thus, we see that

$$\pi_m(H) = -z_1\frac{\partial}{\partial z_1} + z_2\frac{\partial}{\partial z_2}. \tag{4.4}$$

Applying $\pi_m(H)$ to a basis element $z_1^k z_2^{m-k}$, we get

$$\pi_m(H)z_1^k z_2^{m-k} = -kz_1^k z_2^{m-k} + (m-k)z_1^k z_2^{m-k} = (m-2k)z_1^k z_2^{m-k}.$$

Thus, $z_1^k z_2^{m-k}$ is an eigenvector for $\pi_m(H)$ with eigenvalue $(m-2k)$. In particular, $\pi_m(H)$ is diagonalizable.

Let X and Y be the elements

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

in $\mathsf{sl}(2;\mathbb{C})$. Then, (4.3) tells us that

$$\pi_m(X) = -z_2\frac{\partial}{\partial z_1}, \quad \pi_m(Y) = -z_1\frac{\partial}{\partial z_2}$$

so that

$$\pi_m(X)z_1^k z_2^{m-k} = -kz_1^{k-1}z_2^{m-k+1},$$
$$\pi_m(Y)z_1^k z_2^{m-k} = (k-m)z_1^{k+1}z_2^{m-k-1}. \tag{4.5}$$

Proposition 4.8. *The representation π_m is an irreducible representation of* $\mathsf{sl}(2;\mathbb{C})$.

Proof. It suffices to show that every nonzero invariant subspace of V_m is, in fact, equal to V_m. So, let W be such a space. Since W is assumed nonzero, there is at least one nonzero element w in W. Then, w can be written uniquely in the form

$$w = a_0 z_1^m + a_1 z_1^{m-1} z_2 + a_2 z_1^{m-2} z_2^2 + \cdots + a_m z_2^m$$

with at least one of the a_k's nonzero. Let k_0 be the smallest value of k for which $a_k \neq 0$ and consider

$$\pi_m(X)^{m-k_0} w.$$

Since (by (4.5)) each application of $\pi_m(X)$ lowers the power of z_1 by 1, $\pi_m(X)^{m-k_0}$ will kill all the terms in w except $a_{k_0} z_1^{m-k_0} z^{k_0}$. On the other hand, we compute easily that

$$\pi_m(X)^{m-k_0}(z_1^{m-k_0} z_2^{k_0}) = (-1)^{m-k_0}(m-k_0)! z_2^m.$$

We see, then, that $\pi_m(X)^{m-k_0} w$ is a *nonzero* multiple of z_2^m. Since W is assumed invariant, W must contain z_2^m. Furthermore, it follows from (4.5) that $\pi_m(Y)^k z_2^m$ is a *nonzero* multiple of $z_1^k z_2^{m-k}$. Therefore, W must also contain $z_1^k z_2^{m-k}$ for all $0 \leq k \leq m$. Since these elements form a basis for V_m, we see that, in fact, $W = V_m$, as desired. □

4.3.5 Two unitary representations of SO(3)

Let $\mathcal{H} = L^2(\mathbb{R}^3, dx)$, the space of square-integrable functions on $\mathbb{R}^3$. For each $R \in \mathsf{SO}(3)$, define an operator $\Pi_1(R)$ on $\mathcal{H}$ by the formula

$$[\Pi_1(R)f](x) = f\left(R^{-1}x\right).$$

Since Lebesgue measure dx is rotationally invariant, $\Pi_1(R)$ is a unitary operator for each $R \in \mathsf{SO}(3)$. The calculation of the previous subsection shows that the map $R \to \Pi_1(R)$ is a homomorphism of $\mathsf{SO}(3)$ into $U(\mathcal{H})$. This map is strongly continuous and hence constitutes a unitary representation of $\mathsf{SO}(3)$.

Similarly, we may consider the unit sphere $S^2 \subset \mathbb{R}^3$, with the usual surface measure Ω. Of course, any $R \in \mathsf{SO}(3)$ maps S^2 into S^2. For each R, we can define $\Pi_2(R)$ acting on $L^2(S^2, d\Omega)$ by

$$[\Pi_2(R)f](x) = f\left(R^{-1}x\right).$$

Then, Π_2 is a unitary representation of $\mathsf{SO}(3)$.

Neither of the unitary representations Π_1 and Π_2 is irreducible. In the case of Π_2, $L^2(S^2, d\Omega)$ has a very nice decomposition as the orthogonal direct sum of finite-dimensional invariant subspaces. This decomposition is the theory of "spherical harmonics," which are well known in the physics (and mathematics) literature.

4.3.6 A unitary representation of the reals

Let $\mathcal{H} = L^2(\mathbb{R}, dx)$. For each $a \in \mathbb{R}$, define $T_a : \mathcal{H} \to \mathcal{H}$ by

$$(T_a f)(x) = f(x - a).$$

Clearly, T_a is a unitary operator for each $a \in \mathbb{R}$ and, clearly, $T_a T_b = T_{a+b}$. The map $a \to T_a$ is strongly continuous, so T is a unitary representation of $\mathbb{R}$. This representation is not irreducible. The theory of the Fourier transform allows one to determine all the closed, invariant subspaces of $\mathcal{H}$ (Theorem 9.17 of Rudin (1987)).

4.3.7 The unitary representations of the Heisenberg group

Consider the Heisenberg group

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

Now, consider a real, nonzero constant, which, for reasons of historical convention, we will call $\hbar$ ("h bar"). Now, for each $\hbar \in \mathbb{R}\backslash\{0\}$, define a unitary operator $\Pi_\hbar$ on $L^2(\mathbb{R}, dx)$ by

$$\Pi_\hbar \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} f = e^{-i\hbar b} e^{i\hbar c x} f(x - a). \tag{4.6}$$

It is clear that the right-hand side of (4.6) has the same norm as f, so $\Pi_\hbar$ is, indeed, unitary.

Now, compute

$$\begin{aligned} &\Pi_\hbar \begin{pmatrix} 1 & \widetilde{a} & \widetilde{b} \\ 0 & 1 & \widetilde{c} \\ 0 & 0 & 1 \end{pmatrix} \Pi_\hbar \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} f \\ &= e^{-i\hbar \widetilde{b}} e^{i\hbar \widetilde{c} x} e^{-i\hbar b} e^{i\hbar c(x - \widetilde{a})} f(x - \widetilde{a} - a) \\ &= e^{-i\hbar(\widetilde{b} + b + c\widetilde{a})} e^{i\hbar(\widetilde{c} + c)x} f\left(x - (\widetilde{a} + a)\right). \end{aligned}$$

This shows that the map $A \to \Pi_\hbar(A)$ is a homomorphism of the Heisenberg group into $U\left(L^2(\mathbb{R})\right)$. This map is strongly continuous and, so, $\Pi_\hbar$ is a unitary representation of H.

Note that a typical unitary operator $\Pi_\hbar(A)$ consists of first translating f, then multiplying f by the function $e^{i\hbar cx}$, and then multiplying f by the constant $e^{-i\hbar b}$. Multiplying f by the function $e^{i\hbar cx}$ has the effect of translating the Fourier transform of f, or, in physical language, "translating f in momentum space." Now, if U_1 is an ordinary translation and U_2 is a translation of the Fourier transform (i.e., $U_2 =$ multiplication by some $e^{i\hbar cx}$), then U_1 and U_2 will not commute, but $U_1 U_2 U_1^{-1} U_2^{-1}$ will be simply multiplication by a constant of absolute value one. Thus, $\{\Pi_\hbar(A) \,|\, A \in H\}$ is the group of operators on $L^2(\mathbb{R})$ generated by ordinary translations and translations in Fourier space. It is this representation of the Heisenberg group which motivates its name. (See also Exercise 8.)

It follows fairly easily from standard Fourier transform theory (e.g., Theorem 9.17 of Rudin (1987)), that for each $\hbar \in \mathbb{R}\backslash\{0\}$, the representation $\Pi_\hbar$ is irreducible. Furthermore, these are (up to equivalence) almost all of the irreducible unitary representations of H. The only remaining ones are the one-dimensional representations $\Pi_{\alpha,\beta}$ given by

$$\Pi_{\alpha,\beta}\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} = e^{i(\alpha a + \beta c)} I$$

with $\alpha, \beta \in \mathbb{R}$. (The $\Pi_{\alpha,\beta}$'s are the irreducible unitary representations in which the center of H acts trivially.) The fact that the $\Pi_\hbar$'s and the $\Pi_{\alpha,\beta}$'s are all of the (strongly continuous) irreducible unitary representations of H is closely related to the celebrated Stone–Von Neumann theorem in mathematical physics. See, for example, Reed and Simon (1979), Theorem XI.84. See also Exercise 9.

4.4 The Irreducible Representations of su(2)

In this section, we will compute (up to equivalence) all of the finite-dimensional irreducible complex representations of the Lie algebra $\mathsf{su}(2)$. This computation is important for several reasons. In the first place, $\mathsf{su}(2) \cong \mathsf{so}(3)$ and the representations of $\mathsf{so}(3)$ are of physical significance. (The computation we will do here is found in every standard textbook on quantum mechanics, under the heading "angular momentum.") In the second place, the representation theory of $\mathsf{su}(2)$ is an illuminating example of how one uses commutation relations to determine the representations of a Lie algebra. In the third place, in determining the representations of semisimple Lie algebras (Chapters 5 and 6), we will explicitly use the representation theory of $\mathsf{su}(2)$.

Now, every finite-dimensional complex representation π of $\mathsf{su}(2)$ extends by Proposition 4.6 to a complex-linear representation (also called π) of the

complexification of $\mathsf{su}(2)$, namely $\mathsf{sl}(2;\mathbb{C})$. (Recall Section 2.9.) The extension of π to $\mathsf{sl}(2;\mathbb{C})$ is irreducible if and only if the original representation is irreducible, again by Proposition 4.6.

We see, then, that studying the irreducible representations of $\mathsf{su}(2)$ is equivalent to studying the irreducible (complex-linear) representations of $\mathsf{sl}(2;\mathbb{C})$. Passing to the complexified Lie algebra makes our computations easier, in that we can find a nice basis for $\mathsf{sl}(2;\mathbb{C})$ that has no counterpart among the bases of $\mathsf{su}(2)$.

We will use the following basis for $\mathsf{sl}(2;\mathbb{C})$:

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

which have the commutation relations

$$\begin{aligned} [H,X] &= 2X, \\ [H,Y] &= -2Y, \\ [X,Y] &= H. \end{aligned}$$

If V is a (finite-dimensional complex) vector space and A, B, and C are operators on V satisfying

$$\begin{aligned} [A,B] &= 2B, \\ [A,C] &= -2C, \\ [B,C] &= A, \end{aligned}$$

then because of the skew symmetry and bilinearity of brackets, the linear map $\pi : \mathsf{sl}(2;\mathbb{C}) \to \mathsf{gl}(V)$ satisfying

$$\pi(H) = A,\ \pi(X) = B,\ \pi(Y) = C$$

will be a representation of $\mathsf{sl}(2;\mathbb{C})$.

Theorem 4.9. *For each integer $m \geq 0$, there is an irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ with dimension $m+1$. Any two irreducible representations of $\mathsf{sl}(2;\mathbb{C})$ with the same dimension are equivalent. If π is an irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ with dimension $m+1$, then π is equivalent to the representation π_m described in Section 4.3.*

Proof. Let π be an irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ acting on a (finite-dimensional complex) space V. Our strategy is to diagonalize the operator $\pi(H)$. Of course, *a priori*, we do not know that $\pi(H)$ is diagonalizable. However, because we are working over the (algebraically closed) field of complex numbers, $\pi(H)$ must have at least one eigenvector.

The following lemma is the key to the entire proof.

Lemma 4.10. *Let u be an eigenvector of $\pi(H)$ with eigenvalue $\alpha \in \mathbb{C}$. Then,*

$$\pi(H)\pi(X)u = (\alpha+2)\pi(X)u.$$

Thus, either $\pi(X)u = 0$ or $\pi(X)u$ is an eigenvector for $\pi(H)$ with eigenvalue $\alpha + 2$. Similarly,

$$\pi(H)\pi(Y)u = (\alpha - 2)\pi(Y)u$$

so that either $\pi(Y)u = 0$ or $\pi(Y)u$ is an eigenvector for $\pi(H)$ with eigenvalue $\alpha - 2$.

Proof. We call $\pi(X)$ the "raising operator," because it has the effect of raising the eigenvalue of $\pi(H)$ by 2, and we call $\pi(Y)$ the "lowering operator." We know that $[\pi(H), \pi(X)] = \pi([H, X]) = 2\pi(X)$. Thus,

$$\pi(H)\pi(X) - \pi(X)\pi(H) = 2\pi(X)$$

or

$$\pi(H)\pi(X) = \pi(X)\pi(H) + 2\pi(X).$$

Thus,

$$\begin{aligned}\pi(H)\pi(X)u &= \pi(X)\pi(H)u + 2\pi(X)u \\ &= \pi(X)(\alpha u) + 2\pi(X)u \\ &= (\alpha + 2)\pi(X)u.\end{aligned}$$

Similarly, $[\pi(H), \pi(Y)] = -2\pi(Y)$, and, so,

$$\pi(H)\pi(Y) = \pi(Y)\pi(H) - 2\pi(Y)$$

so that

$$\begin{aligned}\pi(H)\pi(Y)u &= \pi(Y)\pi(H)u - 2\pi(Y)u \\ &= \pi(Y)(\alpha u) - 2\pi(Y)u \\ &= (\alpha - 2)\pi(Y)u.\end{aligned}$$

This is what we wanted to show. □

As we have observed, $\pi(H)$ must have at least one eigenvector u ($u \neq 0$), with some eigenvalue $\alpha \in \mathbb{C}$. By the lemma,

$$\pi(H)\pi(X)u = (\alpha + 2)\pi(X)u$$

and, more generally,

$$\pi(H)\pi(X)^n u = (\alpha + 2n)\pi(X)^n u.$$

This means that either $\pi(X)^n u = 0$ or $\pi(X)^n u$ is an eigenvector for $\pi(H)$ with eigenvalue $\alpha + 2n$.

Now, an operator on a finite-dimensional space can have only finitely many distinct eigenvalues. Thus, the $\pi(X)^n u$'s cannot all be different from zero. Thus, there is some $N \geq 0$ such that

$$\pi(X)^N u \neq 0$$

but

$$\pi(X)^{N+1} u = 0.$$

Define $u_0 = \pi(X)^N u$ and $\lambda = \alpha + 2N$. Then,

$$\pi(H)u_0 = \lambda u_0, \tag{4.7}$$
$$\pi(X)u_0 = 0. \tag{4.8}$$

Then, define

$$u_k = \pi(Y)^k u_0$$

for $k \geq 0$. By the second part of the lemma, we have

$$\pi(H)u_k = (\lambda - 2k)\, u_k. \tag{4.9}$$

Since, again, $\pi(H)$ can have only finitely many eigenvalues, the u_k's cannot all be nonzero.

Lemma 4.11. *With the above notation,*

$$\pi(X)u_k = [k\lambda - k(k-1)]u_{k-1} \quad (k > 0),$$
$$\pi(X)\, u_0 = 0.$$

Proof. We proceed by induction on k. In the case $k = 1$, we note that $u_1 = \pi(Y)u_0$. Using the commutation relation $[\pi(X), \pi(Y)] = \pi(H)$, we have

$$\pi(X)u_1 = \pi(X)\pi(Y)u_0 = (\pi(Y)\pi(X) + \pi(H))\, u_0.$$

However, $\pi(X)u_0 = 0$, so we get

$$\pi(X)u_1 = \lambda u_0,$$

which is the lemma in the case $k = 1$.

Now, by definition, $u_{k+1} = \pi(Y)u_k$. Using (4.9) and induction, we have

$$\begin{aligned}
\pi(X)u_{k+1} &= \pi(X)\pi(Y)u_k \\
&= (\pi(Y)\pi(X) + \pi(H))u_k \\
&= \pi(Y)[k\lambda - k(k-1)]u_{k-1} + (\lambda - 2k)u_k \\
&= [k\lambda - k(k-1) + (\lambda - 2k)]u_k.
\end{aligned}$$

Simplifying the last expression gives the lemma. □

Since $\pi(H)$ can have only finitely many eigenvalues, the u_k's cannot all be nonzero. There must, therefore, be a non-negative integer m such that

$$u_k = \pi(Y)^k u_0 \neq 0$$

for all $k \leq m$, but

$$u_{m+1} = \pi(Y)^{m+1} u_0 = 0.$$

Now, if $u_{m+1} = 0$, then, certainly, $\pi(X)u_{m+1} = 0$. Then, by Lemma 4.11,

$$0 = \pi(X)u_{m+1} = [(m+1)\lambda - m(m+1)]\, u_m = (m+1)(\lambda - m)u_m.$$

However, $u_m \neq 0$ and $m + 1 \neq 0$ (since $m \geq 0$). Thus, in order to have $(m+1)(\lambda - m)u_m$ equal to zero, we must have $\lambda = m$, where m is a nonnegative integer. (This also shows that the eigenvalue of $\pi(H)$ that we started with, $\alpha = \lambda - 2N$, must be an integer.)

We have made considerable progress. Given a finite-dimensional irreducible representation π of $\mathsf{sl}(2;\mathbb{C})$, acting on a space V, there exists an integer $m \geq 0$ and nonzero vectors $u_0, \ldots, u_m$ such that (putting λ equal to m)

$$\begin{aligned}
\pi(H)u_k &= (m-2k)u_k, \\
\pi(Y)u_k &= u_{k+1} \quad (k < m), \\
\pi(Y)u_m &= 0, \\
\pi(X)u_k &= [km - k(k-1)]u_{k-1} \quad (k > 0), \\
\pi(X)u_0 &= 0.
\end{aligned} \tag{4.10}$$

The vectors $u_0, \ldots, u_m$ must be linearly independent, since they are eigenvectors of $\pi(H)$ with distinct eigenvalues (Proposition B.1). Moreover, the $(m+1)$-dimensional span of $u_0, \ldots, u_m$ is explicitly invariant under $\pi(H)$, $\pi(X)$, and $\pi(Y)$ and, hence, under $\pi(Z)$ for all $Z \in \mathsf{sl}(2;\mathbb{C})$. Since π is irreducible, this space must be all of V.

We have now shown that every irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ is of the form (4.10). It remains to show that everything of the form (4.10) is a representation and that it is irreducible. That is, if we *define* $\pi(H)$, $\pi(X)$, and $\pi(Y)$ by (4.10) (where the u_k's are basis elements for some $(m+1)$-dimensional vector space), then we want to show that they have the right commutation relations to form a representation of $\mathsf{sl}(2;\mathbb{C})$ and that this representation is irreducible. One way to do this is to show that the representations π_m constructed in the previous section have a basis of the form (4.10). Alternatively, we can directly check that operators defined as in (4.10) really do satisfy the $\mathsf{sl}(2;\mathbb{C})$ commutation relations (Exercise 4), and then prove irreducibility in the same way as in the proof of Proposition 4.8.

We have now shown that there is an irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ in each dimension $m+1$, by writing explicitly (in (4.10)) how H, X, and Y should act in a basis. However, we have shown more than this. We also have shown that any $(m+1)$-dimensional irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ must be of the form (4.10). It follows that any two irreducible representations of $\mathsf{sl}(2;\mathbb{C})$ of dimension $(m+1)$ must be equivalent, for if π_1 and π_2 are two irreducible representations of dimension $(m+1)$, acting on spaces V_1 and V_2, then V_1 has a basis $u_0, \ldots, u_m$ as in (4.10) and V_2 has a similar basis

$\widetilde{u}_0, \ldots, \widetilde{u}_m$. However, then the map $\phi : V_1 \to V_2$ which sends u_k to $\widetilde{u}_k$ will be an isomorphism of representations, as a moment's thought will confirm.

In particular, the $(m+1)$-dimensional representation π_m described in Section 4.3 must be equivalent to (4.10). This can be seen explicitly by introducing the following basis for V_m:

$$u_k = [\pi_m(Y)]^k (z_2^m) = (-1)^k \frac{m!}{(m-k)!} z_1^k z_2^{m-k} \qquad (k \leq m).$$

Then, by definition, $\pi_m(Y)u_k = u_{k+1}$ $(k < m)$, and it is clear that $\pi_m(Y)u_m = 0$. It is easy to see that $\pi_m(H)u_k = (m-2k)u_k$. The only thing left to check is the behavior of $\pi_m(X)$. However, direct computation shows that

$$\pi_m(X)u_k = k(m-k+1)u_{k-1} = [km - k(k-1)]u_{k-1},$$

as required.

This completes the proof of Theorem 4.9. □

If we look carefully at the proof of Theorem 4.9, we see that the argument can tell us something about finite-dimensional, not necessarily irreducible representations of $\mathsf{sl}(2;\mathbb{C})$. In particular, up to and including (4.10), the argument does not use irreducibility, which is used only to show that the vectors in (4.10) span V. Thus, we obtain the following result about arbitrary finite-dimensional representations of $\mathsf{sl}(2;\mathbb{C})$.

Theorem 4.12. *Suppose π is any finite-dimensional, complex-linear representation of $\mathsf{sl}(2;\mathbb{C})$ acting on a space V. Then, we have the following results:*

1. *Every eigenvalue of $\pi(H)$ must be an integer.*
2. *If v is a nonzero element of V such that $\pi(X)v = 0$ and $\pi(H)v = \lambda v$, then there is a non-negative integer m such that $\lambda = m$. Furthermore, the vectors $v, \pi(Y)v, \ldots, \pi(Y)^m v$ are linearly independent and their span is an irreducible invariant subspace of dimension $m+1$.*

4.5 Direct Sums of Representations

One way of generating representations is to take some representations one knows and combine them in some fashion. In this section and the next two, we will consider the three standard methods of obtaining new representations from old, namely direct sums of representations, tensor products of representations, and dual representations.

Definition 4.13. *Let G be a matrix Lie group and let $\Pi_1, \Pi_2, \ldots, \Pi_m$ be representations of G acting on vector spaces $V_1, V_2, \ldots, V_m$. Then, the* ***direct sum*** *of $\Pi_1, \Pi_2, \ldots, \Pi_m$ is a representation $\Pi_1 \oplus \cdots \oplus \Pi_m$ of G acting on the space $V_1 \oplus \cdots \oplus V_m$, defined by*

$$[\Pi_1 \oplus \cdots \oplus \Pi_m(A)]\,(v_1, \ldots, v_m) = (\Pi_1(A)v_1, \ldots, \Pi_m(A)v_m)$$

for all $A \in G$.

Similarly, if $\mathfrak{g}$ *is a Lie algebra, and* $\pi_1, \pi_2, \ldots, \pi_m$ *are representations of* $\mathfrak{g}$ *acting on* $V_1, V_2, \ldots, V_m$, *then we define the* ***direct sum*** *of* $\pi_1, \pi_2, \ldots, \pi_m$, *acting on* $V_1 \oplus \cdots \oplus V_m$ *by*

$$[\pi_1 \oplus \cdots \oplus \pi_m(X)]\,(v_1, \ldots, v_m) = (\pi_1(X)v_1, \ldots, \pi_m(X)v_m)$$

for all $X \in \mathfrak{g}$.

It is straightforward to check that, say, $\Pi_1 \oplus \cdots \oplus \Pi_m$ is really a representation of G.

An important property that some matrix Lie groups and Lie algebras have is the *complete reducibility property.* This means that every finite-dimensional representation is isomorphic to a direct sum of irreducible representations. For such groups, once we know all the irreducible representations, we know all the representations. By no means do all groups have this property. We will discuss this issue further in Section 4.10 of this chapter and in Chapter 6.

4.6 Tensor Products of Representations

Let U and V be finite-dimensional real or complex vector spaces. We wish to define the **tensor product** of U and V, which will be a new vector space $U \otimes V$ "built" out of U and V. We will discuss the idea of this first and then give the precise definition.

We wish to consider a formal "product" of an element u of U with an element v of V, denoted $u \otimes v$. The *space* $U \otimes V$ is then the space of linear combinations of such products, that is, the space of elements of the form

$$a_1 u_1 \otimes v_1 + a_2 u_2 \otimes v_2 + \cdots + a_n u_n \otimes v_n. \tag{4.11}$$

Of course, if "$\otimes$" is to be interpreted as a product, then it should be bilinear; that is, we should have

$$(u_1 + a u_2) \otimes v = u_1 \otimes v + a u_2 \otimes v,$$
$$u \otimes (v_1 + a v_2) = u \otimes v_1 + a u \otimes v_2.$$

We do not assume that the product is commutative. (In fact, the product in the other order, $v \otimes u$, is in a different space, namely $V \otimes U$.)

Now, if $e_1, e_2, \ldots, e_n$ is a basis for U and $f_1, f_2, \ldots, f_m$ is a basis for V, then, using bilinearity, it is easy to see that any element of the form (4.11) can be written as a linear combination of the elements $e_i \otimes f_j$. In fact, it seems reasonable to expect that $\{e_i \otimes f_j \,|\, 1 \leq i \leq n, 1 \leq j \leq m\}$ should be a basis for the space $U \otimes V$. This, in fact, turns out to be the case.

Definition 4.14. *If U and V are finite-dimensional real or complex vector spaces, then a* ***tensor product*** *of U with V is a vector space W, together with a bilinear map $\phi : U \times V \to W$ with the following property: If ψ is any bilinear map of $U \times V$ into a vector space X, then there exists a unique linear map $\widetilde{\psi}$ of W into X such that the following diagram commutes:*

$$\begin{array}{ccc} U \times V & \xrightarrow{\phi} & W \\ & {\scriptstyle\psi}\searrow \quad \swarrow{\scriptstyle\widetilde{\psi}} & \\ & X & \end{array}.$$

Note that the *bilinear* map ψ from $U \times V$ into X turns into the *linear* map $\widetilde{\psi}$ of W into X. This is one of the points of tensor products: Bilinear maps on $U \times V$ turn into linear maps on W.

Theorem 4.15. *If U and V are any finite-dimensional real or complex vector spaces, then a tensor product (W, ϕ) exists. Furthermore, (W, ϕ) is unique up to canonical isomorphism. That is, if (W_1, ϕ_1) and (W_2, ϕ_2) are two tensor products, then there exists a unique vector space isomorphism $\Phi : W_1 \to W_2$ such that the following diagram commutes:*

$$\begin{array}{ccc} U \times V & \xrightarrow{\phi_1} & W_1 \\ & {\scriptstyle\phi_2}\searrow \quad \swarrow{\scriptstyle\Phi} & \\ & W_2 & \end{array}.$$

Suppose that (W, ϕ) is a tensor product and that $e_1, e_2, \ldots, e_n$ is a basis for U and $f_1, f_2, \ldots, f_m$ is a basis for V. Then, $\{\phi(e_i, f_j) \,|\, 1 \le i \le n, 1 \le j \le m\}$ is a basis for W.

Proof. Exercise 10. □

Notation 4.16 *Since the tensor product of U and V is essentially unique, we will let $U \otimes V$ denote an arbitrary tensor product space and we will write $u \otimes v$ instead of $\phi(u, v)$. In this notation, Theorem 4.15 says that $\{e_i \otimes f_j \,|\, 1 \le i \le n, 1 \le j \le m\}$ is a basis for $U \otimes V$, as expected. Note in particular that*

$$\dim(U \otimes V) = (\dim U)(\dim V)$$

(not $\dim U + \dim V$).

The defining property of $U \otimes V$ is called the **universal property** of tensor products. Although it may seem that we are taking a simple idea and making it confusing, in fact there is a point to this universal property. Suppose we want to define a linear map T from $U \otimes V$ into some other space. The most sensible way to define this is to define T on elements of the form $u \otimes v$. (We might try defining it on a basis, but this would force us to worry about whether things depend on the choice of basis.) Now, every element of $U \otimes V$ is a linear combination of things of the form $u \otimes v$. However, this representation

is far from unique. (Since, say, if $u = u_1 + u_2$, then one can rewrite $u \otimes v$ as $u_1 \otimes v + u_2 \otimes v$.)

Thus, if we try to define T by what it does to elements of the form $u \otimes v$, we have to worry about whether T is well defined. This is where the universal property comes in. Suppose that $\psi(u, v)$ is some bilinear expression in (u, v). Then, the universal property says precisely that there is a unique linear map T $(= \widetilde{\psi})$ such that

$$T(u \otimes v) = \psi(u, v).$$

The conclusion is this: We can define a linear map T on $U \otimes V$ by defining it on elements of the form $u \otimes v$, and this will be well defined, *provided* that $T(u \otimes v)$ is bilinear in (u, v). The following proposition illustrates how to make use of this idea.

Proposition 4.17. *Let U and V be finite-dimensional real or complex vector spaces. Let $A : U \to U$ and $B : V \to V$ be linear operators. Then, there exists a unique linear operator from $U \otimes V$ to $U \otimes V$, denoted $A \otimes B$, such that*

$$(A \otimes B)(u \otimes v) = (Au) \otimes (Bv)$$

for all $u \in U$ and $v \in V$.

If A_1 and A_2 are linear operators on U and B_1 and B_2 are linear operators on V, then

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2).$$

Proof. Define a map ψ from $U \times V$ into $U \otimes V$ by

$$\psi(u, v) = (Au) \otimes (Bv).$$

Since A and B are linear and since $\otimes$ is bilinear, ψ will be a bilinear map of $U \times V$ into $U \otimes V$. However, then the universal property says that there is an associated linear map $\widetilde{\psi} : U \otimes V \to U \otimes V$ such that

$$\widetilde{\psi}(u \otimes v) = \psi(u, v) = (Au) \otimes (Bv).$$

Then, $\widetilde{\psi}$ is the desired map $A \otimes B$.

Now, if A_1 and A_2 are operators on U and B_1 and B_2 are operators on V, then compute that

$$\begin{aligned}(A_1 \otimes B_1)(A_2 \otimes B_2)(u \otimes v) &= (A_1 \otimes B_1)(A_2 u \otimes B_2 v) \\ &= A_1 A_2 u \otimes B_1 B_2 v.\end{aligned}$$

This shows that $(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$ are equal on elements of the form $u \otimes v$. Since every element of $U \otimes V$ can be written as a linear combination of things of the form $u \otimes v$ (in fact, of $e_i \otimes f_j$), $(A_1 \otimes B_1)(A_2 \otimes B_2)$ and $(A_1 A_2) \otimes (B_1 B_2)$ must be equal on the whole space. □

We are now ready to define tensor products of representations. There are two different approaches to this, both of which are important. The first approach starts with a representation of a group G acting on a space V and a representation of another group H acting on a space U and produces a representation of the product group $G \times H$ acting on the space $U \otimes V$. The second approach starts with two different representations of the same group G, acting on spaces U and V, and produces a representation of G acting on $U \otimes V$. Both of these approaches can be adapted to apply to Lie algebras.

Definition 4.18. *Let G and H be matrix Lie groups. Let Π_1 be a representation of G acting on a space U and let Π_2 be a representation of H acting on a space V. Then, the* ***tensor product*** *of Π_1 and Π_2 is a representation $\Pi_1 \otimes \Pi_2$ of $G \times H$ acting on $U \otimes V$ defined by*

$$\Pi_1 \otimes \Pi_2(A, B) = \Pi_1(A) \otimes \Pi_2(B)$$

for all $A \in G$ and $B \in H$.

Using the above proposition, it is easy to check that, indeed, $\Pi_1 \otimes \Pi_2$ is a representation of $G \times H$.

Now, if G and H are matrix Lie groups (i.e., G is a closed subgroup of $\mathsf{GL}(n;\mathbb{C})$ and H is a closed subgroup of $\mathsf{GL}(m;\mathbb{C})$), then $G \times H$ can be regarded in an obvious way as a closed subgroup of $\mathsf{GL}(n+m;\mathbb{C})$. Thus, the direct product of matrix Lie groups can be regarded as a matrix Lie group. It is easy to check that the Lie algebra of $G \times H$ is isomorphic to the direct sum of the Lie algebra of G and the Lie algebra of H.

In light of Proposition 4.4, the representation $\Pi_1 \otimes \Pi_2$ of $G \times H$ gives rise to a representation of the Lie algebra of $G \times H$, namely $\mathfrak{g} \oplus \mathfrak{h}$. The following proposition shows that this representation of $\mathfrak{g} \oplus \mathfrak{h}$ is not what one might expect at first.

Proposition 4.19. *Let G and H be matrix Lie groups, let Π_1 and Π_2 be representations of G and H, respectively, and consider the representation $\Pi_1 \otimes \Pi_2$ of $G \times H$. Let $\pi_1 \otimes \pi_2$ denote the associated representation of the Lie algebra of $G \times H$, namely $\mathfrak{g} \oplus \mathfrak{h}$. Then, for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$,*

$$\pi_1 \otimes \pi_2(X, Y) = \pi_1(X) \otimes I + I \otimes \pi_2(Y).$$

Proof. Suppose that $u(t)$ is a smooth curve in U and $v(t)$ is a smooth curve in V. Then, we verify the product rule in the usual way:

$$\begin{aligned}
&\lim_{h \to 0} \frac{u(t+h) \otimes v(t+h) - u(t) \otimes v(t)}{h} \\
&= \lim_{h \to 0} \frac{u(t+h) \otimes v(t+h) - u(t+h) \otimes v(t)}{h} + \frac{u(t+h) \otimes v(t) - u(t) \otimes v(t)}{h} \\
&= \lim_{h \to 0} \left[u(t+h) \otimes \frac{(v(t+h) - v\,(t))}{h} \right] + \lim_{h \to 0} \left[\frac{(u(t+h) - u\,(t))}{h} \otimes v(t) \right].
\end{aligned}$$

Thus,

$$\frac{d}{dt}\left(u(t)\otimes v(t)\right)=\frac{du}{dt}\otimes v(t)+u(t)\otimes\frac{dv}{dt}.$$

This being the case, we can compute $\pi_1\otimes\pi_2(X,Y)$:

$$\begin{aligned}\pi_1\otimes\pi_2(X,Y)(u\otimes v)&=\frac{d}{dt}\Pi_1\otimes\Pi_2(e^{tX},e^{tY})(u\otimes v)\Big|_{t=0}\\&=\frac{d}{dt}\Pi_1(e^{tX})u\otimes\Pi_2(e^{tY})v\Big|_{t=0}\\&=\left(\frac{d}{dt}\Pi_1(e^{tX})u\Big|_{t=0}\right)\otimes v+u\otimes\left(\frac{d}{dt}\Pi_2(e^{tY})v\Big|_{t=0}\right).\end{aligned}$$

This shows that $\pi_1\otimes\pi_2(X,Y)=\pi_1(X)\otimes I+I\otimes\pi_2(Y)$ on elements of the form $u\otimes v$ and, therefore, on the whole space $U\otimes V$. □

Definition 4.20. *Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras and let π_1 and π_2 be representations of $\mathfrak{g}$ and $\mathfrak{h}$, acting on spaces U and V. Then, the* ***tensor product*** *of π_1 and π_2, denoted $\pi_1\otimes\pi_2$, is a representation of $\mathfrak{g}\oplus\mathfrak{h}$ acting on $U\otimes V$, given by*

$$\pi_1\otimes\pi_2(X,Y)=\pi_1(X)\otimes I+I\otimes\pi_2(Y)$$

for all $X\in\mathfrak{g}$ and $Y\in\mathfrak{h}$.

It is easy to check that this indeed defines a representation of $\mathfrak{g}\oplus\mathfrak{h}$. Note that if we defined $\pi_1\otimes\pi_2(X,Y)=\pi_1(X)\otimes\pi_2(Y)$, this would *not* be a representation of $\mathfrak{g}\oplus\mathfrak{h}$, for this is not even a linear map (e.g., we would then have $\pi_1\otimes\pi_2(2X,2Y)=4\pi_1\otimes\pi_2(X,Y)$). Note also that the above definition applies even if π_1 and π_2 do not come from a representation of any matrix Lie group.

Definition 4.21. *Let G be a matrix Lie group and let Π_1 and Π_2 be representations of G, acting on spaces V_1 and V_2. Then, the* ***tensor product*** *of Π_1 and Π_2 is a representation of G acting on $V_1\otimes V_2$ defined by*

$$\Pi_1\otimes\Pi_2(A)=\Pi_1(A)\otimes\Pi_2(A)$$

for all $A\in G$.

Proposition 4.22. *With the above notation, the associated representation of the Lie algebra $\mathfrak{g}$ satisfies*

$$\pi_1\otimes\pi_2(X)=\pi_1(X)\otimes I+I\otimes\pi_2(X)$$

for all $X\in\mathfrak{g}$.

Proof. Using the product rule,

$$\begin{aligned}\pi_1 \otimes \pi_2(X)(u \otimes v) &= \frac{d}{dt}\Big|_{t=0} \Pi_1\left(e^{tX}\right) u \otimes \Pi_2\left(e^{tX}\right) v \\ &= \pi_1(X)\, u \otimes v + v \otimes \pi_2(X)\, u.\end{aligned}$$

This is what we wanted to show. □

Definition 4.23. *If $\mathfrak{g}$ is a Lie algebra and π_1 and π_2 are representations of $\mathfrak{g}$ acting on spaces V_1 and V_2, then the* ***tensor product*** *of π_1 and π_2 is a representation of $\mathfrak{g}$ acting on the space $V_1 \otimes V_2$ defined by*

$$\pi_1 \otimes \pi_2(X) = \pi_1(X) \otimes I + I \otimes \pi_2(X)$$

for all $X \in \mathfrak{g}$.

It is easy to check that $\Pi_1 \otimes \Pi_2$ and $\pi_1 \otimes \pi_2$ are actually representations of G and $\mathfrak{g}$, respectively. There is some ambiguity in the notation, say, $\Pi_1 \otimes \Pi_2$. After all, even if Π_1 and Π_2 are both representations of the same group G, we could still regard $\Pi_1 \otimes \Pi_2$ as a representation of $G \times G$, by taking $H = G$ in Definition 4.18. We will rely on context to make clear whether we are thinking of $\Pi_1 \otimes \Pi_2$ as a representation of $G \times G$ or as representation of G.

Suppose Π_1 and Π_2 are *irreducible* representations of a group G. If we regard $\Pi_1 \otimes \Pi_2$ as a representation of G, it may no longer be irreducible. If it is not irreducible, one can attempt to decompose it as a direct sum of irreducible representations. This process is called the **Clebsch–Gordan** theory. In the case of $\mathrm{SU}(2)$, this theory is relatively simple. (In the physics literature, the problem of analyzing tensor products of representations of $\mathrm{SU}(2)$ is called "addition of angular momentum.") See Exercise 11 and Appendix D.

4.7 Dual Representations

Suppose that π is a representation of a Lie algebra $\mathfrak{g}$ acting on a finite-dimensional vector space V. Let V^* denote the dual space of V, that is, the space of linear functionals on V. (See Section B.7.) If A is a linear operator on V, let A^{tr} denote the dual or transpose operator on V^*,

$$\left(A^{tr}\phi\right)(v) = \phi\left(Av\right)$$

for $\phi \in V^*$, $v \in V$. If $v_1, \ldots, v_n$ is a basis for V, then there is a naturally associated "dual basis" $\phi_1, \ldots, \phi_n$ with the property that $\phi_k(v_l) = \delta_{kl}$. Then, the matrix for A^{tr} in the dual basis is simply the transpose (in the usual matrix sense) of the matrix of A in the original basis. Note that the matrix of A^{tr} is the transpose of the matrix of A and *not* the conjugate transpose. If A and B are linear operators on V, then

$$(AB)^{tr} = B^{tr}A^{tr}.$$

Definition 4.24. *Suppose G is a matrix Lie group and Π is a representation of G acting on a finite-dimensional vector space V. Then, the* ***dual representation*** *Π^* to Π is the representation of G acting on V^* given by*

$$\Pi^*(g) = [\Pi(g^{-1})]^{tr}.$$

Similarly, if π is a representation of a Lie algebra $\mathfrak{g}$ acting on a finite-dimensional vector space V, then π^* is the representation of $\mathfrak{g}$ acting on V^* given by

$$\pi^*(X) = -\pi(X)^{tr}.$$

Note that since the transpose is an order-reversing operation, we cannot simply define $\Pi^*(g) = \Pi(g)^{tr}$. This would not be a representation; we need the inverse in Π^* and the minus sign in π^* in order for the dual representations to actually be representations. The dual representation is also called **contragredient representation**.

The main properties of dual representations are summarized in the following elementary proposition, whose proof is left as an exercise to the reader (Exercise 7).

Proposition 4.25. *If Π is a representation of a matrix Lie group G, then (1) Π^* is irreducible if and only if Π is irreducible and (2) $(\Pi^*)^*$ is isomorphic to Π. Similar statements apply to Lie algebra representations.*

4.8 Schur's Lemma

Let Π and Σ be representations of a matrix Lie group G, acting on spaces V and W. Recall that an intertwining map of representations is a linear map $\phi : V \to W$ with the property that

$$\phi(\Pi(A)v) = \Sigma(A)(\phi(v))$$

for all $v \in V$ and all $A \in G$. Schur's Lemma is an extremely important result which tells us about intertwining maps of irreducible representations. Part of Schur's Lemma applies to both real and complex representations, but part of it applies only to complex representations.

It is desirable to be able to state Schur's Lemma simultaneously for groups and Lie algebras. In order to do so, we need to indulge in a common abuse of notation. If, say, Π is a representation of G acting on a space V, we will refer to V as the representation, without explicit reference to Π.

Theorem 4.26 (Schur's Lemma).

1. *Let V and W be irreducible real or complex representations of a group or Lie algebra and let $\phi : V \to W$ be an intertwining map. Then, either $\phi = 0$ or ϕ is an isomorphism.*

2. *Let V be an irreducible complex representation of a group or Lie algebra and let $\phi : V \to V$ be an intertwining map of V with itself. Then, $\phi = \lambda I$, for some $\lambda \in \mathbb{C}$.*
3. *Let V and W be irreducible complex representations of a group or Lie algebra and let $\phi_1, \phi_2 : V \to W$ be nonzero intertwining maps. Then, $\phi_1 = \lambda \phi_2$, for some $\lambda \in \mathbb{C}$.*

Before proving Schur's Lemma, we obtain two corollaries of it.

Corollary 4.27. *Let Π be an irreducible complex representation of a matrix Lie group G. If A is in the center of G, then $\Pi(A) = \lambda I$. Similarly, if π is an irreducible complex representation of a Lie algebra $\mathfrak{g}$ and if X is in the center of $\mathfrak{g}$ (i.e., $[X, Y] = 0$ for all $Y \in \mathfrak{g}$), then $\pi(X) = \lambda I$.*

Proof. We prove the group case; the proof of the Lie algebra case is similar. If A is in the center of G, then for all $B \in G$,

$$\Pi(A)\Pi(B) = \Pi(AB) = \Pi(BA) = \Pi(B)\Pi(A).$$

However, this says exactly that $\Pi(A)$ is an intertwining map of the space with itself. So by Point 2 of Schur's Lemma, $\Pi(A)$ is a multiple of the identity. □

Corollary 4.28. *An irreducible complex representation of a commutative group or Lie algebra is one dimensional.*

Proof. Again, we prove only the group case. If G is commutative, then the center of G is all of G, so by the previous corollary $\Pi(A)$ is a multiple of the identity for each $A \in G$. However, this means that *every* subspace of V is invariant! Thus, the only way that V can fail to have a nontrivial invariant subspace is for it not to have any nontrivial subspaces. This means that V must be one dimensional. (Recall that we do not allow V to be zero dimensional.) □

We now provide the proof of Schur's Lemma.

Proof. As usual, we will prove just the group case; the proof of the Lie algebra case requires only the obvious notational changes.

Proof of Point 1. Saying that ϕ is an intertwining map means $\phi(\Pi(A)v) = \Sigma(A)(\phi(v))$ for all $v \in V$ and all $A \in G$. Now, suppose that $v \in \ker(\phi)$. Then,

$$\phi(\Pi(A)v) = \Sigma(A)\phi(v) = 0.$$

This shows that $\ker \phi$ is an invariant subspace of V. Since V is irreducible, we must have $\ker \phi = 0$ or $\ker \phi = V$. Thus, ϕ is either one-to-one or zero.

Suppose ϕ is one-to-one. Then, the image of ϕ is a nonzero subspace of W. On the other hand, the image of ϕ is invariant, for if $w \in W$ is of the form $\phi(v)$ for some $v \in V$, then

$$\Sigma(A)w = \Sigma(A)\phi(v) = \phi(\Pi(A)v).$$

Since W is irreducible and image(V) is nonzero and invariant, we must have image$(V) = W$. Thus, ϕ is either zero or one-to-one and onto.

Proof of Point 2. Suppose now that V is an irreducible *complex* representation and that $\phi : V \to V$ is an intertwining map of V to itself. This means that $\phi\Pi(A) = \Pi(A)\phi$ for all $A \in G$ (i.e., that ϕ commutes with all of the $\Pi(A)$'s). Now, since we are working over an algebraically closed field, ϕ must have at least one eigenvalue $\lambda \in \mathbb{C}$. Let U denote the eigenspace for ϕ associated to the eigenvalue λ and let $u \in U$. Then, for each $A \in G$,

$$\phi\left(\Pi(A)u\right) = \Pi(A)\phi(u) = \lambda\Pi(A)u.$$

Thus, applying $\Pi(A)$ to an element of the λ-eigenspace of ϕ yields another element of the λ-eigenspace. Thus, U is invariant.

Since λ is an eigenvalue, $U \neq 0$, and so we must have $U = V$. This means that $\phi(v) = \lambda v$ for all $v \in V$ (i.e., that $\phi = \lambda I$).

Proof of Point 3. If $\phi_2 \neq 0$, then by Point 1, ϕ_2 is an isomorphism. Now, look at $\phi_1 \circ \phi_2^{-1}$. As is easily checked, the composition of two intertwining maps is an intertwining map, so $\phi_1 \circ \phi_2^{-1}$ is an intertwining map of W with itself. Thus, by Point 2, $\phi_1 \circ \phi_2^{-1} = \lambda I$, whence $\phi_1 = \lambda\phi_2$. □

4.9 Group Versus Lie Algebra Representations

We know from Chapter 2 (Theorem 2.21) that every Lie group homomorphism gives rise to a Lie algebra homomorphism. In particular, this shows that every representation of a matrix Lie group gives rise to a representation of the associated Lie algebra. In the case of a simply-connected matrix Lie group G, we have the converse: A Lie algebra homomorphism gives rise to a Lie group homomorphism (Theorem 3.7). This means, in particular, that for a simply-connected matrix Lie group G, there is a natural one-to-one correspondence between the representations of G and the representations of the Lie algebra $\mathfrak{g}$. For non-simply-connected groups, there may be Lie algebra representations for which there is no associated Lie group representation.

It is instructive to see how this general theory works out in the case of $\mathsf{SU}(2)$ (which is simply connected) and $\mathsf{SO}(3)$ (which is not). We have shown (Theorem 4.9) that every irreducible complex representation of $\mathsf{su}(2)$ is equivalent to one of the representations π_m described in Section 4.3. (Recall that the irreducible complex representations of $\mathsf{su}(2)$ are in one-to-one correspondence with the irreducible representations of $\mathsf{sl}(2;\mathbb{C})$.) Each of the representations π_m of $\mathsf{su}(2)$ was constructed from the corresponding representation Π_m of the group $\mathsf{SU}(2)$. Thus, we see, by brute-force computation, that every irreducible complex representation of $\mathsf{su}(2)$ actually comes from a representation of the group $\mathsf{SU}(2)$! This is consistent with the fact that $\mathsf{SU}(2)$ is simply connected (Proposition 1.14).

Let us now consider the situation for $\mathsf{SO}(3)$, which is not simply connected. (See Section 1.5 and Appendix E.) We know from Exercise 16 of Chapter 2 that the Lie algebras $\mathsf{su}(2)$ and $\mathsf{so}(3)$ are isomorphic. In particular, if we take the basis

$$E_1 = \tfrac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \; E_2 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \; E_3 = \tfrac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

for $\mathsf{su}(2)$ and the basis

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

for $\mathsf{so}(3)$, then direct computation shows that $[E_1, E_2] = E_3$, $[E_2, E_3] = E_1$, and $[E_3, E_1] = E_2$, and similarly with the E's replaced by the F's. Thus, the linear map $\phi : \mathsf{su}(2) \to \mathsf{so}(3)$ which takes E_i to F_i will be a Lie algebra isomorphism.

Since $\mathsf{su}(2)$ and $\mathsf{so}(3)$ are isomorphic Lie algebras, they must have "the same" representations. Specifically, if π is a representation of $\mathsf{su}(2)$, then $\pi \circ \phi^{-1}$ will be a representation of $\mathsf{so}(3)$, and every representation of $\mathsf{so}(3)$ is of this form. In particular, the irreducible representations of $\mathsf{so}(3)$ are precisely of the form $\sigma_m = \pi_m \circ \phi^{-1}$. We wish to determine, for a particular m, whether there is a representation Σ_m of the group $\mathsf{SO}(3)$ such that $\Sigma_m(\exp X) = \exp(\sigma_m(X))$ for all X in $\mathsf{so}(3)$.

Proposition 4.29. *Let $\sigma_m = \pi_m \circ \phi^{-1}$ be the irreducible complex representations of the Lie algebra $\mathsf{so}(3)$ $(m \geq 0)$. If m is even, then there is a representation Σ_m of the group $\mathsf{SO}(3)$ such that $\Sigma_m(\exp X) = \exp(\sigma_m(X))$ for all X in $\mathsf{so}(3)$. If m is odd, then there is no such representation of $\mathsf{SO}(3)$.*

Note that the condition that m be even is equivalent to the condition that $\dim V_m = m + 1$ be odd. Thus, it is the odd-dimensional representations of the Lie algebra $\mathsf{so}(3)$ which come from group representations.

In the physics literature, the representations of $\mathsf{su}(2) \cong \mathsf{so}(3)$ are labeled by the parameter $l = m/2$. In terms of this notation, a representation of $\mathsf{so}(3)$ comes from a representation of $\mathsf{SO}(3)$ if and only if l is an integer. The representations with l an integer are called "integer spin"; the others are called "half-integer spin." If one attempts to construct Σ_m by the construction in the proof of Theorem 3.7, then one finds that not all paths are homotopic and the value of the would-be homomorphism Σ_m can depend on the path. Consider, for example, the path in $\mathsf{SO}(3)$ consisting of rotations by angle $2\pi t$ in the (x, y)-plane, which comes back to the identity when $t = 1$. It can be shown that this path is not homotopic to the constant path. If one defines Σ_m along the constant path, then one gets the value $\Sigma_m(I) = I$, as expected. If m is odd, however, and one defines Σ_m along the path of rotations in the (x, y)-plane, then one gets the value $\Sigma_m(I) = -I$. This strongly suggests (and Proposition 4.29 confirms) that there is no way to define Σ_m (m odd)

as a "single-valued" representation of $\mathsf{SO}(3)$. An electron, for example, is a "spin $\frac{1}{2}$" particle, which means that it is described in quantum mechanics in a way that involves the representation σ_1 of $\mathsf{so}(3)$. In the quantum mechanics literature, one finds statements to the effect that performing a 360° rotation on the wave function of the electron gives back the negative of the original wave function. This statement reflects that if one attempts to construct the nonexistent representation Σ_1 of $\mathsf{SO}(3)$, then when defining Σ_1 along a path of rotations in some plane, one gets that $\Sigma_1(I) = -I$.

Proof. Case 1: m odd. In this case, we want to prove that there is no representation Σ_m such that $\Sigma_m(\exp X) = \exp(\sigma_m(X))$ for all X in $\mathsf{so}(3)$. Suppose, to the contrary, that there is such a Σ_m. Then, take $X = 2\pi F_1$. Computing as in Section 2.2, we see that

$$e^{2\pi F_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\pi & -\sin 2\pi \\ 0 & \sin 2\pi & \cos 2\pi \end{pmatrix} = I.$$

Thus, on the one hand, $\Sigma_m\left(e^{2\pi F_1}\right) = \Sigma_m(I) = I$, whereas, on the other hand, $\Sigma_m\left(e^{2\pi F_1}\right) = e^{2\pi\sigma_m(F_1)}$.

Let us compute $e^{2\pi\sigma_m(F_1)}$. By definition, $\sigma_m(F_1) = \pi_m(\phi^{-1}(F_1)) = \pi_m(E_1)$. However, $E_1 = \frac{i}{2}H$, where, as usual,

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

We know that there is a basis $u_0, u_1, \ldots, u_m$ for V_m such that u_k is an eigenvector for $\pi_m(H)$ with eigenvalue $m - 2k$. This means that u_k is also an eigenvector for $\sigma_m(F_1) = \frac{i}{2}\pi_m(H)$, with eigenvalue $\frac{i}{2}(m - 2k)$. Thus, in the basis $\{u_k\}$, we have

$$\sigma_m(F_1) = \begin{pmatrix} \frac{i}{2}m & & & \\ & \frac{i}{2}(m-2) & & \\ & & \ddots & \\ & & & \frac{i}{2}(-m) \end{pmatrix}.$$

But we are assuming that m is odd! This means that $m - 2k$ is an odd integer. Thus, $e^{2\pi\frac{i}{2}(m-2k)} = -1$, and in the basis $\{u_k\}$

$$e^{2\pi\sigma_m(F_1)} = \begin{pmatrix} e^{2\pi\frac{i}{2}m} & & & \\ & e^{2\pi\frac{i}{2}(m-2)} & & \\ & & \ddots & \\ & & & e^{2\pi\frac{i}{2}(-m)} \end{pmatrix} = -I.$$

Thus, on the one hand, $\Sigma_m\left(e^{2\pi F_1}\right) = \Sigma_m(I) = I$, whereas, on the other hand, $\Sigma_m\left(e^{2\pi F_1}\right) = e^{2\pi\sigma_m(F_1)} = -I$. This is a contradiction, so there can be no such group representation Σ_m.

Case 2: m is even. We will use the following:

Lemma 4.30. *There exists a Lie group homomorphism* $\Phi : \mathsf{SU}(2) \to \mathsf{SO}(3)$ *such that*

(1) Φ *maps* $\mathsf{SU}(2)$ *onto* $\mathsf{SO}(3)$,

(2) $\ker \Phi = \{I, -I\}$, *and*

(3) the associated Lie algebra homomorphism is the map $\phi : \mathsf{su}(2) \to \mathsf{so}(3)$ *described earlier, namely the one satisfying* $\phi(E_i) = F_i$ *(*$i = 1, 2, 3$*).*

Proof. Exercise 12. □

Now, consider the representations Π_m of $\mathsf{SU}(2)$. I claim that if m is even, then $\Pi_m(-I) = I$. To see this, note that

$$e^{2\pi E_1} = \exp\begin{pmatrix} \pi i & 0 \\ 0 & -\pi i \end{pmatrix} = -I.$$

Thus, $\Pi_m(-I) = \Pi_m(e^{2\pi E_1}) = e^{\pi_m(2\pi E_1)}$. However, as in Case 1,

$$e^{\pi_m(2\pi E_1)} = \begin{pmatrix} e^{2\pi \frac{i}{2} m} & & & \\ & e^{2\pi \frac{i}{2}(m-2)} & & \\ & & \ddots & \\ & & & e^{2\pi \frac{i}{2}(-m)} \end{pmatrix}.$$

Only, this time, m is even, and so $\frac{i}{2}(m - 2k)$ is an integer, so that $\Pi_m(-I) = e^{\pi_m(2\pi E_1)} = I$.

Since $\Pi_m(-I) = I$, $\Pi_m(-U) = \Pi_m(U)$ for all $U \in \mathsf{SU}(2)$. According to Lemma 4.30, for each $R \in \mathsf{SO}(3)$, there is a unique pair of elements $\{U, -U\}$ such that $\Phi(U) = \Phi(-U) = R$. Since $\Pi_m(U) = \Pi_m(-U)$, it makes sense to define

$$\Sigma_m(R) = \Pi_m(U).$$

It is easy to see that Σ_m is a Lie group homomorphism (hence, a representation). By construction, we have

$$\Pi_m = \Sigma_m \circ \Phi. \tag{4.12}$$

Now, if σ_m denotes the Lie algebra representation associated to Σ_m, then it follows from (4.12) that

$$\pi_m = \sigma_m \circ \phi.$$

However, the Lie algebra homomorphism ϕ takes E_i to F_i, so $\pi_m = \sigma_m \circ \phi$, or $\sigma_m = \pi_m \circ \phi^{-1}$. Thus, Σ_m is the desired representation of $\mathsf{SO}(3)$. □

4.10 Complete Reducibility

Definition 4.31. *A finite-dimensional representation of a group or Lie algebra is said to be* ***completely reducible*** *if it is isomorphic to a direct sum of a finite number of irreducible representations.*

Definition 4.32. *A group or Lie algebra is said to have the* ***complete reducibility property*** *if every finite-dimensional representation of it is completely reducible.*

As we will see in Chapter 6, the complete reducibility property is a very special one that most groups and Lie algebras do not have. If a group or Lie algebra does have the complete reducibility property, then the study of its representations reduces to the study of its *irreducible* representations, which simplifies the analysis considerably.

Proposition 4.33. *If V is a completely reducible representation of a group or Lie algebra, then the following properties hold.*

1. *Every invariant subspace of V is completely reducible.*
2. *Given any invariant subspace U of V, there is another invariant subspace $\tilde{U}$ such that V is the direct sum of U and $\tilde{U}$.*

The proof of this result is tedious but elementary, requiring only Schur's Lemma and basic linear algebra. Exercises 13 through 18 guide the reader through the proof. We will prove later that every finite or compact group has the complete reducibility property. The proof shows directly (i.e., without appealing to Proposition 4.33) that representations of finite and compact groups have Properties 1 and 2 of the proposition, in addition to being completely reducible.

Proposition 4.34. *Let G be a matrix Lie group. Let Π be a finite-dimensional* ***unitary*** *representation of G, acting on a finite-dimensional real or complex Hilbert space V. Then, Π is completely reducible.*

Proof. Let V denote the (finite-dimensional!) Hilbert space on which Π acts and let $\langle \cdot, \cdot \rangle$ denote the inner product on V. Now, let $W \subset V$ be an invariant subspace. Let $W^\perp$ be the orthogonal complement of W; that is, $W^\perp$ is the space of all vectors v in V such that $\langle v, w \rangle = 0$ for all w in W. Then, V is the direct sum of W and $W^\perp$.

I claim that $W^\perp$ is also an invariant subspace. To see this, note that since Π is unitary, $\Pi(A)^* = \Pi(A)^{-1} = \Pi(A^{-1})$ for all $A \in G$. Then, for any $w \in W$ and any $v \in W^\perp$, we have

$$\begin{aligned} \langle \Pi(A)v, w \rangle &= \langle v, \Pi(A)^* w \rangle = \langle v, \Pi(A^{-1})w \rangle \\ &= \langle v, w' \rangle = 0. \end{aligned}$$

In the last step, we have used that $w' = \Pi(A^{-1})w$ is in W, since W is invariant. This shows that $\Pi(A)v$ is orthogonal to every element of W (i.e., that $\Pi(A)v \in W^\perp$).

We have established, then, that for unitary representations, the orthogonal complement of an invariant subspace is, again, invariant. Suppose now that V is not irreducible. Then, we can find an invariant subspace W that is

neither $\{0\}$ nor V, and we decompose V as $W \oplus W^\perp$. Then, W and $W^\perp$ are both invariant subspaces and, thus, unitary representations of G in their own right. Then, W is either irreducible or it splits as an orthogonal direct sum of invariant subspaces, and similarly for $W^\perp$. We continue this process, and since V is finite dimensional, it cannot go on forever. Each time, the dimensions of the spaces get smaller, so eventually we must get irreducible pieces—when the dimension reaches one if not sooner. Thus, we eventually succeed in decomposing V as a direct sum of irreducible invariant subspaces. □

Proposition 4.35. *Every finite group has the complete reducibility property.*

Proof. Suppose that Π is a representation of G, acting on a space V. Choose an arbitrary inner product $\langle \cdot, \cdot \rangle$ on V. Then, define a new inner product $\langle \cdot, \cdot \rangle_G$ on V by

$$\langle v_1, v_2 \rangle_G = \sum_{g \in G} \langle \Pi(g) v_1, \Pi(g) v_2 \rangle .$$

It is very easy to check that indeed $\langle \cdot, \cdot \rangle_G$ is an inner product. Furthermore, if $h \in G$, then

$$\begin{aligned} \langle \Pi(h) v_1, \Pi(h) v_2 \rangle_G &= \sum_{g \in G} \langle \Pi(g)\Pi(h) v_1, \Pi(g)\Pi(h) v_2 \rangle \\ &= \sum_{g \in G} \langle \Pi(gh) v_1, \Pi(gh) v_2 \rangle . \end{aligned}$$

However, as g ranges over G, so does gh. Thus, in fact,

$$\langle \Pi(h) v_1, \Pi(h) v_2 \rangle_G = \langle v_1, v_2 \rangle_G ;$$

that is, Π is a unitary representation with respect to the inner product $\langle \cdot, \cdot \rangle_G$. Thus, Π is isomorphic to a direct sum of irreducibles, by Proposition 4.34. □

There is a variant of the above argument which can be used to prove the following result:

Proposition 4.36. *If G is a compact matrix Lie group, G has the complete reducibility property.*

The argument below is sometimes called "Weyl's unitarian trick."

Proof. This proof requires the notion of Haar measure. (See, for example, Chapter VIII of Knapp (1996) or Section C.4.)

A **left Haar measure** on a matrix Lie group G is a nonzero measure μ on the Borel σ-algebra in G with the following two properties: (1) It is locally finite (i.e., every point in G has a neighborhood with finite measure); (2) it is left-translation invariant. Left-translation invariance means that $\mu(gE) = \mu(E)$ for all $g \in G$ and for all Borel sets $E \subset G$, where

$$gE = \{ge \,|\, e \in E\}\,.$$

It is a fact, which we cannot prove here, that every matrix Lie group has a left Haar measure and that this measure is unique up to multiplication by a constant. (One can analogously define right Haar measure, and a similar theorem holds for it. Left Haar measure and right Haar measure may or may not coincide; a group for which they do is called **unimodular**.)

Now, the key fact for our purpose is that left Haar measure is finite if and only if the group G is compact. Suppose, then, that Π is a finite-dimensional representation of a compact group G acting on a space V. Let $\langle \cdot, \cdot \rangle$ be an arbitrary inner product on V and define a new inner product $\langle \cdot, \cdot \rangle_G$ on V by

$$\langle v_1, v_2 \rangle_G = \int_G \langle \Pi(g)v_1, \Pi(g)v_2 \rangle \; d\mu\,(g)\,,$$

where μ is a left Haar measure. Again, it is easy to check that $\langle \cdot, \cdot \rangle_G$ is an inner product. Furthermore, if $h \in G$, then by the left-invariance of μ,

$$\begin{aligned} \langle \Pi(h)v_1, \Pi(h)v_2 \rangle_G &= \int_G \langle \Pi(g)\Pi(h)v_1, \Pi(g)\Pi(h)v_2 \rangle \; d\mu\,(g) \\ &= \int_G \langle \Pi(gh)v_1, \Pi(gh)v_2 \rangle \; d\mu\,(g) \\ &= \langle v_1, v_2 \rangle_G\,. \end{aligned}$$

So, Π is a unitary representation with respect to $\langle \cdot, \cdot \rangle_G$, and thus completely reducible. Note that the integral defining $\langle \cdot, \cdot \rangle_G$ is convergent because μ is finite. □

4.11 Exercises

1. Prove Point 2 of Proposition 4.5.
2. Suppose that Π is a finite-dimensional unitary representation of a matrix Lie group G (i.e., V is a finite-dimensional Hilbert space, and Π is a continuous homomorphism of G into $U(V)$). Let π be the associated representation of the Lie algebra $\mathfrak{g}$. Show that for each $X \in \mathfrak{g}$, $\pi(X)^* = -\pi(X)$.
3. Show that the adjoint representation and the standard representation are equivalent representations of the Lie algebra $\mathsf{so}(3)$. Show that the adjoint and standard representations of the group $\mathsf{SO}(3)$ are equivalent.
4. Define a vector space with basis $u_0, u_1, \ldots, u_m$. Now, define operators $\pi(H)$, $\pi(X)$, and $\pi(Y)$ by formula (4.10). Verify by direct computation that the operators defined by (4.10) satisfy the commutation relations $[\pi(H), \pi(X)] = 2\pi(X)$, $[\pi(H), \pi(Y)] = -2\pi(Y)$, and $[\pi(X), \pi(Y)] = \pi(H)$. (Thus, $\pi(H)$, $\pi(X)$, and $\pi(Y)$ define a representation of $\mathsf{sl}(2;\mathbb{C})$.) *Hint*: When dealing with $\pi(Y)$, treat the case of u_k, $k < m$, separately from the case of u_m.

5. Consider the standard representation of the Heisenberg group, acting on $\mathbb{C}^3$. Determine all subspaces of $\mathbb{C}^3$ which are invariant under the action of the Heisenberg group. Is this representation completely reducible?
6. Give an example of a representation of the commutative group $\mathbb{R}$ which is not completely reducible.
7. Prove Proposition 4.25.
   Hint: There is a one-to-one correspondence between subspaces of V and subspaces of V^* as follows: Given a subspace W of V, the *annihilator* of W is the subspace of all ϕ in V^* such that ϕ is zero on W. See Section B.7.
8. Consider the unitary representations $\Pi_\hbar$ of the real Heisenberg group. Assume that there is some sort of associated representation $\pi_\hbar$ of the Lie algebra, which should be given by

$$\pi_\hbar(X)f = \frac{d}{dt}\Pi_\hbar\left(e^{tX}\right)f\Big|_{t=0}.$$

   (We have not proved any theorem of this sort for infinite-dimensional unitary representations.)
   Computing in a purely formal manner (i.e., ignoring all technical issues) compute

$$\pi_\hbar\begin{pmatrix}0&1&0\\0&0&0\\0&0&0\end{pmatrix}, \pi_\hbar\begin{pmatrix}0&0&0\\0&0&1\\0&0&0\end{pmatrix}, \pi_\hbar\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}.$$

   Verify (still formally) that these operators have the right commutation relations to generate a representation of the Lie algebra of the real Heisenberg group; that is, verify that on this basis, $\pi_\hbar[X,Y] = [\pi_\hbar(X), \pi_\hbar(Y)]$. Why is this computation not rigorous?
9. Consider the Heisenberg group over the field $\mathbb{Z}/p$ of integers mod p, with p prime, namely

$$H_p = \left\{ \begin{pmatrix}1&a&b\\0&1&c\\0&0&1\end{pmatrix} \middle| a,b,c \in \mathbb{Z}/p \right\}.$$

   This is a subgroup of the group $\mathsf{GL}(3;\mathbb{Z}/p)$ and has p^3 elements.
   Let V_p denote the space of complex-valued functions on $\mathbb{Z}/p$, which is a p-dimensional complex vector space. For each nonzero $n \in \mathbb{Z}/p$, define a complex representation of H_p by the formula

$$(\Pi_n f)(x) = e^{-i2\pi nb/p}e^{i2\pi ncx/p}f(x-a), \quad x \in \mathbb{Z}/p.$$

   (These representations are analogous to the unitary representations of the real Heisenberg group, with the quantity $2\pi n/p$ playing the role of $\hbar$.) Note that these representations are defined over $\mathbb{C}$ rather than $\mathbb{Z}/p$.

(a) Show that for each n, Π_n is actually a representation of H_p and that it is irreducible.
(b) Determine (up to equivalence) all of the one-dimensional complex representations of H_p.
(c) Show that every irreducible complex representation of H_p is either one dimensional or equivalent to one of the Π_n's.

10. Prove Theorem 4.15.
Hints: For existence, choose bases $\{e_i\}$ and $\{f_j\}$ for U and V. Then, define a space W which has as a basis $\{w_{ij} \,|\, 1 \leq i \leq n, 1 \leq j \leq m\}$. Define $\phi(e_i, f_j) = w_{ij}$ and extend by bilinearity. For uniqueness, use the universal property.
11. Recall the spaces V_m introduced in Section 4.3, viewed as representations of the Lie algebra $\mathsf{sl}(2;\mathbb{C})$. In particular, consider the space V_1 (which has dimension 2).
(a) Regard $V_1 \otimes V_1$ as a representation of $\mathsf{sl}(2;\mathbb{C})$, as in Definition 4.23. Show that this representation is not irreducible.
(b) Now, view $V_1 \otimes V_1$ as a representation of $\mathsf{sl}(2;\mathbb{C}) \oplus \mathsf{sl}(2;\mathbb{C})$, as in Definition 4.20. Show that this representation is irreducible.
12. *Proof of Lemma 4.30.*
Let $\{E_1, E_2, E_3\}$ be the usual basis for $\mathsf{su}(2)$ and let $\{F_1, F_2, F_3\}$ be the basis for $\mathsf{so}(3)$ introduced in Section 4.9. Identify $\mathsf{su}(2)$ with $\mathbb{R}^3$ by identifying the basis $\{E_1, E_2, E_3\}$ with the standard basis for $\mathbb{R}^3$. Consider ad_{E_1}, ad_{E_2}, and ad_{E_3} as operators on $\mathsf{su}(2)$, hence on $\mathbb{R}^3$. Show that $\mathrm{ad}_{E_i} = F_i$, for $i = 1, 2, 3$. It follows that ad is a Lie algebra isomorphism of $\mathsf{su}(2)$ onto $\mathsf{so}(3)$.
Now, consider $\mathrm{Ad} : \mathsf{SU}(2) \to \mathsf{GL}(\mathsf{su}(2)) = \mathsf{GL}(3;\mathbb{R})$. Show that the image of Ad is precisely $\mathsf{SO}(3)$. Show that the kernel of Ad is $\{I, -I\}$.
Show that $\mathrm{Ad} : \mathsf{SU}(2) \to \mathsf{SO}(3)$ is the homomorphism Φ required by Lemma 4.30.
13. Suppose V is a finite-dimensional representation of a group or Lie algebra and that W is a nontrivial invariant subspace of V. Show that there exists a nontrivial *irreducible* invariant subspace for V that is contained in W.
14. Suppose V is a finite-dimensional representation of a group or Lie algebra and that W and W' are invariant subspaces of V with $W' \subset W$. Suppose that U is an invariant subspace for V such that $V = W' \oplus U$. Show that $W \cap U$ is an invariant subspace of V and that $W = W' \oplus (W \cap U)$.
15. Suppose that V_1 and V_2 are *inequivalent* irreducible representations of a group or Lie algebra, and consider the associated representation $V_1 \oplus V_2$. Regard V_1 and V_2 as subpaces of $V_1 \oplus V_2$ in the obvious way. Following the outline below, show that V_1 and V_2 are the only nontrivial invariant subspaces of $V_1 \oplus V_2$.
(a) First assume that U is a nontrivial *irreducible* invariant subspace. Let $P_1 : V_1 \oplus V_2 \to V_1$ be the projection onto the first factor and let P_2 be the projection onto the second factor. Show that P_1 and P_2 are intertwining maps. Show that $U = V_1$ or $U = V_2$.

(b) Using Exercise 13, show that V_1 and V_2 are the only nontrivial invariant subspaces of $V_1 \oplus V_2$.

16. Suppose that V is an irreducible finite-dimensional representation of a group or Lie algebra, and consider the associated representation $V \oplus V$. Show that every nontrivial invariant subspace U of $V \oplus V$ is equivalent to V and is of the form

$$U = \{(\lambda_1 v, \lambda_2 v) | v \in V\},$$

for some constants λ_1 and λ_2, not both zero.

17. Suppose V is a completely reducible finite-dimensional representation of some group or Lie algebra, in which case V is equivalent to a representation of the form

$$(V_1 \oplus \cdots \oplus V_1) \oplus (V_2 \oplus \cdots \oplus V_2) \oplus \cdots \oplus (V_k \oplus \cdots \oplus V_k),$$

where $V_1, \ldots, V_k$ are pairwise inequivalent irreducible representations and where V_l occurs n_l times, $l = 1, \ldots, k$. Suppose U is a nontrivial irreducible invariant subspace of V. Show that U is contained in $V_l \oplus \cdots \oplus V_l$ for some l and that, as a subspace of $V_l \oplus \cdots \oplus V_l$, U is of the form

$$\{(\lambda_1 v, \lambda_2 v, \ldots, \lambda_{n_l} v) | v \in V_l\},$$

where the λ's are constants that are not all equal to zero.

18. Using the results and methods of the five preceding exercises, prove Proposition 4.33.

Part II

Semisimple Theory

5

The Representations of SU(3)

5.1 Introduction

There is a theory of the representations of semisimple groups and Lie algebras (discussed in Chapters 6, 7, and 8) that includes as a special case the representation theory of SU(3). However, I feel that it is worthwhile to examine the case of SU(3) separately, before going on to the general theory. I feel this way partly because SU(3) is an important group in physics, but chiefly because the general semisimple theory is difficult to digest. Considering a nontrivial example makes what is going on much clearer. In fact, all of the elements of the general theory are present already in the case of SU(3), so we do not lose too much by considering at first just this case.

The main result of this chapter is Theorem 5.9, which states that an irreducible finite-dimensional representation of SU(3) can be classified in terms of its "highest weight." This is analogous to labeling the irreducible representations V_m of SU(2) or $\mathsf{sl}(2;\mathbb{C})$ by the highest eigenvalue of $\pi_m(H)$. (The highest eigenvalue of $\pi_m(H)$ in V_m is precisely m.) In the next two chapters, we will look at the analogous results for general semisimple Lie algebras.

The group SU(3) is simply connected (Appendix E), and so the finite-dimensional representations of SU(3) are in one-to-one correspondence with the finite-dimensional representations of the Lie algebra su(3). Meanwhile, the complex representations of su(3) are in one-to-one correspondence with the complex-linear representations of the complexified Lie algebra $\mathsf{su}(3)_{\mathbb{C}} = \mathsf{sl}(3;\mathbb{C})$ (Proposition 4.6). Moreover, a representation of SU(3) is irreducible if and only if the associated representation of su(3) is irreducible, and this holds if and only if the associated complex-linear representation of $\mathsf{sl}(3;\mathbb{C})$ is irreducible. (This follows from Proposition 4.5, Proposition 4.6, and the connectedness of SU(3).) Thus, we have the following result.

Proposition 5.1. *There is a one-to-one correspondence between the finite-dimensional complex representations* Π *of* SU(3) *and the finite-dimensional*

complex-linear representations π *of* $\mathsf{sl}(3;\mathbb{C})$. *This correspondence is determined by the property that*

$$\Pi\left(e^X\right) = e^{\pi(X)}$$

for all $X \in \mathsf{su}(3) \subset \mathsf{sl}(3;\mathbb{C})$.

The representation Π *is irreducible if and only if the representation* π *is irreducible.*

Since $\mathsf{SU}(3)$ is compact, Proposition 4.36 tells us that all of the finite-dimensional representations of $\mathsf{SU}(3)$ are direct sums of irreducible representations. The above proposition then implies that the same holds for $\mathsf{sl}(3;\mathbb{C})$; that is, $\mathsf{sl}(3;\mathbb{C})$ has the complete reducibility property. Complete reducibility will be an essential ingredient even in the classification of irreducible representations. (See the proof of Proposition 5.16.)

Moreover, we can apply the same reasoning to the simply-connected group $\mathsf{SU}(2)$, its Lie algebra $\mathsf{su}(2)$, and its complexified Lie algebra $\mathsf{sl}(2;\mathbb{C})$. Thus, we have established the following.

Proposition 5.2. *Every finite-dimensional representation of* $\mathsf{sl}(2;\mathbb{C})$ *or* $\mathsf{sl}(3;\mathbb{C}$ *decomposes as a direct sum of irreducible invariant subspaces.*

We will use the following basis for $\mathsf{sl}(3;\mathbb{C})$:

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

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

$$Y_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad Y_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Note that the span of $\{H_1, X_1, Y_1\}$ is a subalgebra of $\mathsf{sl}(3;\mathbb{C})$ which is isomorphic to $\mathsf{sl}(2;\mathbb{C})$ (as can be seen by ignoring the third row and the third column in each matrix). Similarly, the span of $\{H_2, X_2, Y_2\}$ is a subalgebra isomorphic to $\mathsf{sl}(2;\mathbb{C})$. Thus, we have the following commutation relations:

$$\begin{aligned} [H_1, X_1] &= 2X_1, \quad [H_2, X_2] = 2X_2, \\ [H_1, Y_1] &= -2Y_1, \quad [H_2, Y_2] = -2Y_2, \\ [X_1, Y_1] &= H_1, \quad [X_2, Y_2] = H_2. \end{aligned}$$

We now list all of the commutation relations among the basis elements which involve at least one of H_1 and H_2. (This includes some repetitions of the above commutation relations.)

$$
\begin{aligned}
&[H_1, H_2] = \quad 0;\\
&\\
&[H_1, X_1] = \ 2X_1, \ [H_1, Y_1] = -2Y_1,\\
&[H_2, X_1] = -X_1, \ [H_2, Y_1] = \quad Y_1;\\
&\\
&[H_1, X_2] = -X_2, \ [H_1, Y_2] = \quad Y_2,\\
&[H_2, X_2] = \ 2X_2, \ [H_2, Y_2] = -2Y_2;\\
&\\
&[H_1, X_3] = \quad X_3, \ [H_1, Y_3] = \ -Y_3,\\
&[H_2, X_3] = \quad X_3, \ [H_2, Y_3] = \ -Y_3.
\end{aligned}
\tag{5.1}
$$

Finally, we list all of the remaining commutation relations.

$$
\begin{aligned}
&[X_1, Y_1] = \quad H_1,\\
&[X_2, Y_2] = \quad H_2,\\
&[X_3, Y_3] = H_1 + H_2;\\
&\\
&[X_1, X_2] = \ X_3, \quad [Y_1, Y_2] = -Y_3,\\
&[X_1, Y_2] = \ 0, \quad [X_2, Y_1] = \ 0;\\
&\\
&[X_1, X_3] = \ 0, \quad [Y_1, Y_3] = \ 0,\\
&[X_2, X_3] = \ 0, \quad [Y_2, Y_3] = \ 0;\\
&\\
&[X_2, Y_3] = \ Y_1, \quad [X_3, Y_2] = \ X_1,\\
&[X_1, Y_3] = -Y_2, \quad [X_3, Y_1] = -X_2.
\end{aligned}
$$

All of the analysis we will do for the representations of $\mathsf{sl}(3;\mathbb{C})$ will be in terms of the above basis. From now on, all representations of $\mathsf{sl}(3;\mathbb{C})$ will be assumed to be finite dimensional and complex linear.

5.2 Weights and Roots

Our basic strategy in classifying the representations of $\mathsf{sl}(3;\mathbb{C})$ is to simultaneously diagonalize $\pi(H_1)$ and $\pi(H_2)$. (See Section B.8 for information on simultaneous diagonalization.) Since H_1 and H_2 commute, $\pi(H_1)$ and $\pi(H_2)$ will also commute (for any representation π) and so there is at least a chance that $\pi(H_1)$ and $\pi(H_2)$ can be simultaneously diagonalized. (Compare Proposition B.13.)

Definition 5.3. *If (π, V) is a representation of $\mathsf{sl}(3;\mathbb{C})$, then an ordered pair $\mu = (m_1, m_2) \in \mathbb{C}^2$ is called a* ***weight*** *for π if there exists $v \neq 0$ in V such that*

$$
\begin{aligned}
\pi(H_1)v &= m_1 v,\\
\pi(H_2)v &= m_2 v.
\end{aligned}
\tag{5.2}
$$

A nonzero vector v satisfying (5.2) is called a ***weight vector*** *corresponding to the weight μ. If $\mu = (m_1, m_2)$ is a weight, then the space of all vectors v satisfying (5.2) (including the zero vector) is the* ***weight space*** *corresponding to the weight μ. The* ***multiplicity*** *of a weight is the dimension of the corresponding weight space.*

Thus, a weight is simply a pair of simultaneous eigenvalues for $\pi(H_1)$ and $\pi(H_2)$. (See Section B.8 for a discussion of simultaneous eigenvectors and eigenvalues.) It is easily shown that equivalent representations have the same weights and multiplicities.

Proposition 5.4. *Every representation of $\mathsf{sl}(3;\mathbb{C})$ has at least one weight.*

Proof. Since we are working over the complex numbers, $\pi(H_1)$ has at least one eigenvalue $m_1 \in \mathbb{C}$. Let $W \subset V$ be the eigenspace for $\pi(H_1)$ with eigenvalue m_1. Since $[H_1, H_2] = 0$, $\pi(H_2)$ commutes with $\pi(H_1)$, and, so, by Proposition B.4, $\pi(H_2)$ must map W into itself. Thus, $\pi(H_2)$ can be viewed as an operator on W. Then, the restriction of $\pi(H_2)$ to W must have at least one eigenvector w with eigenvalue $m_2 \in \mathbb{C}$ and w is a simultaneous eigenvector for $\pi(H_1)$ and $\pi(H_2)$ with eigenvalues m_1 and m_2. □

Now, every representation π of $\mathsf{sl}(3;\mathbb{C})$ can be viewed, by restriction, as a representation of the subalgebra $\{H_1, X_1, Y_1\} \cong \mathsf{sl}(2;\mathbb{C})$. Note that even if π is irreducible as a representation of $\mathsf{sl}(3;\mathbb{C})$, there is no reason to expect that it will still be irreducible as a representation of the subalgebra $\{H_1, X_1, Y_1\}$. Nevertheless, π restricted to $\{H_1, X_1, Y_1\}$ must be *some* finite-dimensional representation of $\mathsf{sl}(2;\mathbb{C})$. The same reasoning applies to the restriction of π to the subalgebra $\{H_2, X_2, Y_2\}$, which is also isomorphic to $\mathsf{sl}(2;\mathbb{C})$.

Now, recall Theorem 4.12, which tells us that in any finite-dimensional representation of $\mathsf{sl}(2;\mathbb{C})$, irreducible or not, all of the eigenvalues of $\pi(H)$ must be integers. Theorem 4.12 has the following corollary.

Corollary 5.5. *If π is a representation of $\mathsf{sl}(3;\mathbb{C})$, then all of the weights of π are of the form*

$$\mu = (m_1, m_2)$$

with m_1 and m_2 being integers.

Proof. Apply Theorem 4.12 to the restriction of π to $\{H_1, X_1, Y_1\}$ and to the restriction of π to $\{H_2, X_2, Y_2\}$. □

Our strategy now is to begin with one simultaneous eigenvector for $\pi(H_1)$ and $\pi(H_2)$ and then to apply $\pi(X_i)$ or $\pi(Y_i)$ and see what the effect is. The following definition is relevant in this context.

Definition 5.6. *An ordered pair $\alpha = (a_1, a_2) \in \mathbb{C}^2$ is called a* ***root*** *if*

1. *a_1 and a_2 are not both zero, and*

2. there exists a nonzero $Z \in \mathsf{sl}(3;\mathbb{C})$ such that

$$[H_1, Z] = a_1 Z,$$
$$[H_2, Z] = a_2 Z.$$

The element Z is called a ***root vector*** *corresponding to the root α.*

Condition 2 in the definition says that Z is a simultaneous eigenvector for ad_{H_1} and ad_{H_2}. This means that Z is a weight vector for the adjoint representation with weight (a_1, a_2). Thus, taking into account Condition 1, we may say that the roots are precisely the *nonzero weights of the adjoint representation.* Corollary 5.5 then tells us that for any root, both a_1 and a_2 must be integers, which we can also see directly in (5.3). The commutation relations (5.1) tell us what the roots for $\mathsf{sl}(3;\mathbb{C})$ are. There are six roots:

$$\begin{array}{cc} \alpha & \mathbf{Z} \\ (2,-1) & X_1 \\ (-1,2) & X_2 \\ (1,1) & X_3 \\ (-2,1) & Y_1 \\ (1,-2) & Y_2 \\ (-1,-1) & Y_3 \end{array} \tag{5.3}$$

Note that H_1 and H_2 are also simultaneous eigenvectors for ad_{H_1} and ad_{H_2}, but they are not root vectors because the simultaneous eigenvalues are both zero. Since the vectors in (5.3) together with H_1 and H_2 form a basis for $\mathsf{sl}(3;\mathbb{C})$, it is not hard to show that the roots listed in (5.3) are the only roots (Exercise 1). These six roots form a "root system," conventionally called A_2. For more information, see Chapters 6, 7, and 8.

It is convenient to single out the two roots corresponding to X_1 and X_2 and give them special names:

$$\alpha_1 = (2,-1),$$
$$\alpha_2 = (-1,2). \tag{5.4}$$

The roots α_1 and α_2 are called the **positive simple roots**. They have the property that all of the roots can be expressed as linear combinations of α_1 and α_2 with *integer* coefficients, and these coefficients are (for each root) either all greater than or equal to zero or all less than or equal to zero. This is verified by direct computation:

$$\begin{aligned} (2,-1) &= \alpha_1, \\ (-1,2) &= \alpha_2, \\ (1,1) &= \alpha_1 + \alpha_2, \\ (-2,1) &= -\alpha_1, \\ (1,-2) &= -\alpha_2, \\ (-1,-1) &= -\alpha_1 - \alpha_2. \end{aligned}$$

(The decision to designate α_1 and α_2 as the positive simple roots is arbitrary; any other pair of roots with similar properties would do just as well. We simply choose one set of positive simple roots and hold to that choice throughout the chapter. See Section 6.8.)

The significance of the roots for the representation theory of $\mathsf{sl}(3;\mathbb{C})$ is contained in the following lemma. Although its proof is very easy, this lemma plays a crucial role in the classification of the representations of $\mathsf{sl}(3;\mathbb{C})$. Note that this lemma is the analog of Lemma 4.10, which was the key to the classification of the representations of $\mathsf{sl}(2;\mathbb{C})$.

Lemma 5.7. *Let $\alpha = (a_1, a_2)$ be a root and Z_α a corresponding root vector in $\mathsf{sl}(3;\mathbb{C})$. Let π be a representation of $\mathsf{sl}(3;\mathbb{C})$, $\mu = (m_1, m_2)$ a weight for π, and $v \neq 0$ a corresponding weight vector. Then,*

$$\pi(H_1)\pi(Z_\alpha)v = (m_1 + a_1)\pi(Z_\alpha)v,$$
$$\pi(H_2)\pi(Z_\alpha)v = (m_2 + a_2)\pi(Z_\alpha)v.$$

Thus, either $\pi(Z_\alpha)v = 0$ or $\pi(Z_\alpha)v$ is a new weight vector with weight

$$\mu + \alpha = (m_1 + a_1, m_2 + a_2).$$

Proof. The definition of a root tells us that we have the commutation relation $[H_1, Z_\alpha] = a_1 Z_\alpha$. Thus,

$$\begin{aligned}\pi(H_1)\pi(Z_\alpha)v &= (\pi(Z_\alpha)\pi(H_1) + a_1\pi(Z_a))\, v \\ &= \pi(Z_\alpha)(m_1 v) + a_1\pi(Z_\alpha)v \\ &= (m_1 + a_1)\pi(Z_\alpha)v.\end{aligned}$$

A similar argument allows us to compute $\pi(H_2)\pi(Z_\alpha)v$. □

5.3 The Theorem of the Highest Weight

We see then that if we have a representation with a weight $\mu = (m_1, m_2)$, then by applying the root vectors X_1, X_2, X_3, Y_1, Y_2, and Y_3, we can get some new weights of the form $\mu + \alpha$, where α is the root. Of course, some of the time, $\pi(Z_\alpha)v$ will be zero, in which case $\mu + \alpha$ is not necessarily a weight. In fact, since our representation is finite dimensional, there can be only finitely many weights, so we must get zero quite often. By analogy to the classification of the representations of $\mathsf{sl}(2;\mathbb{C})$, we would like to single out in each representation a "highest" weight and then work from there. The following definition gives the "right" notion of highest.

Definition 5.8. *Let $\alpha_1 = (2, -1)$ and $\alpha_2 = (-1, 2)$ be the roots introduced in (5.4). Let μ_1 and μ_2 be two weights. Then, μ_1 is **higher** than μ_2 (or, equivalently, μ_2 is **lower** than μ_1) if $\mu_1 - \mu_2$ can be written in the form*

$$\mu_1 - \mu_2 = a\alpha_1 + b\alpha_2$$

with $a \geq 0$ and $b \geq 0$. This relationship is written as $\mu_1 \succeq \mu_2$ or $\mu_2 \preceq \mu_1$.

If π is a representation of $\mathsf{sl}(3;\mathbb{C})$, then a weight μ_0 for π is said to be a ***highest weight*** *if for all weights μ of π, $\mu \preceq \mu_0$.*

Note that the relation of "higher" is only a *partial* ordering; that is, one can easily have μ_1 and μ_2 such that μ_1 is neither higher nor lower than μ_2. For example, $\alpha_1 - \alpha_2$ is neither higher nor lower than 0. This, in particular, means that a finite set of weights need not have a highest element (e.g., the set $\{0, \alpha_1 - \alpha_2\}$ has no highest element). Note also that the coefficients a and b do not have to be integers, even if both μ_1 and μ_2 have integer entries. For example, $(1,0)$ is higher than $(0,0)$ since $(1,0) = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$.

We are now ready to state the main theorem regarding the irreducible representations of $\mathsf{sl}(3;\mathbb{C})$, commonly called the theorem of the highest weight.

Theorem 5.9.

1. *Every irreducible representation π of $\mathsf{sl}(3;\mathbb{C})$ is the direct sum of its weight spaces; that is, $\pi(H_1)$ and $\pi(H_2)$ are simultaneously diagonalizable in every irreducible representation.*
2. *Every irreducible representation of $\mathsf{sl}(3;\mathbb{C})$ has a unique highest weight μ_0, and two equivalent irreducible representations have the same highest weight.*
3. *Two irreducible representations of $\mathsf{sl}(3;\mathbb{C})$ with the same highest weight are equivalent.*
4. *If π is an irreducible representation of $\mathsf{sl}(3;\mathbb{C})$, then the highest weight μ_0 of π is of the form*
$$\mu_0 = (m_1, m_2)$$
with m_1 and m_2 being non-negative integers.
5. *If m_1 and m_2 are non-negative integers, then there exists an irreducible representation π of $\mathsf{sl}(3;\mathbb{C})$ with highest weight $\mu_0 = (m_1, m_2)$.*

An ordered pair (m_1, m_2) with m_1 and m_2 being non-negative integers is called a **dominant integral element**. Theorem 5.9 tell us that the highest weight of each irreducible representation of $\mathsf{sl}(3;\mathbb{C})$ is a dominant integral element and, conversely, that every dominant integral element occurs as the highest weight of some irreducible representation. Since $(1,0) = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$ and $(0,1) = \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2$, we see that every dominant integral element is higher than zero. However, if μ has integer coefficients and is higher than zero, this does not necessarily mean that μ is dominant integral. (For example, $\alpha_1 = (2,-1)$ is higher than zero but is not dominant integral.)

Figure 5.1 shows the roots and dominant integral elements for $\mathsf{sl}(3;\mathbb{C})$. This picture is made using the obvious basis for the space of weights; that is, the x-coordinate is the eigenvalue of H_1 and the y-coordinate is the eigenvalue of H_2. Once we have introduced the Weyl group (Section 5.6), we will see

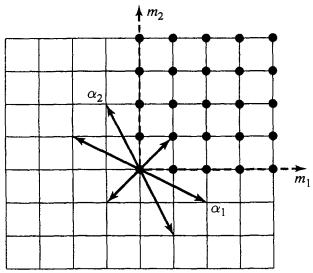


Fig. 5.1. Roots and dominant integral elements for $\mathsf{sl}(3;\mathbb{C})$ (in obvious basis)

the same picture (Figure 5.2) rendered using a Weyl-invariant inner product, which will give a more geometric view of the situation.

Note the parallels between this result and the classification of the irreducible representations of $\mathsf{sl}(2;\mathbb{C})$: In each irreducible representation of $\mathsf{sl}(2;\mathbb{C})$, $\pi(H)$ is diagonalizable, and there is a largest eigenvalue of $\pi(H)$. Two irreducible representations of $\mathsf{sl}(2;\mathbb{C})$ with the same largest eigenvalue are equivalent. The highest eigenvalue is always a non-negative integer, and, conversely, for every non-negative integer m, there is an irreducible representation with highest eigenvalue m.

Note, however, that in the classification of the representations of $\mathsf{sl}(3;\mathbb{C})$, the notion of "highest" does not mean what we might have thought it should mean; that is, $(m_1, m_2) \succeq (n_1, n_2)$ does *not* mean $m_1 \geq n_1$ and $m_2 \geq n_2$, as we might have guessed. (For example, the weight $(1,1)$ is higher than the weights $(-1,2)$ and $(2,-1)$.) Nevertheless, the condition on which weights can be highest weights *is* the obvious one: m_1 and m_2 must be non-negative integers.

It is possible to obtain much more information about the irreducible representations besides the highest weight. For example, we have the following formula for the dimension of the representation with highest weight (m_1, m_2).

Theorem 5.10. *The dimension of the irreducible representation with highest weight* (m_1, m_2) *is*

$$\frac{1}{2}(m_1+1)(m_2+1)(m_1+m_2+2).$$

See Humphreys (1972, Section 24.3). (Humphreys refers to $\mathsf{sl}(3;\mathbb{C})$ as A_2.) We will not prove this formula here. It is a consequence of the Weyl character formula, which is discussed in the context of general semisimple Lie algebras in Sections 7.4 and 7.6.

5.4 Proof of the Theorem

It will take us some time to prove Theorem 5.9. The proof will consist of a series of propositions.

Proposition 5.11. *In every irreducible representation (π, V) of $\mathsf{sl}(3;\mathbb{C})$, $\pi(H_1)$ and $\pi(H_2)$ can be simultaneously diagonalized; that is, V is the direct sum of its weight spaces.*

Proof. Let W be the direct sum of the weight spaces in V. Equivalently, W is the space of all vectors $w \in V$ such that w can be written as a linear combination of simultaneous eigenvectors for $\pi(H_1)$ and $\pi(H_2)$. Since (Proposition 5.4) π always has at least one weight, $W \neq \{0\}$.

On the other hand, Lemma 5.7 tells us that if Z_α is a root vector corresponding to the root α, then $\pi(Z_\alpha)$ maps the weight space corresponding to μ into the weight space corresponding to $\mu + \alpha$. Thus, W is invariant under the action of all of the root vectors, namely under the action X_1, X_2, X_3, Y_1, Y_2, and Y_3. Since W is certainly invariant under the action of H_1 and H_2, W is invariant under all of $\mathsf{sl}(3;\mathbb{C})$. Thus, by irreducibility, $W = V$. □

Definition 5.12. *A representation (π, V) of $\mathsf{sl}(3;\mathbb{C})$ is said to be a* ***highest weight cyclic representation with weight*** *$\mu_0 = (m_1, m_2)$ if there exists $v \neq 0$ in V such that*

1. *v is a weight vector with weight μ_0,*
2. *$\pi(X_1)v = \pi(X_2)v = 0$,*
3. *the smallest invariant subspace of V containing v is all of V.*

The vector v is called a ***cyclic vector*** *for π.*

Proposition 5.13. *Let (π, V) be a highest weight cyclic representation of $\mathsf{sl}(3;\mathbb{C})$ with weight μ_0. Then,*

1. *π has highest weight μ_0 and*
2. *the weight space corresponding to the highest weight μ_0 is one dimensional.*

Before turning to the proof of this proposition, let us record a simple lemma that applies to arbitrary Lie algebras and which will be useful also in the setting of general semisimple Lie algebras.

Lemma 5.14. *Suppose that $\mathfrak{g}$ is any Lie algebra and that π is a representation of $\mathfrak{g}$. Suppose that $X_1, \ldots, X_m$ is an ordered basis for $\mathfrak{g}$ as a vector space. Then, any expression of the form*

$$\pi(X_{i_1})\pi(X_{i_2})\cdots\pi(X_{i_N}), \tag{5.5}$$

can be expressed as a linear combination of terms of the form

$$\pi(X_m)^{k_m}\pi(X_{m-1})^{k_{m-1}}\cdots\pi(X_1)^{k_1} \tag{5.6}$$

where in each term $k_1 + \cdots + k_m \leq N$. Here, the k_i's are non-negative integers (zero is allowed!).

Proof. Let us think about how this works in the case $N \leq 2$. If $N = 1$, there is nothing to do: Any expression of the form $\pi(X_i)$ is of the form (5.6) with $k_i = 1$ and all the other k_j's equal to zero. If $N = 2$, we consider an expression of the form $\pi(X_i)\pi(X_j)$. If $i \geq j$, then this is already of the form (5.6) (with most of the k_i's equal to zero). If $i < j$, then we write

$$\begin{aligned}\pi(X_i)\pi(X_j) &= \pi(X_j)\pi(X_i) + \pi([X_i, X_j]) \\ &= \pi(X_j)\pi(X_i) + \sum_{k=1}^{m} c_{ijk}\pi(X_k),\end{aligned} \tag{5.7}$$

where the c_{ijk}'s are the structure constants for this basis of $\mathfrak{g}$, and the right-hand side is now a linear combination of terms of the form (5.6).

The proof for the general case is by induction on N. Assume, then, that the result holds for a product of N or fewer terms and consider an expression of the form (5.5) with $N+1$ factors. By our induction hypothesis, we can assume that the last N factors are in the desired form and we need only consider an expression of the form

$$\pi(X_i)\pi(X_m)^{k_m}\pi(X_{m-1})^{k_{m-1}}\cdots\pi(X_1)^{k_1}$$

with $k_1 + \cdots + k_m \leq N$. Now, we move the factor of $\pi(X_i)$ to the right one step at a time until it is in the right spot. Each time we have $\pi(X_i)\pi(X_j)$ somewhere in the expression we can move the $\pi(X_i)$ to the right by using (5.7). As we move $\pi(X_i)$ to the right, we will generate multiple commutator terms, each of which has one fewer factor and, thus, can be handled by the induction hypothesis. Thus, we ultimately get several terms with $N-1$ factors, together with one term having N factors and being of the form (5.6) (once $\pi(X_i)$ finally gets to the right spot). □

We now proceed with the proof of Proposition 5.13.

Proof. Let v be as in the definition. Consider the subspace W of V spanned by elements of the form

$$w = \pi(Y_{i_1})\pi(Y_{i_2})\cdots\pi(Y_{i_n})v \tag{5.8}$$

with each i_l equal to 1, 2, or 3 and $n \geq 0$. (If $n = 0$, it is understood that w in (5.8) is equal to v.) I assert that W is invariant. To see this, it suffices to check that W is invariant under each of the basis elements, which we do by using the lemma. We take as our basis for $\mathfrak{sl}(3;\mathbb{C})$ the elements X_1, X_2, X_3, H_1, H_2, Y_1, Y_2, and Y_3, in that order. If we multiply an element w by (π applied to) some Lie algebra element, the lemma tells us that we can rewrite the resulting vector as a linear combination of terms in which the $\pi(X_i)$'s act first, the $\pi(H_i)$'s act second, and the $\pi(Y_i)$'s act last, and all of these are applied to the vector v. However, v is annihilated by the $\pi(X_i)$'s, so any term having a positive power of any X_i is simply zero and we are left with

the $\pi(H_i)$'s and the $\pi(Y_i)$'s acting on v (in that order). Furthermore, v is an eigenvector for $\pi(H_1)$ and $\pi(H_2)$, so any factors of $\pi(H_i)$ acting on v can be replaced by constants in front of the whole expression. That leaves only factors of $\pi(Y_i)$ applied to v, which means that we are getting a linear combination of vectors of the form (5.8). This shows that W is invariant. Since by definition W contains v, we must have $W = V$.

Now, Y_1 is a root vector with root $-\alpha_1$, Y_2 is a root vector with root $-\alpha_2$, and Y_3 is a root vector with root $-\alpha_1 - \alpha_2$. So, applying Lemma 5.7 repeatedly, we see that each element of the form (5.8) is either zero or a weight vector with weight $\mu_0 - n_1\alpha_1 - n_2\alpha_2$. Thus, $V = W$ is spanned by v together with weight vectors with weights lower than μ_0. Thus, μ_0 is the highest weight for V.

Furthermore, every element w of W can be written as a $w = cv + v_1 + \cdots + v_m$, where $v_1, \ldots, v_m$ are weight vectors with distinct weights lower than μ_0. (If at first the weights are not distinct, then simply combine all the terms corresponding to each distinct weight.) It then follows from Proposition B.14 that such a vector can be a weight vector with weight μ_0 only if all the v_k's $(k = 1, \ldots, m)$ are zero, in which case, w is a multiple of v. (Apply the proposition to $(cv - w) + v_1 + \cdots + v_m$.) Thus, the weight space corresponding to μ_0 is spanned by v and the corresponding weight space is one dimensional. □

Proposition 5.15. *Every irreducible representation of* $\mathsf{sl}(3;\mathbb{C})$ *is a highest weight cyclic representation, with a unique highest weight* μ_0.

Proof. Uniqueness is immediate, since by the previous proposition, μ_0 is the highest weight, and two distinct weights cannot both be highest.

We have already shown that every irreducible representation is the direct sum of its weight spaces. Since the representation is finite dimensional, there can be only finitely many weights. It follows that there must exist a weight μ_0 such that there is no weight $\mu \neq \mu_0$ with $\mu \succeq \mu_0$. This says that there is no weight higher than μ_0 (which is *not* the same as saying that μ_0 is highest). However, if there is no weight higher than μ_0, then for any nonzero weight vector v with weight μ_0, we must have

$$\pi(X_1)v = \pi(X_2)v = 0.$$

(For otherwise, say, $\pi(X_1)v$ will be a weight vector with weight $\mu_0 + \alpha_1 \succ \mu_0$.)

Since π is assumed irreducible, the smallest invariant subspace containing v must be the whole space; therefore, the representation is highest weight cyclic. □

Proposition 5.16. *Every highest weight cyclic representation of* $\mathsf{sl}(3;\mathbb{C})$ *is irreducible.*

Proof. Let (π, V) be a highest weight cyclic representation with highest weight μ_0 and cyclic vector v. By complete reducibility (Proposition 5.2), V decomposes as a direct sum of irreducible representations

$$V \cong \bigoplus_i V_i. \tag{5.9}$$

By Proposition 5.11, each of the V_i's is the direct sum of its weight spaces. Since the weight μ_0 occurs in V, it must occur in some V_i. (This follows from Proposition B.14.) On the other hand, Proposition 5.13 says that the weight space corresponding to μ_0 is one dimensional; that is, v is (up to a constant) the *only* vector in V with weight μ_0. Thus, V_i must contain v. However, then V_i is an invariant subspace containing v, so $V_i = V$. Thus, there is only one term in the sum (5.9), and V is irreducible. □

Proposition 5.17. *Two irreducible representations of* $\mathsf{sl}(3;\mathbb{C})$ *with the same highest weight are equivalent.*

Proof. We now know that a representation is irreducible if and only if it is highest weight cyclic. Suppose that (π, V) and (σ, W) are two such representations with the same highest weight μ_0. Let v and w be the cyclic vectors for V and W, respectively. Now, consider the representation $V \oplus W$ and let U be smallest invariant subspace of $V \oplus W$ which contains the vector (v, w).

By definition, U is a highest weight cyclic representation, therefore irreducible by Proposition 5.16. Consider the two "projection" maps $P_1 : V \oplus W \to V$, $P_1(v,w) = v$ and $P_2 : V \oplus W \to W$, $P_1(v,w) = w$. It is easy to check that P_1 and P_2 are intertwining maps of representations. Therefore, the restrictions of P_1 and P_2 to $U \subset V \oplus W$ will also be intertwining maps.

Now, neither $P_1|_U$ nor $P_2|_U$ is the zero map (since both are nonzero on (v, w)). Moreover, U, V, and W are all irreducible. Therefore, by Schur's Lemma, $P_1|_U$ is an isomorphism of U with V, and $P_2|_U$ is an isomorphism of U with W. Thus, $V \cong U \cong W$. □

Proposition 5.18. *If* π *is an irreducible representation of* $\mathsf{sl}(3;\mathbb{C})$*, then the highest weight of* π *is of the form*

$$\mu = (m_1, m_2)$$

with m_1 *and* m_2 *being non-negative integers.*

Proof. We already know that *all* of the weights of π are of the form (m_1, m_2), with m_1 and m_2 being integers. We must show that if $\mu_0 = (m_1, m_2)$ is the highest weight, then m_1 and m_2 are both non-negative. For this, we again use what we know about the representations of $\mathsf{sl}(2;\mathbb{C})$. If π is an irreducible representation of $\mathsf{sl}(3;\mathbb{C})$ with highest weight $\mu_0 = (m_1, m_2)$ and if $v \neq 0$ is a weight vector with weight μ_0, then we must have $\pi(X_1)v = \pi(X_2)v = 0$. (Otherwise, μ_0 would not be highest.) Theorem 4.12, applied to the restrictions of π to $\{H_1, X_1, Y_1\}$ and to $\{H_2, X_2, Y_2\}$, shows that m_1 and m_2 must be non-negative. □

Proposition 5.19. *If m_1 and m_2 are non-negative integers, then there exists an irreducible representation of $\mathsf{sl}(3;\mathbb{C})$ with highest weight $\mu = (m_1, m_2)$.*

Proof. Note that the trivial representation is an irreducible representation with highest weight $(0,0)$. So, we need only construct representations with at least one of m_1 and m_2 positive.

First, we construct two irreducible representations with highest weights $(1,0)$ and $(0,1)$. (These are the so-called **fundamental representations**.) We consider first the standard representation of $\mathsf{sl}(3;\mathbb{C})$, acting on $\mathbb{C}^3$ in the obvious way. This representation is easily shown to be irreducible. The simultaneous eigenvectors for H_1 and H_2 in the standard representation are the standard basis elements e_1, e_2, and e_3, which have weights $(1,0)$, $(-1,1)$, and $(0,-1)$, respectively. The highest weight for the standard representation is $(1,0)$.

To construct an irreducible representation with weight $(0,1)$, we modify the standard representation. Specifically, we define

$$\pi(Z) = -Z^{tr} \tag{5.10}$$

for all $Z \in \mathsf{sl}(3;\mathbb{C})$. Using the fact that $(AB)^{tr} = B^{tr}A^{tr}$, it is easy to check that

$$-[Z_1, Z_2]^{tr} = \left[-Z_1^{tr}, -Z_2^{tr}\right],$$

so that π is really a representation. (This is isomorphic to the dual of the standard representation, as defined in Section 4.7.) It is also easily checked that this representation is irreducible. The simultaneous eigenvectors for H_1 and H_2 in this representation are again e_1, e_2, and e_3, but this time with weights $(-1,0)$, $(1,-1)$, and $(0,1)$. The highest weight for this representation is $(0,1)$.

Let (π_1, V_1) denote $\mathbb{C}^3$ acted on by the standard representation and let v_1 denote a weight vector corresponding to the highest weight $(1,0)$. (So, $v_1 = (1,0,0)$.) Let (π_2, V_2) denote $\mathbb{C}^3$ acted on by the representation (5.10) and let v_2 denote a weight vector for the highest weight $(0,1)$. (So, $v_2 = (0,0,1)$.) Now, consider the representation

$$V_1 \otimes V_1 \otimes \cdots \otimes V_1 \otimes V_2 \otimes V_2 \otimes \cdots \otimes V_2,$$

where V_1 occurs m_1 times and V_2 occurs m_2 times. Note that the action of $\mathsf{sl}(3;\mathbb{C})$ on this space is

$$\begin{aligned} Z \to\ & (\pi_1(Z) \otimes I \otimes \cdots \otimes I) \\ & + (I \otimes \pi_1(Z) \otimes I \otimes \cdots \otimes I) + \cdots + (I \otimes \cdots \otimes I \otimes \pi_2(Z)). \end{aligned} \tag{5.11}$$

Let π_{m_1,m_2} denote this representation.

Consider the vector

$$v_{m_1,m_2} = v_1 \otimes v_1 \otimes \cdots \otimes v_1 \otimes v_2 \otimes v_2 \otimes \cdots \otimes v_2.$$

Then, applying (5.11) shows that

$$\begin{aligned}
\pi_{m_1,m_2}(H_1)v_{m_1,m_2} &= m_1 v_{m_1,m_2}, \\
\pi_{m_1,m_2}(H_2)v_{m_1,m_2} &= m_2 v_{m_1,m_2}, \\
\pi_{m_1,m_2}(X_1)v_{m_1,m_2} &= 0, \\
\pi_{m_1,m_2}(X_2)v_{m_1,m_2} &= 0.
\end{aligned} \tag{5.12}$$

Now, the representation π_{m_1,m_2} is *not* irreducible (unless $(m_1, m_2) = (1, 0)$ or $(0, 1)$). However, if we let W denote the smallest invariant subspace containing the vector v_{m_1,m_2}, then, in light of (5.12), W will be highest weight cyclic with highest weight (m_1, m_2). Therefore, by Proposition 5.16, W is irreducible with highest weight (m_1, m_2).

Thus, W is the representation we want. □

We have now completed the proof of Theorem 5.9.

5.5 An Example: Highest Weight (1, 1)

To obtain the irreducible representation with highest weight $(1, 1)$, we are supposed to take the tensor product of the irreducible representations with highest weights $(1, 0)$ and $(0, 1)$, and then extract a certain invariant subspace. Let us establish some notation for the representations $(1, 0)$ and $(0, 1)$. In the standard representation, the weight vectors for

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

are the standard basis elements for $\mathbb{C}^3$, namely e_1, e_2, and e_3. The corresponding weights are $(1, 0)$, $(-1, 1)$, and $(0, -1)$. The highest weight is $(1, 0)$.

Recall that

$$Y_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; Y_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Thus,

$$\begin{aligned}
Y_1(e_1) &= e_2, \;\; Y_2(e_1) = 0, \\
Y_1(e_2) &= 0, \;\; Y_2(e_2) = e_3, \\
Y_1(e_3) &= 0, \;\; Y_2(e_3) = 0.
\end{aligned} \tag{5.13}$$

Now, the representation with highest weight $(0, 1)$ is the representation $\pi(Z) = -Z^{tr}$, for $Z \in \mathfrak{sl}(3; \mathbb{C})$. Let us define

$$\overline{Z} = -Z^{tr}$$

for all $Z \in \mathfrak{sl}(3; \mathbb{C})$. Thus, $\pi(Z) = \overline{Z}$. Note that

$$\overline{H_1} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \overline{H_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The weight vectors are again e_1, e_2, and e_3, with weights $(-1,0)$, $(1,-1)$, and $(0,1)$, respectively. The highest weight is $(0,1)$.

Define new basis elements

$$\begin{aligned} f_1 &= e_3, \\ f_2 &= -e_2, \\ f_3 &= e_1. \end{aligned}$$

Then, since

$$\overline{Y_1} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \overline{Y_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

we have

$$\begin{aligned} \overline{Y_1}(f_1) &= 0, \quad \overline{Y_2}(f_1) = f_2, \\ \overline{Y_1}(f_2) &= f_3, \quad \overline{Y_2}(f_2) = 0, \\ \overline{Y_1}(f_3) &= 0, \quad \overline{Y_2}(f_3) = 0. \end{aligned} \tag{5.14}$$

Note that the highest weight vector is $f_1 = e_3$.

So, to obtain an irreducible representation with highest weight $(1,1)$, we are supposed to take the tensor product of the representations with highest weights $(1,0)$ and $(0,1)$, and then take the smallest invariant subspace containing the vector $e_1 \otimes f_1$. In light of the proof of Proposition 5.13, this smallest invariant subspace is obtained by starting with $e_1 \otimes f_1$ and applying all possible combinations of Y_1 and Y_2.

Recall that if π_1 and π_2 are two representations of the Lie algebra $\mathfrak{sl}(3;\mathbb{C})$, then

$$\begin{aligned} (\pi_1 \otimes \pi_2)(Y_1) &= \pi_1(Y_1) \otimes I + I \otimes \pi_2(Y_1), \\ (\pi_1 \otimes \pi_2)(Y_2) &= \pi_1(Y_2) \otimes I + I \otimes \pi_2(Y_2). \end{aligned}$$

In our case, we want $\pi_1(Y_i) = Y_i$ and $\pi_2(Y_i) = \overline{Y_i}$. Thus,

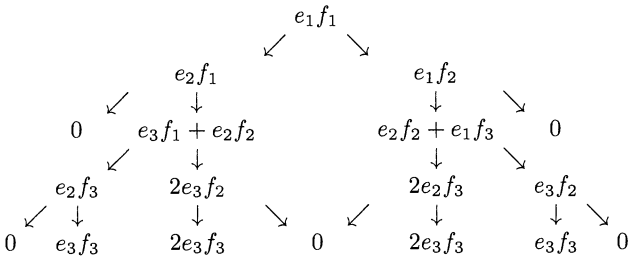
$$\begin{aligned} (\pi_1 \otimes \pi_2)(Y_1) &= Y_1 \otimes I + I \otimes \overline{Y_1}, \\ (\pi_1 \otimes \pi_2)(Y_2) &= Y_2 \otimes I + I \otimes \overline{Y_2}. \end{aligned}$$

The actions of Y_i and $\overline{Y_i}$ are described in (5.13) and (5.14).

Note that $\pi_1 \otimes \pi_2$ is *not* an irreducible representation. The representation $\pi_1 \otimes \pi_2$ has dimension 9, whereas the smallest invariant subspace containing $e_1 \otimes f_1$ has, as it turns out, dimension 8.

So, it remains only to begin with $e_1 \otimes f_1$, apply $Y_1 \otimes I + I \otimes \overline{Y_1}$ and $Y_2 \otimes I + I \otimes \overline{Y_2}$ repeatedly until we get zero, and then figure out what dependence relations exist among the vectors we get. This calculation is contained in the

following chart. Here, there are two arrows coming out of each vector. Of these, the left arrow indicates the action of $Y_1 \otimes I + I \otimes \overline{Y_1}$ and the right arrow indicates the action of $Y_2 \otimes I + I \otimes \overline{Y_2}$. To save space, I have omitted the tensor product symbol and written, for example, $e_2 f_2$ instead of $e_2 \otimes f_2$.



A basis for the space spanned by these vectors is $e_1 f_1$, $e_2 f_1$, $e_1 f_2$, $e_3 f_1 + e_2 f_2$, $e_2 f_2 + e_1 f_3$, $e_2 f_3$, $e_3 f_2$, and $e_3 f_3$. (These vectors are linearly independent and every vector listed above is a constant multiple of one of these.) So, the dimension of this representation is 8; it is (isomorphic to) the adjoint representation.

The weights for this representation are $(1,1)$, $(-1,2)$, $(2,-1)$, $(0,0)$, $(1,-2)$, $(-2,1)$, and $(-1,-1)$. Each weight has multiplicity 1 except for $(0,0)$, which has multiplicity 2 because $e_3 f_1 + e_2 f_2$ and $e_2 f_2 + e_1 f_3$ are both weight vectors with weight $(0,0)$.

5.6 The Weyl Group

There is an important symmetry to the representations of $\mathsf{sl}(3;\mathbb{C})$ involving something called the Weyl group. (Our treatment will follow the compact group approach, which is apparently different from the Lie algebra approach, but ultimately equivalent to it.) To understand the idea behind the Weyl group symmetry, let us observe that the representations of $\mathsf{sl}(3;\mathbb{C})$ are, in a certain sense, invariant under the adjoint action of $\mathsf{SU}(3)$. What I mean by this is the following. Let π be a finite-dimensional representation of $\mathsf{sl}(3;\mathbb{C})$ acting on a vector space V and let Π be the associated representation of $\mathsf{SU}(3)$ acting on the same space. For any $A \in \mathsf{SU}(3)$, we can define a new representation π_A of $\mathsf{sl}(3;\mathbb{C})$, acting on the same vector space V, by setting

$$\pi_A(X) = \pi(AXA^{-1}).$$

Since the adjoint action of A on $\mathsf{sl}(3;\mathbb{C})$ is a Lie algebra automorphism, this is, again, a representation of $\mathsf{sl}(3;\mathbb{C})$. This new representation is easily seen to be equivalent to the original representation; direct calculation shows that $\Pi(A)$ is an intertwining map between (π, V) and (π_A, V). We may say, then,

that the adjoint action of $\mathsf{SU}(3)$ is a symmetry of the set of equivalence classes of representations of $\mathsf{sl}(3;\mathbb{C})$.

Now, we have analyzed the representations of $\mathsf{sl}(3;\mathbb{C})$ by simultaneously diagonalizing the operators $\pi(H_1)$ and $\pi(H_2)$. Of course, this means that any linear combination of $\pi(H_1)$ and $\pi(H_2)$ is also simultaneously diagonalized. So, what really counts is the two-dimensional subspace $\mathfrak{h}$ of $\mathsf{sl}(3;\mathbb{C})$ spanned by H_1 and H_2. (This space is called a **Cartan subalgebra** of $\mathsf{sl}(3;\mathbb{C})$. See Chapter 6 for more information.) Now, in general, the adjoint action of $A \in \mathsf{SU}(3)$ will not preserve the space $\mathfrak{h}$ and so the equivalence of π and π_A does not (in general) tell us anything about the weights of π. However, there are elements A in $\mathsf{SU}(3)$ for which Ad_A *does* preserve $\mathfrak{h}$. These elements make up the Weyl group for $\mathsf{SU}(3)$ and (as we shall see below) give rise to a symmetry of the set of weights of any representation π. So, we may say that the Weyl group is the "residue" of the adjoint symmetry of the representations (discussed in the previous paragraph) that is left after we focus our attention on the subspace $\mathfrak{h}$ of $\mathsf{sl}(3;\mathbb{C})$.

Definition 5.20. *Let $\mathfrak{h}$ be the two-dimensional subspace of $\mathsf{sl}(3;\mathbb{C})$ spanned by H_1 and H_2. Let Z be the subgroup of $\mathsf{SU}(3)$ consisting of those $A \in \mathsf{SU}(3)$ such that $\mathrm{Ad}_A(H) = H$ for all $H \in \mathfrak{h}$. Let N be the subgroup of $\mathsf{SU}(3)$ consisting of those $A \in \mathsf{SU}(3)$ such that $\mathrm{Ad}_A(H)$ is an element of $\mathfrak{h}$ for all H in $\mathfrak{h}$.*

It is a straightforward exercise (Exercise 8) to verify that Z and N are actually subgroups of $\mathsf{SU}(3)$ and to verify that Z is a normal subgroup of N. This leads us to the definition of the Weyl group.

Definition 5.21. *The **Weyl group** of $\mathsf{SU}(3)$, denoted W, is the quotient group N/Z.*

We can define an action of W on $\mathfrak{h}$ as follows. For each element w of W, choose an element A of the corresponding equivalence class in N. Then for H in $\mathfrak{h}$ we define the action $w \cdot H$ of w on H by

$$w \cdot H = \mathrm{Ad}_A(H).$$

To see that this action is well defined, suppose B is another element of the same equivalence class as A. Then $B = AC$ with $C \in Z$ and, thus,

$$\mathrm{Ad}_B(H) = \mathrm{Ad}_A \mathrm{Ad}_C(H) = \mathrm{Ad}_A(H),$$

by the definition of Z. It is easily seen that W is isomorphic to the group of linear transformations of $\mathfrak{h}$ that can be expressed as Ad_A for some $A \in N$.

The following proposition will allow us to compute W explicitly.

Proposition 5.22. *The group Z consists precisely of the diagonal matrices inside $\mathsf{SU}(3)$, namely the matrices of the form*

$$A = \begin{pmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{i\phi} & 0 \\ 0 & 0 & e^{-i(\theta+\phi)} \end{pmatrix} \tag{5.15}$$

for θ and ϕ in $\mathbb{R}$. The group N consists of precisely those matrices $A \in \mathrm{SU}(3)$ such that for each $k = 1, 2, 3$, there exist $l \in \{1, 2, 3\}$ and $\theta \in \mathbb{R}$ such that $Ae_k = e^{i\theta} e_l$. Here, e_1, e_2, e_3 is the standard basis for $\mathbb{C}^3$.

The Weyl group $W = N/Z$ is isomorphic to the permutation group on three elements.

Proof. Suppose A is in Z, which means that A commutes with all elements of $\mathfrak{h}$. Then, certainly, A must commute with H_1. Now, the matrix H_1 has eigenvalues 1, -1, and 0. The corresponding eigenspaces are the span of e_1, the span of e_2, and the span of e_3. Since A commutes with H_1, it must preserve each of these eigenspaces (Proposition B.4). This means that Ae_k must be a multiple of e_k for each $k = 1, 2, 3$. This is the same as saying that A is diagonal. If A is also to be unitary and have determinant 1 then it must be of the form in the proposition. Conversely, any matrix of the form (5.15) does indeed commute not only with H_1 but also with H_2 and, thus, with every element of $\mathfrak{h}$. So, Z consists precisely of the form (5.15).

Now, suppose that A is in N. Then, AH_1A^{-1} must be in $\mathfrak{h}$ and therefore must be diagonal. Now, H_1 has eigenvectors e_1, e_2, and e_3 with *distinct* eigenvalues 1, -1, 0. Then, AHA^{-1} will have eigenvectors Ae_1, Ae_2, and Ae_3 with the same eigenvalues, 1, -1, 0. Since the eigenvalues of AHA^{-1} are distinct, the *only* eigenvectors it has are multiples of Ae_1, multiples of Ae_2, and multiples of Ae_3. (This would not be the case if AHA^{-1} had a repeated eigenvalue.) On the other hand, AHA^{-1} is diagonal, which means it has e_1, e_2, and e_3 as eigenvectors. The only way these two descriptions of the eigenvectors of AHA^{-1} can agree is if each Ae_k is a constant multiple of some e_l. The constant must have absolute value 1 if A is unitary.

Conversely, suppose A is in $\mathrm{SU}(3)$ and A takes each e_k to $e^{i\theta}e_l$. Then, the eigenvectors for AH_1A^{-1} will still be e_1, e_2, and e_3 (but with the eigenvalues possibly in a different order) and so AH_1A^{-1} will be diagonal. Furthermore, since H_1 has trace zero, AH_1A^{-1} will also have trace zero. However, $\mathfrak{h}$ consists of all diagonal 3×3 matrices with trace zero, and so this shows that AH_1A^{-1} is in $\mathfrak{h}$. The same argument shows that AH_2A^{-1} is in $\mathfrak{h}$.

Now, let us think about what Ad_A looks like as a linear transformation of $\mathfrak{h}$, for A in N. For $k = 1, 2, 3$, let $\sigma(k)$ be the element of $\{1, 2, 3\}$ such that A maps e_k to a multiple of $e_{\sigma(k)}$. Since A is invertible, the map $k \to \sigma(k)$ must be a permutation of the set $\{1, 2, 3\}$. Now each element H of $\mathfrak{h}$ is a diagonal matrix, which means that H has e_1, e_2, and e_3 as eigenvectors; the diagonal entries of H are precisely the corresponding eigenvalues λ_1, λ_2, and λ_3. Then, AHA^{-1} has $e_{\sigma(k)}$ as an eigenvector with eigenvalue λ_k. This means that the diagonal entry of H that was originally in the k^{th} spot of H is now in the $\sigma(k)^{\text{th}}$ spot.

We see, then, that A acts on $\mathfrak{h}$ by permuting the diagonal entries of each $H \in \mathfrak{h}$ according to the permutation σ. Thus, the group of linear transformations of $\mathfrak{h}$ of the form Ad_A, $A \in N$, is isomorphic to the permutation group on three elements. This group of linear transformations is (isomorphic to) the Weyl group $W = N/Z$.

Note that although each entry of A maps each e_k to some constant multiple $e^{i\theta_k}$ of $e_{\sigma(k)}$, the action of Ad_A on $\mathfrak{h}$ depends only on the value of $\sigma(k)$ and not on the constants $e^{i\theta_k}$. This reflects that if one multiplies any $A \in N$ on the right by some $B \in Z$, then $\mathrm{Ad}_{AB}(H) = \mathrm{Ad}_A \mathrm{Ad}_B(H) = \mathrm{Ad}_A(H)$, since by the definition of Z, the adjoint action of B on $\mathfrak{h}$ is just the identity. □

In the case of $\mathsf{SU}(3)$, it is possible to identify the Weyl group with a certain subgroup of N, instead of as the quotient group N/Z. See Exercise 9. Exercise 10 asks one to verify by direct calculation that the action of a particular element A of N is as described in the above proof.

We want to show that the Weyl group is a symmetry of the weights of any finite-dimensional representation of $\mathsf{sl}(3;\mathbb{C})$. To understand this, we need to adopt a less basis-dependent view of the weights. We have defined a weight as a pair (m_1, m_2) of simultaneous eigenvalues for $\pi(H_1)$ and $\pi(H_2)$. However, if a vector v is an eigenvector for $\pi(H_1)$ and $\pi(H_2)$ then it is also an eigenvector for $\pi(H)$ for any element H of the space $\mathfrak{h}$ spanned by H_1 and H_2. Furthermore, the eigenvalues must depend linearly on H since if H and J are any two elements of $\mathfrak{h}$ and $\pi(H)v = \lambda_1 v$ and $\pi(J) = \lambda_2 v$, then

$$\begin{aligned}\pi(aH + bJ)v &= (a\pi(H) + b\pi(J))v \\ &= (a\lambda_1 + b\lambda_2)v.\end{aligned}$$

So, we may make the following basis-independent notion of a weight.

Definition 5.23. *Let $\mathfrak{h}$ be the subspace of $\mathsf{sl}(3;\mathbb{C})$ spanned by H_1 and H_2 and let π be a finite-dimensional representation of $\mathsf{sl}(3;\mathbb{C})$ acting on a vector space V. A linear functional $\mu \in \mathfrak{h}^*$ is called a **weight** for π if there exists a nonzero vector v in V such that*

$$\pi(H)v = \mu(H)v$$

*for all H in $\mathfrak{h}$. Such a vector v is called a **weight vector** with weight μ.*

So, a weight is just a collection of simultaneous eigenvalues of all the elements H of $\mathfrak{h}$, which, as we have noted, must depend linearly on H and which, therefore, define a linear functional on $\mathfrak{h}$. Since H_1 and H_2 span $\mathfrak{h}$, the linear functional μ is determined by the value of $\mu(H_1)$ and $\mu(H_2)$, and thus our new notion of weight is equivalent to our old notion of a weight as just a pair of simultaneous eigenvalues of $\pi(H_1)$ and $\pi(H_2)$. The reason for adopting this basis-independent approach is that the action of the Weyl group does not preserve the basis $\{H_1, H_2\}$ for $\mathfrak{h}$.

The Weyl group is (or may be thought of as) a group of linear transformations of $\mathfrak{h}$. This means that W acts linearly on $\mathfrak{h}$, and we denote this action as

$w \cdot H$. We can define an associated action on the dual space $\mathfrak{h}^*$ as in the definition of the dual representation in Chapter 4. Thus, for $\mu \in \mathfrak{h}^*$ and $w \in W$, we define $w \cdot \mu$ to be the element of $\mathfrak{h}^*$ given by

$$(w \cdot \mu)(H) = \mu(w^{-1} \cdot H). \tag{5.16}$$

We now come to the main point of the Weyl group from the point of view of representation theory, namely that the weights of any representation are invariant under the action of the Weyl group.

Theorem 5.24. *Suppose that π is any finite-dimensional representation of $\mathsf{sl}(3;\mathbb{C})$ and that $\mu \in \mathfrak{h}^*$ is a weight for π. Then, for any $w \in W$, $w \cdot \mu$ is also a weight of π, and the multiplicity of $w \cdot \mu$ is the same as the multiplicity of μ.*

Proof. Suppose that μ is a weight for a representation (π, V) of $\mathsf{sl}(3;\mathbb{C})$ and suppose that v is a weight vector with weight μ. Then, let Π be the associated representation of the simply-connected group $\mathsf{SU}(3)$ and consider the vector $\Pi(A)v$, for $A \in N$. We want to show that $\Pi(A)v$ is, again, a weight vector. So, we compute

$$\begin{aligned}\pi(H)\Pi(A)v &= \Pi(A)\Pi(A)^{-1}\pi(H)\Pi(A)v \\ &= \Pi(A)\pi(A^{-1}HA)v \\ &= \mu(A^{-1}HA)\Pi(A)v.\end{aligned}$$

Here, we have used that A is in N, which guarantees that $A^{-1}HA$ is, again, in $\mathfrak{h}$. However, $A^{-1}HA$ is nothing but $w^{-1} \cdot H$, where w is the Weyl group element represented by A. Thus, by (5.16), $\mu(A^{-1}HA) = (w \cdot \mu)(H)$. This shows that $\Pi(A)v$ is, again, a weight vector, with weight $w \cdot \mu$ and thus $w \cdot \mu$ is, again, a weight for (π, V). The same sort of reasoning shows that $\Pi(A)$ is an invertible map of the weight space associated to the weight μ onto the weight space with weight $w \cdot \mu$, whose inverse is $\Pi(A)^{-1}$. This means that the two weight spaces have the same dimension and, therefore, μ and $w \cdot \mu$ have the same multiplicity. □

Note that since the roots are nothing but the nonzero weights of the adjoint representation, this result tells us that the roots are invariant under the action of the Weyl group. In order to visualize the action of the Weyl group, it is convenient to identify $\mathfrak{h}^*$ with $\mathfrak{h}$ by means of an inner product on $\mathfrak{h}$ that is invariant under the action of the Weyl group. Recall that $\mathfrak{h}$ is a subspace of the space of diagonal matrices, and we use on the space of diagonal matrices the inner product obtained by identifying with $\mathbb{C}^3$ in the obvious way. (This inner product is, if one prefers, the restriction to the diagonal matrices of the Hilbert–Schmidt inner product $\langle A, B\rangle = \mathrm{trace}(A^*B)$. See Section B.6.) Since the Weyl group acts by permuting the diagonal entries, this inner product (restricted to the subspace $\mathfrak{h}$) is preserved by the action of W.

We now use this inner product on $\mathfrak{h}^*$ to identify $\mathfrak{h}$. Given any element α of $\mathfrak{h}$, the map $H \to \langle \alpha, H \rangle$ is a linear functional on $\mathfrak{h}$ (i.e., an element of $\mathfrak{h}^*$). Every linear functional on $\mathfrak{h}$ can be represented in this way for a unique α in $\mathfrak{h}$ (Section B.7). We will now simply identify each linear functional with the corresponding element of $\mathfrak{h}$. Thus, we will now regard a weight for (π, V) as a nonzero element of $\mathfrak{h}$ with the property that there exists a nonzero v in V such that

$$\pi(H)v = \langle \alpha, H \rangle v \tag{5.17}$$

for all H in $\mathfrak{h}$. This is the same as Definition 5.23 except that, now, α lives in $\mathfrak{h}$ and we write $\langle \alpha, H \rangle$ instead of $\alpha(H)$ on the right. The roots, being weights for the adjoint representation, are viewed in a similar way.

Now that the roots and weights live in $\mathfrak{h}$ instead of $\mathfrak{h}^*$, we can use the above inner product on $\mathfrak{h}$. Furthermore, it can be shown (Exercise 11) that under our identification of $\mathfrak{h}^*$ with $\mathfrak{h}$, the action of W on $\mathfrak{h}^*$ (described in (5.16)) coincides with the adjoint action of W on $\mathfrak{h}$.

We are now ready to begin calculating. I claim that with our new point of view the roots α_1 and α_2 are identified with the following elements of $\mathfrak{h}$:

$$\alpha_1 = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & & \\ & 1 & \\ & & -1 \end{pmatrix}.$$

To check this, we note that these matrices are indeed in $\mathfrak{h}$ since the diagonal entries sum to zero. Then, direct calculation shows that $\langle \alpha_1, H_1 \rangle = 2$, $\langle \alpha_1, H_2 \rangle = -1$ and $\langle \alpha_2, H_1 \rangle = -1$, $\langle \alpha_2, H_2 \rangle = 2$, in agreement with our earlier definition (5.4) of α_1 and α_2. So, then, we can compute the lengths and angles as $\|\alpha_1\|^2 = \langle \alpha_1, \alpha_1 \rangle = 2$, $\|\alpha_2\|^2 = \langle \alpha_2, \alpha_2 \rangle = 2$, and $\langle \alpha_1, \alpha_2 \rangle = -1$. This means that (with respect to this inner product) α_1 and α_2 both have length $\sqrt{2}$ and the angle θ between them satisfies $\cos\theta = -1/2$, so that $\theta = 120°$.

We now consider the dominant integral elements, which are the possible highest weights of irreducible representations of $\mathfrak{sl}(3;\mathbb{C})$. With our new point of view, these are the elements μ of $\mathfrak{h}$ such that $\langle \mu, H_1 \rangle$ and $\langle \mu, H_2 \rangle$ are nonnegative integers. We begin by considering the **fundamental weights** μ_1 and μ_2 defined by

$$\begin{aligned} \langle \mu_1, H_1 \rangle &= 1, \quad \langle \mu_2, H_1 \rangle = 0, \\ \langle \mu_1, H_2 \rangle &= 0, \quad \langle \mu_2, H_2 \rangle = 1. \end{aligned}$$

A little trial and error shows that these can be expressed in terms of α_1 and α_2 as follows:

$$\begin{aligned} \mu_1 &= \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2, \\ \mu_2 &= \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2. \end{aligned}$$

Plugging in the expressions for α_1 and α_2, we get

$$\mu_1 = \begin{pmatrix} \frac{2}{3} & & \\ & -\frac{1}{3} & \\ & & -\frac{1}{3} \end{pmatrix}, \quad \mu_1 = \begin{pmatrix} \frac{1}{3} & & \\ & \frac{1}{3} & \\ & & -\frac{2}{3} \end{pmatrix}.$$

An elementary calculation then shows that μ_1 and μ_2 each have length $\sqrt{6}/3$ and that the angle between them is 60°. The set of dominant integral elements is then precisely the set of linear combinations of μ_1 and μ_2 with non-negative integer coefficients. Note that $\mu_1 + \mu_2 = \alpha_1 + \alpha_2$, an observation that helps in drawing Figure 5.2 below.

We are now finally ready to draw some pictures. Figure 5.2 shows the same information as Figure 5.1, namely, the roots and the dominant integral elements, but now drawn relative to a Weyl-invariant inner product. We draw only the two-dimensional *real* subspace of $\mathfrak{h}$ consisting of those elements μ such that $\langle \mu, H_1 \rangle$ and $\langle \mu, H_2 \rangle$ are real, since all the roots and weights have this property. In this figure, the arrows indicate the roots, the black dots indicate dominant integral elements (i.e., points μ such that $\langle \mu, H_1 \rangle$ and $\langle \mu, H_2 \rangle$ are non-negative integers), and the triangular grid indicates integral elements (i.e., points μ such that $\langle \mu, H_1 \rangle$ and $\langle \mu, H_2 \rangle$ are integers).

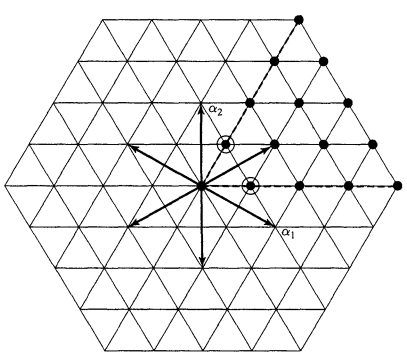


Fig. 5.2. Roots and dominant integral elements for $\mathfrak{sl}(3;\mathbb{C})$ (using Weyl-invariant inner product)

Let us see how the Weyl group acts on Figure 5.2. Let $(1,2,3)$ denote the cyclic permutation that takes 1 to 2 to 3 to 1, and let $w_{(1,2,3)}$ denote the corresponding Weyl group element (Exercise 10). Then, $w_{(1,2,3)}$ acts by cyclically permuting the diagonal entries of each element of H. Thus, $w_{(1,2,3)}$ takes α_1 to α_2 and takes α_2 to $-(\alpha_1 + \alpha_2)$. This action is a 120° rotation, counterclockwise in Figure 5.2. Next, let $(1,2)$ be the permutation that interchanges 1 and 2 and let $w_{(1,2)}$ be the corresponding Weyl group element. Then, $w_{(1,2)}$ acts by interchanging the first two diagonal entries of each element of H, and

thus takes α_1 to $-\alpha_1$ and takes α_2 to $\alpha_1 + \alpha_2$. This corresponds to a reflection about the line perpendicular to α_1. The reader is invited to calculate the action of the remaining Weyl group elements. The Weyl group consists of six elements: the identity, clockwise and counterclockwise rotations by 120°, and three reflections—about the line perpendicular to α_1, about the line perpendicular to α_2, and about the line perpendicular to $\alpha_1 + \alpha_2$. This is the symmetry of an equilateral triangle centered at the origin, as indicated in Figure 5.3.

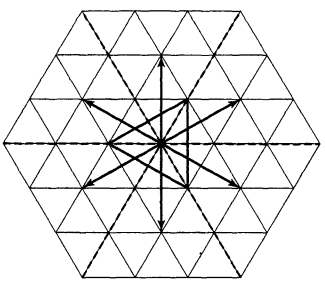


Fig. 5.3. Weyl group for $\mathsf{sl}(3;\mathbb{C})$

5.7 Weight Diagrams

In this section, we consider weight diagrams for $\mathsf{sl}(3;\mathbb{C})$ (i.e., pictures of the weights of various representations of $\mathsf{sl}(3;\mathbb{C})$). (These weight diagrams should not be confused with Dynkin diagrams, which are discussed in Chapter 8.) The action of the Weyl group is critical to understanding which weights arise in a representation with a given highest weight μ_0.

Definition 5.25. *Let $v_1, \ldots, v_n$ be a finite collection of points in a vector space V. The* ***convex hull*** *of $v_1, \ldots, v_n$ is the set of all vectors in V that can be expressed as*

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n,$$

where $c_1, \ldots, c_n$ are non-negative real constants satisfying $c_1 + \cdots + c_n = 1$.

Equivalently, the convex hull is the smallest convex subset of V containing all of the points $v_1, \ldots, v_n$.

To make the weight diagrams, we make use of the following result. As in (5.17), we regard the weights (and so also the roots) as elements of $\mathfrak{h}$.

Theorem 5.26. *Suppose that π is an irreducible representation of $\mathsf{sl}(3;\mathbb{C})$ with highest weight μ_0. Then, an element μ of $\mathfrak{h}$ is a weight of the representation π if and only if the following two conditions are satisfied:*

1. *μ is contained in the convex hull of the orbit of μ_0 under the Weyl group.*
2. *$\mu_0 - \mu$ is expressible as a linear combination of α_1 and α_2 with integer coefficients.*

Let us think first about Condition 1. The Weyl group of SU(3) has six elements and the orbit of a "generic" point in $\mathfrak{h}$ will consist of six points. These six points will form the vertices of a hexagon. The simplest way to see how this will work out is first to apply to μ_0 a reflection (say, about the line perpendicular to α_1) and then to apply 120° clockwise and counterclockwise rotations to the resulting pair of points to get a total of six points. Suppose, however, that one starts with a dominant integral element μ_0 that is on the edge of the set of all dominant integral elements (these are the elements of the form $(m_1, 0)$ or $(0, m_2)$ in our old view of the weights). Then, μ_0 is left invariant by either the reflection about the line perpendicular to α_1 or by the reflection about the line perpendicular to α_2. In that case, the orbit of μ_0 is a triangle (unless $\mu_0 = 0$, in which, case the orbit is just a single point). Finally, the convex hull of the orbit of μ_0 will be a filled-in triangle or hexagon (or a single point if $\mu_0 = 0$).

Let us think now about Condition 2. Condition 2 implies that μ must be an integral element (i.e., that $\langle \mu, H_1 \rangle$ and $\langle \mu, H_2 \rangle$ are integers), since μ_0, α_1, and α_2 all have this property. However, not every integral element will satisfy Condition 2. Suppose, for example, that $\mu_0 = (1, 0)$ and $\mu = (0, 0)$. Then, $\mu_0 - \mu = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$ and the coefficients are not integral. So, $(0, 0)$ is not a weight of the irreducible representation with highest weight $(1, 0)$. In most cases, there will be integral elements contained in the convex hull of the orbit of μ_0 that are not weights of the representation with highest weight μ_0.

I will give only the main idea of the proof of Theorem 5.26. It is not too hard to show that Conditions 1 and 2 are both necessary conditions for the weights of the representation with highest weight μ_0. See Exercises 13 and 14. Showing that the conditions are sufficient is an $\mathfrak{sl}(2;\mathbb{C})$ argument and makes use of the fact that there can be no "gaps" in the eigenvalues of H in a finite-dimensional representation of $\mathfrak{sl}(2;\mathbb{C})$: If a non-negative integer k is an eigenvalue for H in some representation, then so are $k-2, k-4, \ldots, -k$. (One starts with the orbit of μ_0 under the Weyl group and then uses the just-mentioned result, applied to various $\mathfrak{sl}(2;\mathbb{C})$ subalgebras of $\mathfrak{sl}(3;\mathbb{C})$, to "fill in" all the elements satisfying Conditions 1 and 2.)

Theorem 5.26 tells us which weights occur in a given representation of $\mathfrak{sl}(3;\mathbb{C})$ but not what the multiplicities of the weights are. It can be shown that the multiplicities obey the following simple pattern. The weights occur in "rings" in which the rings toward the outside are hexagons and the rings toward the inside are triangles. The weights in the outermost ring have multiplicity 1. The multiplicities then increase by 1 each time one moves inward one ring, until the rings become triangles, at which point the multiplicities stabilize. The situation for the multiplicities in other semisimple Lie algebras is more complicated—see Section 7.6.

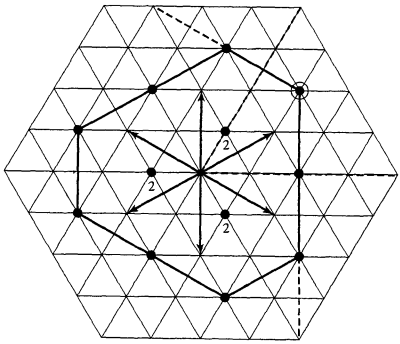


Fig. 5.4. Highest weight (1,2)

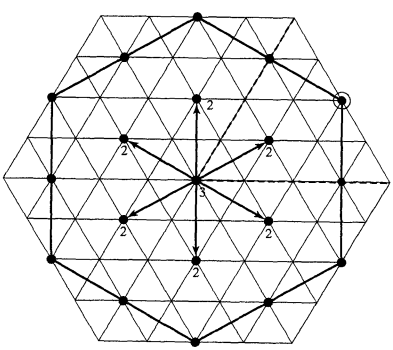


Fig. 5.5. Highest weight (2,2)

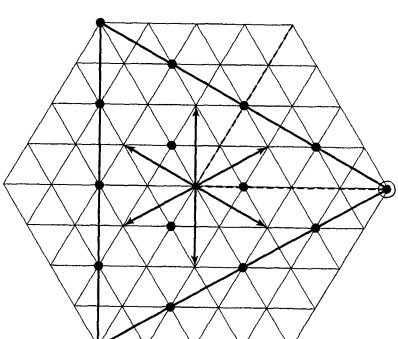


Fig. 5.6. Highest weight (4,0)

Figures 5.4, 5.5, and 5.6 show the weights and multiplicities for three irreducible representations, with highest weights $(1,2)$, $(2,2)$, and $(4,0)$, respectively. In each figure, a black dot indicates a weight of the representation, with the highest weight being circled. A number next to a dot indicates the multiplicity of the corresponding weight. A dot without a number indicates a weight of multiplicity one. In Figure 5.4, the dashed lines extending from the highest weights indicate the boundary of the set of points that are lower than $(1,2)$. The dimensions of these representations are 15, 27, and 15, respectively, as can be computed either from the dimension formula (Theorem 5.10) or by adding up the multiplicities of all the weights.

5.8 Exercises

1. Show that the roots listed in (5.3) are the only roots.
2. Let π be an irreducible finite-dimensional representation of $\mathsf{sl}(3;\mathbb{C})$ acting on a space V and let π^* be the dual representation to π, acting on V^*, as defined in Section 4.7. Show that the weights of π^* are the negatives of the weights of π.
   Hint: Choose a basis for V in which both $\pi(H_1)$ and $\pi(H_2)$ are diagonal.
3. As in Exercise 2, let π be an irreducible representation of $\mathsf{sl}(3;\mathbb{C})$ and let π^* be the dual representation to π. Show that if π has highest weight (m_1,m_2), π^* has highest weight (m_2,m_1).
   Hint: Establish this first in the cases $(m_1,m_2)=(1,0)$ and $(m_1,m_2)=(0,1)$.
4. Consider the adjoint representation of $\mathsf{sl}(3;\mathbb{C})$ as a representation of $\mathsf{sl}(2;\mathbb{C})$ by restricting the adjoint representation to the subalgebra spanned by X_1, Y_1, and H_1. Decompose this representation as a direct sum of irreducible representations of $\mathsf{sl}(2;\mathbb{C})$. Which representations occur and with what multiplicity?
5. Following the method of Section 5.5, work out the representation of $\mathsf{sl}(3;\mathbb{C})$ with highest weight $(2,0)$, acting on a subspace of $\mathbb{C}^3\otimes\mathbb{C}^3$. Determine all the weights of this representation and their multiplicity (i.e., the dimension of the corresponding weight space). Verify that the dimension formula (Theorem 5.10) holds in this case.
6. Consider the nine-dimensional representation of $\mathsf{sl}(3;\mathbb{C})$ considered in Section 5.5, namely the tensor product of the representations with highest weights $(1,0)$ and $(0,1)$. Decompose this representation as a direct sum of irreducibles. Do the same for the tensor product of two copies of the irreducible representation with highest weight $(1,0)$. (Compare Exercise 5.)
7. Let V_m denote the space of homogeneous polynomials on $\mathbb{C}^3$ of degree m. By imitating Section 4.3, construct a representation of $\mathsf{SU}(3)$ acting on V_m. Find the weights for the associated action of $\mathsf{sl}(3;\mathbb{C})$ on V_1 and V_2.

Show that V_1 and V_2 are irreducible representations (of $\mathsf{SU}(3)$ or $\mathsf{sl}(3;\mathbb{C})$). What are the highest weights of these representations?

8. Show that Z and N (defined in Definition 5.20) are subgroups of $\mathsf{SU}(3)$. Show that Z is a normal subgroup of N.
9. For each permutation σ of $\{1,2,3\}$, let A_σ be the matrix such that $Ae_k = \mathrm{sgn}(\sigma)\, e_{\sigma(k)}$, where $\mathrm{sgn}(\sigma)$ is the sign of the permutation σ, equal to 1 for even permutations and equal to -1 for odd permutations. Show that the matrices A_σ form a subgroup of N that is isomorphic to W.
10. Consider the matrix A in $\mathsf{SU}(3)$ given by

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

which maps e_1 to e_2, e_2 to e_3, and e_3 to e_1. Let H be an arbitrary element of $\mathfrak{h}$ and let λ_1, λ_2, and λ_3 be the diagonal entries of H (which must sum to zero). Compute by hand AHA^{-1} and verify that this is related to H as described in Section 5.6, namely that λ_1 gets shifted into the second spot, λ_2 gets shifted into the third spot, and λ_3 gets shifted into the first spot.

11. Show that under the identification of $\mathfrak{h}^*$ with $\mathfrak{h}$ described in Section 5.6, the action of W on $\mathfrak{h}^*$ (described in (5.16)) coincides with the adjoint action of W on $\mathfrak{h}$.
12. Regard the Weyl group as a group of linear transformations of $\mathfrak{h}$. Show that $-I$ is not an element of the Weyl group. Which representations of $\mathsf{sl}(3;\mathbb{C})$ have the property that their weights are invariant under $-I$?
13. Using the proof of Proposition 5.13, show that every weight μ of an irreducible representation with highest weight μ_0 must satisfy Condition 2 of Theorem 5.26.
14. This exercise asks one to "prove" geometrically the following result. Let μ_0 be a dominant integral element and μ any integral element. If $w \cdot \mu$ is lower than μ_0 for all $w \in W$, then μ is contained in the convex hull of the W-orbit of μ_0.

    To see why this result is true, make a picture of a typical dominant integral element μ_0 and its W-orbit. Now, take a typical point μ that is *not* in the convex hull of the orbit of μ_0 and draw its W-orbit. Show that the W-orbit of μ contains at least one point that is not lower than μ_0.

    This result (along with the invariance of the weights under the action of the Weyl group) shows that Condition 1 of Theorem 5.26 is a necessary condition for μ to be a weight of the representation with highest weight μ_0.

6

Semisimple Lie Algebras

In this chapter, we will consider a class of Lie algebras (the complex semisimple ones) that are sufficiently similar to $\mathsf{sl}(3;\mathbb{C})$ that their representations can be described, similarly to $\mathsf{sl}(3;\mathbb{C})$, by a "theorem of the highest weight." We will not come to the representations themselves until the next chapter; in this chapter, we develop the structures needed to state the theorem of the highest weight. Although this chapter could be understood simply as a description of the structure of semisimple Lie algebras, without any mention of representation theory, I think it is helpful to have the representations in mind. The representation theory, especially in light of our experience with $\mathsf{sl}(3;\mathbb{C})$, motivates the notions of Cartan subalgebras, roots, and the Weyl group.

We will give three equivalent characterizations of semisimple Lie algebras (and there are several other commonly used ones). The first characterization is the one that we will take as our definition and which presumably accounts for the term "semisimple": A semisimple Lie algebra is one which is isomorphic to a direct sum of simple Lie algebras. The second characterization is that a complex Lie algebra is semisimple if and only if it is isomorphic to the complexification of the Lie algebra of a compact simply-connected group. This characterization shows, for example, that $\mathsf{sl}(n;\mathbb{C}) \cong \mathsf{su}(n)_{\mathbb{C}}$ is semisimple. The third characterization is that a Lie algebra $\mathfrak{g}$ is semisimple if and only if it has the complete reducibility property, that is, if and only if every finite-dimensional representation of $\mathfrak{g}$ decomposes as a direct sum of irreducibles.

Before getting into the details of semisimple Lie algebras, let us briefly outline what our strategy will be in classifying their representations and what structures we will need to carry out this strategy. We will look for commuting elements $H_1, \ldots, H_r$ in our Lie algebra that we will try to simultaneously diagonalize in each representation. We should find as many such elements as possible, and if they are going to be simultaneously diagonalizable in every representation, they must certainly be diagonalizable in the adjoint representation. This leads (in basis-independent language) to the definition of a **Cartan subalgebra**. The nonzero sets of simultaneous eigenvalues for $\mathrm{ad}_{H_1}, \ldots, \mathrm{ad}_{H_r}$ are called **roots** and the corresponding simultaneous eigenvectors are called

root vectors. The root vectors will serve to raise and lower the eigenvalues of $\pi(H_1), \ldots, \pi(H_r)$ in each representation π. We will also have the **Weyl group**, which is an important symmetry of the roots and also of the weights in each representation. Finally, we will introduce the notion of **positive roots**, in terms of which the notion of "highest weight" will be defined.

One crucial part of the structure of semisimple Lie algebras is the existence of certain special subalgebras isomorphic to $\mathsf{sl}(2;\mathbb{C})$. Several times over the course of this chapter and the next one, we will make use of our knowledge of the representations of $\mathsf{sl}(2;\mathbb{C})$. In particular, if X, Y, and H are the usual basis elements for $\mathsf{sl}(2;\mathbb{C})$, then we will use repeatedly that the eigenvalues of $\sigma(H)$ in any finite-dimensional representation of $\mathsf{sl}(2;\mathbb{C})$ must be integers (Theorem 4.12). In view of the importance of this result, it is worthwhile now to recall why this is so. From the Lie algebra point of view, we began with an eigenvector for $\sigma(H)$ and then used $\sigma(X)$ and $\sigma(Y)$ to raise and lower the eigenvalues for $\sigma(H)$ in increments of 2. Since the representation is finite dimensional, this chain of eigenvalues must terminate in both directions. The calculations of Section 4.4 (especially Lemma 4.11) show that this can happen only if the highest eigenvalue m of $\sigma(H)$ is a non-negative integer. In that case, all of the other eigenvalues of $\sigma(H)$ are of the form $m - 2k$ and, so, are also integers.

From the group point of view, we recall that because $\mathsf{SU}(2)$ is simply connected, for each finite-dimensional representation σ of $\mathsf{sl}(2;\mathbb{C}) \cong \mathsf{su}(2)_{\mathbb{C}}$ there is a representation Σ of $\mathsf{SU}(2)$ such that $\Sigma(\exp X) = \exp \sigma(X)$ for all X in $\mathsf{su}(2)$. We note that $2\pi i H$ is in $\mathsf{su}(2)$ and that $\exp(2\pi i H) = I$ in $\mathsf{SU}(2)$. Thus,

$$\exp(2\pi i \sigma(H)) = \Sigma(\exp(2\pi i H)) = \Sigma(I) = I.$$

This can happen only if the eigenvalues of $\sigma(H)$ are integers. After all, if λ is an eigenvalue for $\sigma(H)$, then $\exp(2\pi i \lambda)$ is an eigenvalue for $\exp(2\pi i \sigma(H)) = I$, so $\exp(2\pi i \lambda) = 1$ and λ must be an integer.

6.1 Complete Reducibility and Semisimple Lie Algebras

Recall (Section 4.10) that a group or Lie algebra is said to have the *complete reducibility* property if every finite-dimensional representation of it decomposes as a direct sum of irreducible invariant subspaces. Recall also (Proposition 4.36) that a connected compact matrix Lie group always has this property. It follows that the Lie algebra of a compact *simply-connected* matrix Lie group also has the complete reducibility property, since, in that case, there is a one-to-one correspondence between the representations of the compact group and its Lie algebra. Since there is a one-to-one correspondence between the representations of a real Lie algebra and the complex-linear representations of its complexification, we see also that if a complex Lie algebra $\mathfrak{g}$ is isomorphic to the complexification of the Lie algebra of a compact simply-connected group,

then $\mathfrak{g}$ has the complete reducibility property. In Chapter 5, we applied this reasoning to $\mathsf{sl}(2;\mathbb{C})$ (the complexification of the Lie algebra of $\mathsf{SU}(2)$) and to $\mathsf{sl}(3;\mathbb{C})$ (the complexification of the Lie algebra of $\mathsf{SU}(3)$).

In this chapter, we will study complex Lie algebras that are isomorphic to the complexification of the Lie algebra of a compact simply-connected matrix Lie group. As it turns out, such Lie algebras are precisely the complex semisimple Lie algebras, which we now define. Although we will mostly be concerned with *complex* semisimple Lie algebras, there is a brief discussion of real semisimple Lie algebras at the end of this section.

Definition 6.1. *If $\mathfrak{g}$ is a complex Lie algebra, then an* ***ideal*** *in $\mathfrak{g}$ is a complex subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ with the property that for all X in $\mathfrak{g}$ and H in $\mathfrak{h}$, we have $[X,H]$ in $\mathfrak{h}$.*

Note that the definition of an ideal is stronger than that of a subalgebra. For a subalgebra, we require only that the bracket of two elements of the subalgebra remain in the subalgebra. For an ideal, we require that the bracket of an element of the ideal with *any element of* $\mathfrak{g}$ be, again, in the ideal. Any Lie algebra $\mathfrak{g}$ has two "trivial" examples of ideals: $\mathfrak{g}$ itself and the zero ideal $\mathfrak{h} = \{0\}$.

Definition 6.2. *A complex Lie algebra $\mathfrak{g}$ is called* ***indecomposable*** *if the only ideals in $\mathfrak{g}$ are $\mathfrak{g}$ and $\{0\}$. A complex Lie algebra $\mathfrak{g}$ is called* ***simple*** *if $\mathfrak{g}$ is indecomposable and $\dim \mathfrak{g} \geq 2$.*

The term "indecomposable" is not a standard one, but since there does not seem to be any standard term for this concept, I have invented one. Note that the only indecomposable Lie algebras that are not simple are the one-dimensional ones and that any two one-dimensional Lie algebras are isomorphic, since all brackets must be zero. A one-dimensional Lie algebra has no nontrivial subalgebras and, hence, certainly no nontrivial ideals. Thus one-dimensional Lie algebras are indecomposable but not simple.

There is an analogy between finite-dimensional Lie algebras and finite groups. Subalgebras in the Lie algebra setting are the analogs of subgroups in the finite group setting, and ideals in the Lie algebra setting are the analogs of normal subgroups in the finite group setting. In this analogy, the one-dimensional Lie algebras (which are precisely the Lie algebras having no nontrivial subalgebras) are the analogs of the cyclic groups of prime order (which are precisely the groups having no nontrivial subgroups). However, there is a discrepancy in terminology: cyclic groups of prime order are called simple but one-dimensional Lie algebras are not called simple. This terminological convention is important to bear in mind in the following definition.

Definition 6.3. *A complex Lie algebra is called* ***reductive*** *if it is isomorphic to a direct sum of indecomposable Lie algebras. A complex Lie algebra is called* ***semisimple*** *if it isomorphic to a direct sum of simple Lie algebras.*

Note that a reductive Lie algebra is a direct sum of indecomposable algebras, which are either simple or one-dimensional commutative. Thus, a reductive Lie algebra is one that decomposes as a direct sum of a semisimple algebra (coming from the simple terms in the direct sum) and a commutative algebra (coming from the one-dimensional terms in the direct sum).

We will assume (in the spirit of this book) that the complex semisimple Lie algebras we study are given to us as subalgebras of some $\mathfrak{gl}(n;\mathbb{C})$. There is no loss of generality in this since by Ado's Theorem every finite-dimensional Lie algebra has a faithful finite-dimensional representation. In fact, for semisimple Lie algebras, the adjoint representation is always faithful, as is easily shown. (See Exercise 1.)

Proposition 6.4. *A complex Lie algebra $\mathfrak{g}$ is reductive precisely if the adjoint representation is completely reducible.*

Proof. An ideal is precisely an invariant subspace for the adjoint representation, as a moment's thought will confirm. So, if the adjoint representation decomposes as a direct sum of irreducibles, then $\mathfrak{g}$ decomposes (as a vector space) as $\mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_m$, where each $\mathfrak{g}_k$ is an ideal and where $\mathfrak{g}_k$ contains no ideals of $\mathfrak{g}$ other than $\mathfrak{g}_k$ itself and $\{0\}$. Now, if $X \in \mathfrak{g}_k$ and $Y \in \mathfrak{g}_l$ $(k \neq l)$, then $[X,Y] = 0$, since $[X,Y]$ must be in both $\mathfrak{g}_k$ and $\mathfrak{g}_l$ (because both $\mathfrak{g}_k$ and $\mathfrak{g}_l$ are ideals). This means that $\mathfrak{g}$ must be the direct sum (in the Lie algebra sense) of the $\mathfrak{g}_k$'s.

Now, I claim that each $\mathfrak{g}_k$ must be indecomposable when viewed as a Lie algebra in its own right. After all, suppose that $\mathfrak{h}$ is an ideal in $\mathfrak{g}_k$. Then, $\mathfrak{h}$ is also an ideal in $\mathfrak{g}$. (The commutator of an element of $\mathfrak{g}_k$ with an element $\mathfrak{h}$ will remain in $\mathfrak{h}$ by assumption. The commutator of an element of $\mathfrak{h}$ with an element of $\mathfrak{g}_l$, $l \neq k$, will be zero.) This means, by our assumptions on the $\mathfrak{g}_k$'s, that $\mathfrak{h} = \{0\}$ or $\mathfrak{h} = \mathfrak{g}_k$.

So, if the adjoint representation decomposes as a sum of irreducibles, then $\mathfrak{g}$ is reductive. Conversely, if $\mathfrak{g}$ is reductive, then $\mathfrak{g}$ is a direct sum of indecomposable algebras, which are then irreducible invariant subspaces for the adjoint representation. □

Corollary 6.5. *The complexification of the Lie algebra of a connected compact matrix Lie group is reductive.*

This follows from the above proposition and Proposition 4.36 (stating that connected compact groups have the complete reducibility property). Note that the Lie algebra of a compact Lie group may be only reductive and not semisimple. For example, the Lie algebra of S^1 is one dimensional and, thus, not semisimple.

Theorem 6.6. *A complex Lie algebra is semisimple if and only if it is isomorphic to the complexification of the Lie algebra of a simply-connected compact matrix Lie group.*

I will not prove this result. Nevertheless, let us discuss the ideas behind it. One direction is fairly easy, namely proving that if $\mathfrak{g}$ is the complexification of the Lie algebra of a compact simply-connected group K, then $\mathfrak{g}$ is semisimple. We have already shown that $\mathfrak{g}$ is reductive, even if K is not simply connected. Thus $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$, with $\mathfrak{g}_1$ semisimple and $\mathfrak{g}_2$ commutative. It can be shown that the Lie algebra $\mathfrak{k}$ of K decomposes as $\mathfrak{k} = \mathfrak{k}_1 \oplus \mathfrak{k}_2$, where $\mathfrak{g}_1 = \mathfrak{k}_1 + i\mathfrak{k}_1$ and $\mathfrak{g}_2 = \mathfrak{k}_2 + i\mathfrak{k}_2$. Then K decomposes as $K_1 \times K_2$, where K_1 and K_2 are simply connected and where K_2 is commutative. However, a simply-connected commutative Lie group is isomorphic to $\mathbb{R}^n$, which is noncompact for $n \geq 1$. Thus, the compactness of K means that $\mathfrak{k}_2 = \{0\}$, in which case $\mathfrak{g}_2 = \{0\}$ and $\mathfrak{g} = \mathfrak{g}_1$ is semisimple.

For the other direction, given a complex semisimple Lie algebra, we must find the correct real form whose corresponding simply-connected group is compact. For this, see Varadarajan (1974).

Definition 6.7. *If $\mathfrak{g}$ is a complex semisimple Lie algebra, then a* ***compact real form*** *of $\mathfrak{g}$ is real subalgebra $\mathfrak{k}$ of $\mathfrak{g}$ with the property that every X in $\mathfrak{g}$ can be written uniquely as $X = X_1 + iX_2$ with X_1 and X_2 in $\mathfrak{k}$ and such that there is a compact simply-connected matrix Lie group K_1 such that the Lie algebra $\mathfrak{k}_1$ of K_1 is isomorphic to $\mathfrak{k}$.*

Theorem 6.6 tells us that every complex semisimple Lie algebra has a compact real form. The compact real form is not unique, but it is "unique up to conjugation," as explained in Section 6.10.

Note that K itself is not necessarily simply connected. Consider, for example, the complex Lie algebra $\mathsf{so}(3;\mathbb{C}) \subset \mathsf{gl}(3;\mathbb{C})$ which is the complexification of $\mathsf{so}(3)$. We note that $\mathsf{so}(3)$ is isomorphic to $\mathsf{su}(2)$, which is the Lie algebra of the compact simply-connected group $\mathsf{SU}(2)$. This means that $\mathsf{so}(3;\mathbb{C})$ is semisimple and that $\mathsf{so}(3)$ is a compact real form of $\mathsf{so}(3;\mathbb{C})$. However, the subgroup of $\mathsf{GL}(3;\mathbb{C})$ whose Lie algebra is $\mathsf{so}(3)$ is the group $\mathsf{SO}(3)$, which is not simply connected. So, in this case, we have $K = \mathsf{SO}(3)$ and $K_1 = \mathsf{SU}(2)$.

Proposition 6.8. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra. If $\mathfrak{g}$ is a subalgebra of $\mathsf{gl}(n;\mathbb{C})$ and $\mathfrak{k}$ is a compact real form of $\mathfrak{g}$, then the connected Lie subgroup K of $\mathsf{GL}(n;\mathbb{C})$ whose Lie algebra is $\mathfrak{k}$ is compact.*

Proof. The definition of a compact real form is that there is a simply-connected compact matrix Lie group K_1 whose Lie algebra $\mathfrak{k}_1$ is isomorphic to $\mathfrak{k}$. Let $\phi : \mathfrak{k}_1 \to \mathfrak{k} \subset \mathsf{gl}(n;\mathbb{C})$ be a Lie algebra isomorphism. By Theorem 3.7, there is an associated Lie group homomorphism $\Phi : K_1 \to \mathsf{GL}(n;\mathbb{C})$, and let K be the image of this homomorphism. Since the image of a compact set under a continuous map is compact, K is compact (and hence closed). Furthermore, since the image of $\mathfrak{k}_1$ is $\mathfrak{k}$, Proposition 3.16 tells us that K is the connected Lie subgroup of $\mathsf{GL}(n;\mathbb{C})$ with Lie algebra $\mathfrak{k}$. □

A corollary of Theorem 6.6 is the following result, which we will make use of in the next chapter.

Corollary 6.9. *Every complex semisimple Lie algebra has the complete reducibility property.*

This holds because the representations of $\mathfrak{g}$ are in one-to-one correspondence with the representations of K_1, and compact groups have the complete reducibility property (Theorem 4.36). Actually, it is not hard to prove (Exercise 2) that among complex Lie algebras, *only* the semisimple ones have the complete reducibility property. Thus, complete reducibility is sometimes taken as the definition of semisimplicity for Lie algebras. For an algebraic proof of complete reducibility of semisimple Lie algebras, see Humphreys (1972).

Up to now, we have considered only *complex* semisimple Lie algebras, since these are the ones whose representations we will consider. (Working over $\mathbb{C}$ instead of $\mathbb{R}$ allows us to find nice bases for our Lie algebras.) Nevertheless, we can define the terms *ideal*, *indecomposable*, *simple*, *reductive*, and *semisimple* for real Lie algebras in precisely the same way as for the complex case.

Proposition 6.10. *If $\mathfrak{g}$ is a real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ its complexification, then $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g}_{\mathbb{C}}$ is semisimple.*

I will not prove this result—see Varadarajan (1974). Note that the proposition does *not* hold if the word "semisimple" is replaced by "simple." If $\mathfrak{g}_{\mathbb{C}}$ is simple, then $\mathfrak{g}$ must be simple (since if $\mathfrak{h}$ were a nontrivial ideal in $\mathfrak{g}$, then $\mathfrak{h}_{\mathbb{C}}$ would be a nontrivial ideal in $\mathfrak{g}_{\mathbb{C}}$), but the converse of this statement does not hold. For example, it can be shown that the six-dimensional real Lie algebra $\mathsf{so}(3,1)$ is simple. However, its complexification $\mathsf{so}(3,1;\mathbb{C})$ is isomorphic to $\mathsf{so}(4;\mathbb{C})$, which, in turn, is isomorphic to $\mathsf{sl}(2;\mathbb{C}) \oplus \mathsf{sl}(2;\mathbb{C})$, and so the complexification is not simple.

As a consequence of the above proposition and Theorem 6.6, we see that the real Lie algebra of a compact simply-connected group is semisimple. However, not every real semisimple Lie algebra is of this sort. Consider, for example, $\mathsf{sl}(n;\mathbb{R})$. Of course, $\mathsf{SL}(n;\mathbb{R})$ is noncompact, and there can be no compact simply-connected Lie group whose Lie algebra is $\mathsf{sl}(n;\mathbb{R})$, since such a group would then be the universal cover of $\mathsf{SL}(n;\mathbb{R})$, and the universal cover of a noncompact group is noncompact.

Nevertheless, $\mathsf{sl}(n;\mathbb{R})$ *is* semisimple, by Proposition 6.10, because its complexification is $\mathsf{sl}(n;\mathbb{C})$, which is also the complexification of $\mathsf{su}(n)$, which is the Lie algebra of a compact simply-connected group. So, a group G whose Lie algebra $\mathfrak{g}$ is real semisimple should be thought of as being "almost compact." This is to be understood not in any topological sense but rather in the sense that G has a compact simply-connected "cousin" K with the property that $\mathfrak{k}_{\mathbb{C}}$ is isomorphic to $\mathfrak{g}_{\mathbb{C}}$. For example, $\mathsf{SL}(n;\mathbb{R})$ has $\mathsf{SU}(n)$ as its cousin and $\mathsf{SO}(n,k)$ has $\mathsf{Spin}(n+k)$ (the simply-connected double cover of $\mathsf{SO}(n+k)$) as its cousin for $n+k \geq 3$.

When working with finite-dimensional representations, one can always extend the representation to the complexification, and so it is easier to work

only with complex semisimple Lie algebras. This will be our approach in the next chapter.

There are several other equivalent characterizations of semisimple Lie algebras, for example, that the Lie algebra have no nonzero solvable ideals or that the Killing form ($B(X, Y) = \text{trace}(\text{ad}_X \text{ad}_Y)$) be nondegenerate.

6.2 Examples of Reductive and Semisimple Lie Algebras

Let us consider some examples of Lie algebras that are reductive or semisimple, starting with the complex case. The following table lists the complex Lie algebras that we have encountered in this book that are either reductive or semisimple. An entry of "reductive" in the table means actually "reductive but not semisimple."

$\mathsf{sl}(n;\mathbb{C})$ $(n \geq 2)$	semisimple
$\mathsf{so}(n;\mathbb{C})$ $(n \geq 3)$	semisimple
$\mathsf{so}(2;\mathbb{C})$	reductive
$\mathsf{gl}(n;\mathbb{C})$ $(n \geq 1)$	reductive
$\mathsf{sp}(n;\mathbb{C})$ $(n \geq 1)$	semisimple

To verify the results of this table, we use Theorem 6.6. First, $\mathsf{sl}(n;\mathbb{C})$ is the complexification of $\mathsf{su}(n)$, which is the Lie algebra of the compact simply-connected group $\mathsf{SU}(n)$. Next, $\mathsf{so}(n;\mathbb{C})$ is the complexification of $\mathsf{so}(n)$, which is the Lie algebra of the compact group $\mathsf{SO}(n)$. Unfortunately, $\mathsf{SO}(n)$ is not simply connected. However, $\mathsf{so}(n)$ is also the Lie algebra of $\mathsf{Spin}(n)$, which is compact and simply connected for all $n \geq 3$ (Bröcker and tom Dieck (1985)). Meanwhile, $\mathsf{so}(2;\mathbb{C})$ is one-dimensional commutative and thus reductive but not semisimple.

Next, $\mathsf{gl}(n;\mathbb{C})$ is the complexification of $\mathsf{u}(n)$, which is the Lie algebra of the compact group $\mathsf{U}(n)$. This means that $\mathsf{gl}(n;\mathbb{C})$ is reductive. However, the center of a semisimple Lie algebra must be trivial, and the center of $\mathsf{gl}(n;\mathbb{C})$ is nontrivial, containing all the multiples of the identity. Note that $\mathsf{gl}(n;\mathbb{C}) \cong \mathsf{sl}(n;\mathbb{C}) \oplus \mathbb{C}$, where $\mathsf{sl}(n;\mathbb{C})$ is semisimple and $\mathbb{C}$ is one-dimensional commutative. So, $\mathsf{gl}(n;\mathbb{C})$ is reductive but not semisimple. Finally, $\mathsf{sp}(n;\mathbb{C})$ is the complexification of $\mathsf{sp}(n)$, which is the Lie algebra of the compact simply-connected group $\mathsf{Sp}(n)$.

All of the above-listed semisimple algebras are actually simple, except for $\mathsf{so}(4;\mathbb{C})$, which is isomorphic to $\mathsf{sl}(2;\mathbb{C}) \oplus \mathsf{sl}(2;\mathbb{C})$. In Chapter 8, we will discuss the classification of complex simple Lie algebras. It turns out that every complex simple Lie algebra is isomorphic to one of $\mathsf{sl}(n;\mathbb{C})$, $\mathsf{so}(n;\mathbb{C})$ $(n \neq 4)$, $\mathsf{sp}(n;\mathbb{C})$, or to one of five "exceptional" Lie algebras conventionally called G_2, F_4, E_6, E_7, and E_8.

We now make a table of the real Lie algebras we have encountered in this book that are either reductive or semisimple. Again, "reductive" means

actually "reductive but not semisimple." In each case, the complexification of the listed Lie algebra is isomorphic to one of the complex Lie algebras in the above table. Note that there can be several nonisomorphic real Lie algebras whose complexifications are isomorphic to the same complex semisimple Lie algebra.

$\mathsf{su}(n)$ $(n \geq 2)$	semisimple
$\mathsf{so}(n)$ $(n \geq 3)$	semisimple
$\mathsf{so}(2)$	reductive
$\mathsf{sp}(n)$ $(n \geq 1)$	semisimple
$\mathsf{so}(n,k)$ $(n+k \geq 3)$	semisimple
$\mathsf{so}(1,1)$	reductive
$\mathsf{sp}(n;\mathbb{R})$ $(n \geq 1)$	semisimple
$\mathsf{sl}(n;\mathbb{R})$ $(n \geq 2)$	semisimple
$\mathsf{gl}(n;\mathbb{R})$ $(n \geq 1)$	reductive

The other Lie algebras we have examined in this book, namely the Lie algebras of the Heisenberg group, the Euclidean group, and the Poincaré group, are neither reductive nor semisimple.

6.3 Cartan Subalgebras

We now begin to develop the structure that we will use (in the next chapter) in describing the representations of complex semisimple Lie algebras. These same structures are used to give a classification of semisimple Lie algebras, as discussed in Chapter 8. See Section 6.9 for how these structures come out in the case $\mathfrak{g} = \mathsf{sl}(n;\mathbb{C})$.

Definition 6.11. *If $\mathfrak{g}$ is a complex semisimple Lie algebra, then a **Cartan subalgebra** of $\mathfrak{g}$ is a complex subspace $\mathfrak{h}$ of $\mathfrak{g}$ with the following properties:*

1. *For all H_1 and H_2 in $\mathfrak{h}$, $[H_1, H_2] = 0$.*
2. *For all X in $\mathfrak{g}$, if $[H, X] = 0$ for all H in $\mathfrak{h}$, then X is in $\mathfrak{h}$.*
3. *For all H in $\mathfrak{h}$, ad_H is diagonalizable.*

Condition 1 says that $\mathfrak{h}$ is a commutative subalgebra of $\mathfrak{g}$. Condition 2 says that $\mathfrak{h}$ is a *maximal* commutative subalgebra (i.e., not contained in any larger commutative subalgebra). Condition 3 says that each ad_H $(H \in \mathfrak{h})$ is diagonalizable. Since the H's in $\mathfrak{h}$ commute, the ad_H's also commute, and thus they are *simultaneously* diagonalizable. (It is a standard result in linear algebra that any commuting family of diagonalizable matrices is simultaneously diagonalizable; see Section B.8.)

Of course, the definition of a Cartan subalgebra makes sense in any Lie algebra, semisimple or not. However, if $\mathfrak{g}$ is not semisimple, then $\mathfrak{g}$ may not have any Cartan subalgebras. (See Exercise 3.) Even in the semisimple case we must prove that a Cartan subalgebra exists.

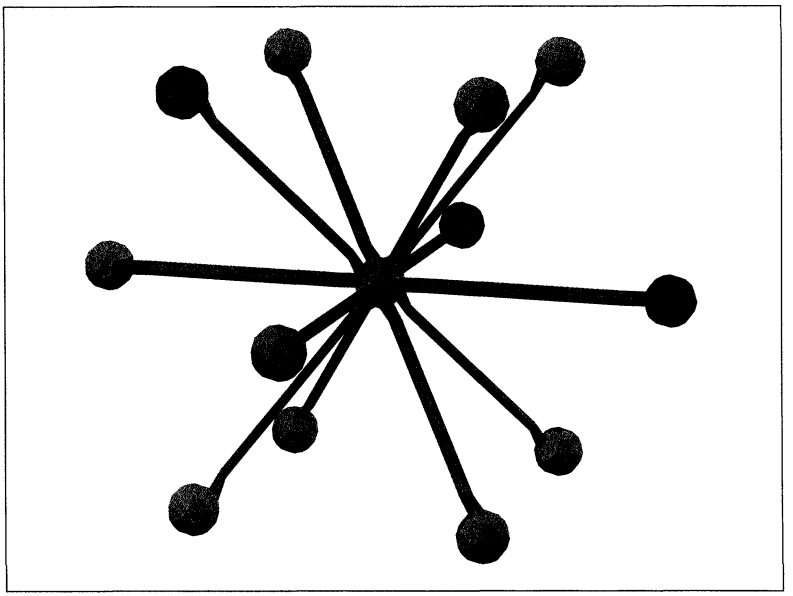


Plate 1: The A_3 root system

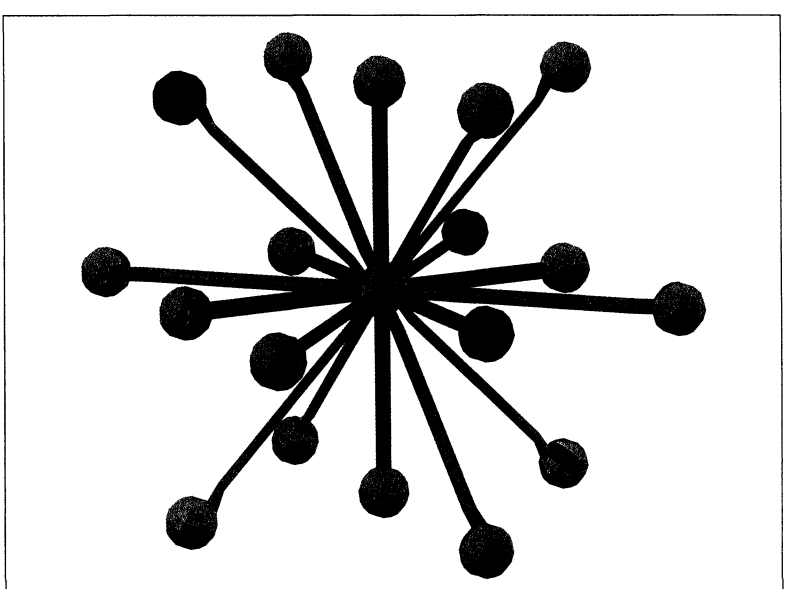


Plate 2: The B_3 root system

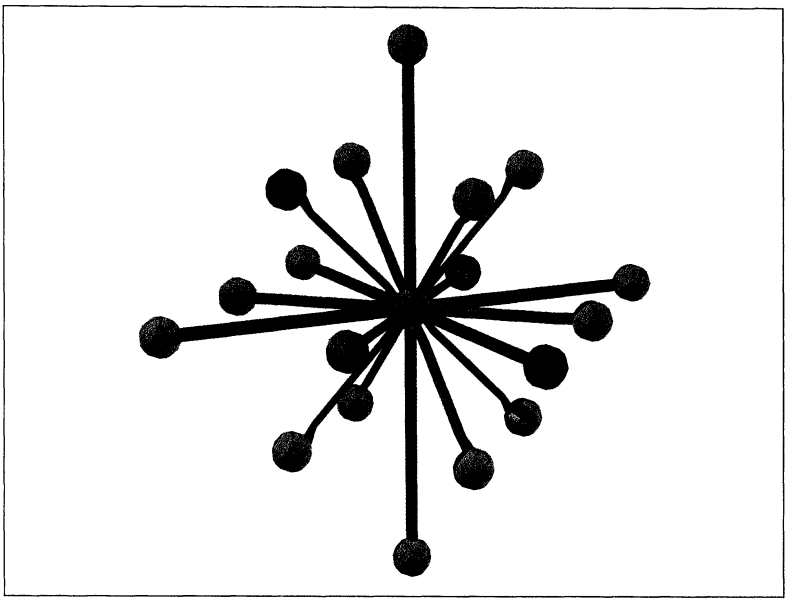


Plate 3: The C_3 root system

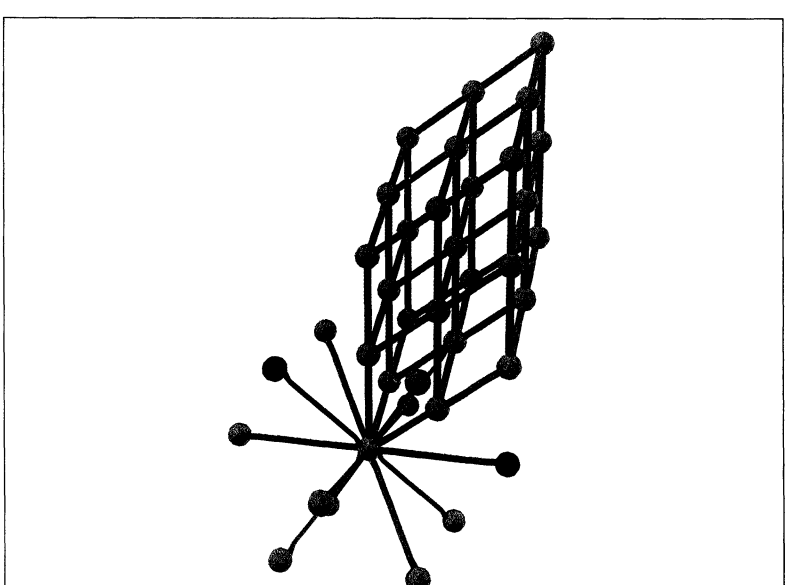


Plate 4: Dominant integral elements for A_3

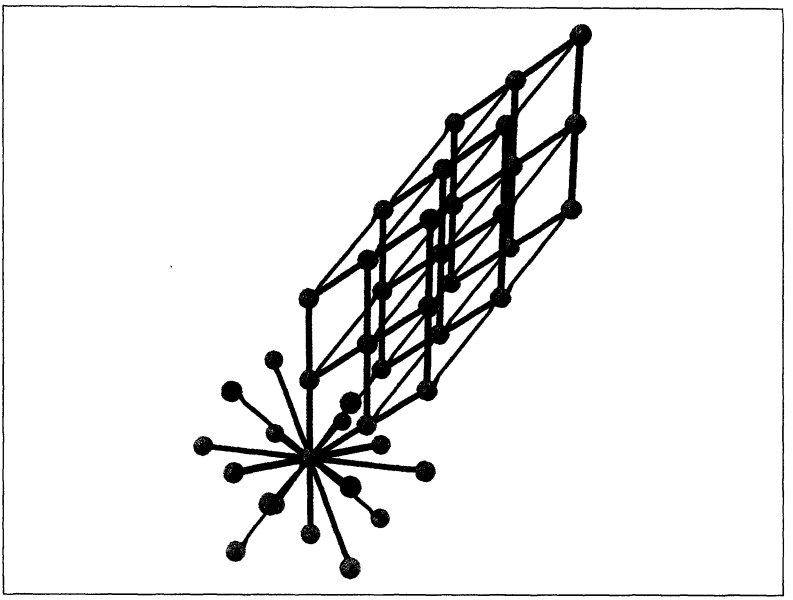


Plate 5: Dominant integral elements for B_3

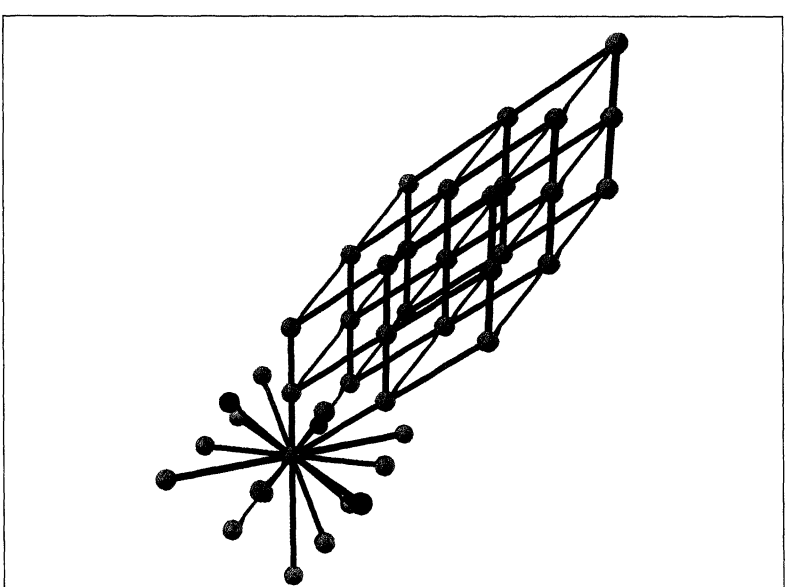


Plate 6: Dominant integral elements for C_3

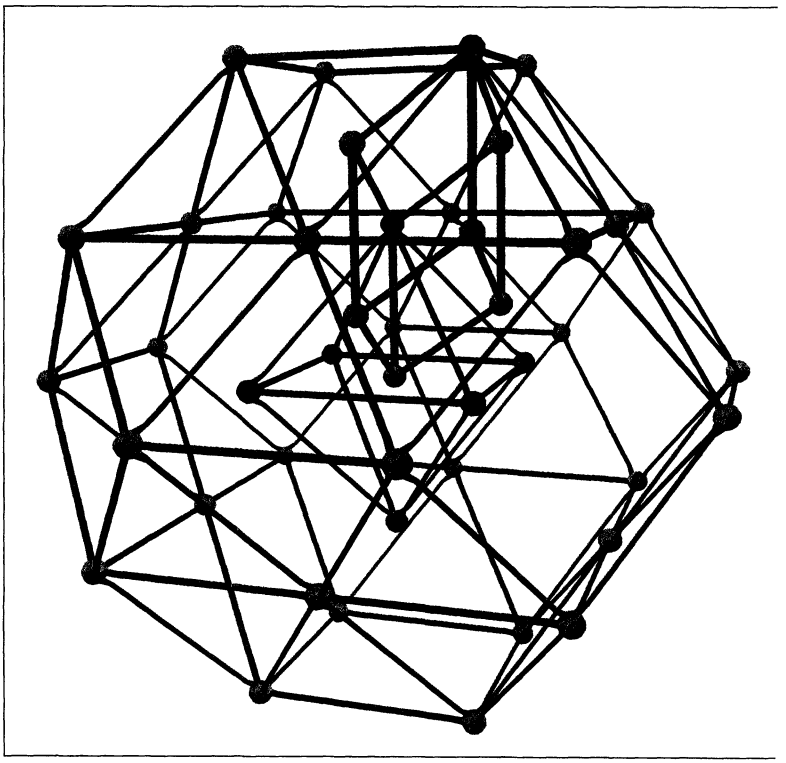


Plate 7: Representative weight diagram for A_3

Proposition 6.12. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra, let $\mathfrak{k}$ be a compact real form of $\mathfrak{g}$, and let $\mathfrak{t}$ be any maximal commutative subalgebra of $\mathfrak{k}$. Define $\mathfrak{h} \subset \mathfrak{g}$ to be $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$. Then, $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$.*

Note that $\mathfrak{k}$ (or any other Lie algebra) contains a maximal commutative subalgebra. After all, let $\mathfrak{t}_1$ be any one-dimensional subspace of $\mathfrak{k}$. Then, $\mathfrak{t}_1$ is a commutative subalgebra of $\mathfrak{k}$. If $\mathfrak{t}_1$ is maximal, then we are done; if not, then we chose some commutative subalgebra $\mathfrak{t}_2$ properly containing $\mathfrak{t}_1$. Then, if $\mathfrak{t}_2$ is maximal, we are done, and if not, we chose a commutative subalgebra $\mathfrak{t}_3$ properly containing $\mathfrak{t}_2$. Since $\mathfrak{k}$ is finite dimensional, this process cannot go on forever and we will eventually get a maximal commutative subalgebra.

Proof. It is clear that $\mathfrak{h}$ is a commutative subalgebra of $\mathfrak{g}$. We must first show that $\mathfrak{h}$ is *maximal* commutative. So, suppose that $X \in \mathfrak{g}$ commutes with every element of $\mathfrak{h}$, which certainly means that it commutes with every element of $\mathfrak{t}$. Then, write $X = X_1 + iX_2$ with X_1 and X_2 in $\mathfrak{k}$. Then, for H in $\mathfrak{t}$, we have

$$[H, X_1 + iX_2] = [H, X_1] + i[H, X_2] = 0,$$

where $[H, X_1]$ and $[H, X_2]$ are in $\mathfrak{k}$ (since $\mathfrak{k}$ is a real subalgebra). However, since every element of $\mathfrak{g}$ has a *unique* decomposition as an element of $\mathfrak{k}$ plus an element of $i\mathfrak{k}$, we see that $[H, X_1]$ and $[H, X_2]$ must separately be zero. Since this holds for all H in $\mathfrak{t}$ and since $\mathfrak{t}$ is maximal commutative, we must have X_1 and X_2 in $\mathfrak{t}$, which means that $X = X_1 + iX_2$ is in $\mathfrak{h}$. This shows that $\mathfrak{h}$ is maximal commutative.

We assume, as usual, that $\mathfrak{g}$ is given as a subalgebra of $\mathfrak{gl}(n;\mathbb{C})$. Let K be the subgroup of $\mathsf{GL}(n;\mathbb{C})$ whose Lie algebra is $\mathfrak{k}$. According to Proposition 6.8, K is compact. So, by the averaging method of Section 4.10, there exists a real-valued inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{k}$ that is invariant under the adjoint action of K. This inner product can then be extended to a complex-valued inner product on $\mathfrak{g}$, also denoted $\langle \cdot, \cdot \rangle$, that is invariant under the adjoint action of K and that takes real values on $\mathfrak{k}$. This means that for each A in K, Ad_A is a unitary operator on $\mathfrak{g}$ (with respect to the inner product $\langle \cdot, \cdot \rangle$). It then follows that for each X in $\mathfrak{k}$, $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is skew self-adjoint. (This is by the same argument that shows that the Lie algebra of $\mathsf{U}(n)$ consists of skew-self-adjoint matrices.) Thus, in particular, ad_H is skew for each H in $\mathfrak{t}$, and a skew operator on a finite-dimensional complex inner product space is automatically diagonalizable. (See Appendix B.)

Finally, if H is any element of $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$, then $H = H_1 + iH_2$, with H_1 and H_2 in $\mathfrak{t}$. Since H_1 and H_2 commute, ad_{H_1} and ad_{H_2} also commute, and, therefore, ad_H will be diagonalizable (since a linear combination of commuting diagonalizable operators is diagonalizable). This shows that $\mathfrak{h}$ is a Cartan subalgebra. □

It is possible to prove that every Cartan subalgebra of $\mathfrak{g}$ arises as in Proposition 6.12 (for some compact real form $\mathfrak{k}$ and some maximal commutative

subalgebra $\mathfrak{t}$ of $\mathfrak{k}$) and also that Cartan subalgebras are "unique up to conjugation." (See Section 6.10 for more precise statements.) In particular, all Cartan subalgebras of a given complex semisimple Lie algebra have the same dimension. In light of this result, the following definition makes sense.

Definition 6.13. *If $\mathfrak{g}$ is a complex semisimple Lie algebra, then the* ***rank*** *of $\mathfrak{g}$ is the dimension of any Cartan subalgebra.*

6.4 Roots and Root Spaces

From now on we assume that we have chosen a compact real form $\mathfrak{k}$ of $\mathfrak{g}$ and a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$, and we consider the Cartan subalgebra $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$. We assume also that we have chosen (as in the proof of Proposition 6.12) an inner product on $\mathfrak{g}$ that is invariant under the adjoint action of K and that takes real values on $\mathfrak{k}$.

Definition 6.14. *A* ***root*** *of $\mathfrak{g}$ (relative to the Cartan subalgebra $\mathfrak{h}$) is a nonzero linear functional α on $\mathfrak{h}$ such that there exists a nonzero element X of $\mathfrak{g}$ with*

$$[H, X] = \alpha(H)X$$

for all H in $\mathfrak{h}$.

The set of all roots is denoted R.

The condition on X says that X is an eigenvector for each ad_H, with eigenvalue $\alpha(H)$. Note that if X is actually an eigenvector for each ad_H with H in $\mathfrak{h}$, then the eigenvalues must depend linearly on H. That is why we insist that α be a linear functional on $\mathfrak{h}$. (See Section B.8.) So, a root is just a (nonzero) collection of simultaneous eigenvalues for the ad_H's. Note any element of $\mathfrak{h}$ is a simultaneous eigenvector for all the ad_H's, with all eigenvalues equal to zero, but we only call α a root if α is nonzero. Of course, for any root α, *some* of the $\alpha(H)$'s may be equal to zero; we just require that not all of them be zero.

Proposition 6.15. *If α is a root, $\alpha(H)$ is imaginary for all H in $\mathfrak{t}$.*

Proof. By the proof of Proposition 6.12, there exists an inner product on $\mathfrak{g}$ such that ad_H is skew self-adjoint for all H in $\mathfrak{t}$. The eigenvalues of a skew operator are necessarily imaginary and each $\alpha(H)$ is an eigenvalue for ad_H. □

Note that the set of linear functionals on $\mathfrak{h}$ that are imaginary on $\mathfrak{t}$ forms a *real* vector space whose real dimension equals the complex dimension of $\mathfrak{h}$. If $\mathfrak{t}^*$ denotes the space of real-valued linear functionals on $\mathfrak{t}$, then the roots are contained in $i\mathfrak{t}^* \subset \mathfrak{h}^*$.

Definition 6.16. *If α is a root, then the* ***root space*** $\mathfrak{g}_\alpha$ *is the space of all X in $\mathfrak{g}$ for which $[H, X] = \alpha(H)X$ for all H in $\mathfrak{h}$. An element of $\mathfrak{g}_\alpha$ is called a* ***root vector*** *(for the root α).*

More generally, if α is any element of $\mathfrak{h}^*$, we define $\mathfrak{g}_\alpha$ to be the space of all X in $\mathfrak{g}$ for which $[H, X] = \alpha(H)X$ for all H in $\mathfrak{h}$ (but we do not call $\mathfrak{g}_\alpha$ a root space unless α is actually a root).

Taking $\alpha = 0$, we see that $\mathfrak{g}_0$ is the set of all elements of $\mathfrak{g}$ that commute with every element of $\mathfrak{h}$. Since $\mathfrak{h}$ is a maximal commutative subalgebra, we conclude that $\mathfrak{g}_0 = \mathfrak{h}$. If α is not zero and not a root, then $\mathfrak{g}_\alpha = \{0\}$.

Now, since $\mathfrak{h}$ is commutative, the operators ad_H, $H \in \mathfrak{h}$, all commute. Furthermore, by the definition of a Cartan subalgebra, each ad_H, $H \in \mathfrak{h}$, is diagonalizable. It follows (Proposition B.13) that the ad_H's, $H \in \mathfrak{h}$, are simultaneously diagonalizable. As a result, $\mathfrak{g}$ can be decomposed as the direct sum of $\mathfrak{h}$ and the root spaces $\mathfrak{g}_\alpha$. (Here, we make use of Proposition B.14.) Thus, we have established the following.

Proposition 6.17. *The Lie algebra $\mathfrak{g}$ can be decomposed as a direct sum as follows:*

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha.$$

This means that every element of $\mathfrak{g}$ can be written uniquely as a sum of an element of $\mathfrak{h}$ and one element from each root space $\mathfrak{g}_\alpha$.

Proposition 6.18. *For any α and β in $\mathfrak{h}^*$, $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.*

More explicitly, this means that if X is in $\mathfrak{g}_\alpha$ and Y is in $\mathfrak{g}_\beta$, then $[X, Y]$ is in $\mathfrak{g}_{\alpha+\beta}$. In particular, if X is in $\mathfrak{g}_\alpha$ and Y is in $\mathfrak{g}_{-\alpha}$, then $[X, Y]$ is in $\mathfrak{h}$. Furthermore, if X is in $\mathfrak{g}_\alpha$, Y is in $\mathfrak{g}_\beta$, and $\alpha + \beta$ is neither zero nor a root, then $[X, Y] = 0$.

Proof. We use that ad_H is a derivation, which means that $[H, [X, Y]] = [[H, X], Y] + [X, [H, Y]]$. This identity is equivalent to the Jacobi identity (Exercise 22 from Chapter 2). So, if X is in $\mathfrak{g}_\alpha$ and Y is in $\mathfrak{g}_\beta$ then we have for all H in $\mathfrak{h}$,

$$\begin{aligned}[H, [X, Y]] &= [\alpha(H)X, Y] + [X, \beta(H)Y] \\ &= (\alpha(H) + \beta(H))[X, Y].\end{aligned}$$

This shows that $[X, Y]$ is in $\mathfrak{g}_{\alpha+\beta}$. □

Proposition 6.19.

1. *If $\alpha \in \mathfrak{h}^*$ is a root, then so is $-\alpha$.*
2. *The roots span $\mathfrak{h}^*$.*

Proof. Since α is a root, there exists a nonzero element X of $\mathfrak{g}$ such that $[H, X] = \alpha(H)X$ for all H in $\mathfrak{h}$ and thus, in particular, for all H in $\mathfrak{t}$. Then, X can be written uniquely as $X = X_1 + iX_2$ with X_1 and X_2 in $\mathfrak{k}$. So, for H in $\mathfrak{t}$, we have

$$[H, X] = [H, X_1] + i[H, X_2],$$

where $[H, X_1]$ and $[H, X_2]$ are in $\mathfrak{k}$, since $\mathfrak{k}$ is a real subalgebra. Recall that $\alpha(H)$ is imaginary for H in $\mathfrak{t}$. So, write $\alpha(H) = ia$, with a real. Then,

$$[H, X] = iaX = -aX_2 + iaX_1.$$

Since each element of $\mathfrak{g}$ has a *unique* decomposition as a sum of an element of $\mathfrak{k}$ and an element $i\mathfrak{k}$, we must have $[H, X_1] = -aX_2$ and $[H, X_2] = aX_1$.

Now, put $Y = X_1 - iX_2$. Then,

$$\begin{aligned}
[H, Y] &= [H, X_1] - i[H, X_2] \\
&= -aX_2 - iaX_1 \\
&= -ia(X_1 - iX_2) \\
&= -iaY.
\end{aligned}$$

Thus, $[H, Y] = -\alpha(H)Y$ for all H in $\mathfrak{t}$, and thus also for all H in $\mathfrak{h}$. This shows that $-\alpha$ is another root.

For Point 2, suppose that the roots did not span $\mathfrak{h}^*$. Then, there would exist a nonzero $H \in \mathfrak{h}$ such that $\alpha(H) = 0$ for all $\alpha \in R$. Then, $[H, H_1] = 0$ for all H_1 in $\mathfrak{h}$, and also $[H, X] = \alpha(H)X = 0$ for X in $\mathfrak{g}_\alpha$. Thus, by Proposition 6.17, H would commute with all elements of $\mathfrak{g}$; that is, H would be in the center of $\mathfrak{g}$. However, the center of a semisimple Lie algebra must be zero (Exercise 1), and we have a contradiction. □

We now come to the first substantial result about roots and root spaces, the proof of which will occupy the remainder of this section.

Theorem 6.20.

1. *If α is a root, then the only multiples of α that are roots are α and $-\alpha$.*
2. *If α is a root, then the root space $\mathfrak{g}_\alpha$ is one dimensional.*
3. *For each root α, we can find nonzero elements X_α in $\mathfrak{g}_\alpha$, Y_α in $\mathfrak{g}_{-\alpha}$, and H_α in $\mathfrak{h}$ such that*

$$\begin{aligned}
[H_\alpha, X_\alpha] &= 2X_\alpha, \\
[H_\alpha, Y_\alpha] &= -2Y_\alpha, \\
[X_\alpha, Y_\alpha] &= H_\alpha.
\end{aligned}$$

The element H_α is unique (i.e., independent of the choice of X_α and Y_α).

Point 3 of the theorem tells us that X_α, Y_α, and H_α span a subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}(2;\mathbb{C})$. The elements H_α of $\mathfrak{h}$ given in Point 3 of the theorem are called the **co-roots**. Their properties are closely related to the properties of the roots themselves.

In preparation for the proof of Theorem 6.20, we choose (as in the proof of Proposition 6.12) an inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$ that is invariant under the adjoint action of K and that takes real values on $\mathfrak{k}$. This means that for each X in $\mathfrak{k}$, ad_X is skew; that is,

$$\langle \mathrm{ad}_X Y, Z \rangle = - \langle Y, \mathrm{ad}_X Z \rangle \tag{6.1}$$

for all X in $\mathfrak{k}$ and all Y and Z in $\mathfrak{g}$. Now, suppose X is any element of $\mathfrak{g}$ with $X = X_1 + iX_2$ ($X_1, X_2 \in \mathfrak{k}$) and define

$$X^* = -X_1 + iX_2. \tag{6.2}$$

The motivation for this definition is that if $\mathfrak{g} = \mathfrak{sl}(n;\mathbb{C})$ and $\mathfrak{k} = \mathfrak{su}(n)$, then X^* is the usual adjoint of X.

It follows from (6.1) that for any X in $\mathfrak{g}$ (not necessarily in $\mathfrak{k}$), we have

$$\langle \mathrm{ad}_X Y, Z \rangle = \langle Y, \mathrm{ad}_{X^*} Z \rangle . \tag{6.3}$$

Furthermore, looking at the proof of Proposition 6.19, we see that if X is in the root space $\mathfrak{g}_\alpha$, then X^* is in the root space $\mathfrak{g}_{-\alpha}$. (The element Y in the proof of that proposition is $-X^*$.)

We now come to a simple but crucial calculation.

Lemma 6.21. *Suppose that X is in $\mathfrak{g}_\alpha$, Y is in $\mathfrak{g}_{-\alpha}$, and H is in $\mathfrak{h}$. Then, $[X,Y]$ is in $\mathfrak{h}$ and*

$$\langle [X,Y], H \rangle = \alpha(H) \langle Y, X^* \rangle ,$$

where X^* is defined by (6.2).

Proof. That $[X,Y]$ is in $\mathfrak{h}$ follows from Proposition 6.18. Then, using (6.3), we compute that

$$\begin{aligned} \langle [X,Y], H \rangle &= \langle \mathrm{ad}_X Y, H \rangle = \langle Y, \mathrm{ad}_{X^*} H \rangle \\ &= \langle Y, [X^*, H] \rangle = - \langle Y, [H, X^*] \rangle . \end{aligned} \tag{6.4}$$

However, since X is in $\mathfrak{g}_\alpha$, X^* is in $\mathfrak{g}_{-\alpha}$, and, so, (6.4) becomes

$$\langle [X,Y], H \rangle = - \langle Y, -\alpha(H) X^* \rangle = \alpha(H) \langle Y, X^* \rangle .$$

Recall that we take the inner product to have the complex conjugate in the first factor. □

With this lemma established, we turn to the proof of Theorem 6.20.

Proof. The proof will be in several steps. Throughout the proof α will be a fixed root.

Step 1. If X is in $\mathfrak{g}_\alpha$ and Y is in $\mathfrak{g}_{-\alpha}$, then $[X,Y]$ is in $\mathfrak{h}$ and $[X,Y]$ is orthogonal to all elements of $\mathfrak{h}$ for which $\alpha(H)$ is zero.

This is an immediate consequence of Lemma 6.21.

Let $\ker\alpha$ denote the space of all H in $\mathfrak{h}$ for which $\alpha(H)=0$ and let $(\ker\alpha)^\perp$ denote the orthogonal complement of $\ker\alpha$ in $\mathfrak{h}$. Then, if $\mathfrak{h}$ has dimension r, $\ker\alpha$ will have dimension $r-1$ and $(\ker\alpha)^\perp$ will have dimension 1. Thus, Step 1 is telling us that $[\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}]$ is contained in the *one-dimensional* subspace $(\ker\alpha)^\perp$ of $\mathfrak{h}$. This result will be important for us as we continue with the proof of Theorem 6.20.

Step 2. Let X be a nonzero element of $\mathfrak{g}_\alpha$, so that X^* is a nonzero element of $\mathfrak{g}_{-\alpha}$. Then, $[X,X^*]\neq 0$ and $\alpha([X,X^*])$ is real and strictly positive.

To see that $[X,X^*]$ is not zero, we apply Lemma 6.21 with $Y=X^*$ and with H any element of $\mathfrak{h}$ for which $\alpha(H)\neq 0$. The lemma then shows that $\langle[X,Y],H\rangle\neq 0$, which means that $[X,Y]\neq 0$.

Now apply Lemma 6.21 with $Y=X^*$ and with $H=[X,X^*]$. This gives

$$\langle[X,X^*],[X,X^*]\rangle=\alpha([X,X^*])\,\langle X,X^*\rangle\,.$$

From the positivity of the inner product and the fact that $[X,X^*]$ is nonzero, we conclude that $\alpha([X,X^*])$ is real and strictly positive.

Step 3. We can choose nonzero elements $X_\alpha\in\mathfrak{g}_\alpha$, $Y_\alpha\in\mathfrak{g}_{-\alpha}$, and $H_\alpha\in\mathfrak{h}$ such that $[H,X_\alpha]=2X_\alpha$, $[H,Y_\alpha]=-2Y_\alpha$, and $[X_\alpha,Y_\alpha]=H_\alpha$.

We initially take X to be any nonzero element of $\mathfrak{g}_\alpha$, Y to be X^*, and H to be $[X,Y]=[X,X^*]$. Then, $[H,X]=\alpha(H)X$, where $\alpha(H)=\alpha([X,X^*])>0$, and $[H,Y]=-\alpha(H)Y$. We now set

$$H_\alpha=\frac{2}{\alpha(H)}H,$$

$$X_\alpha=\sqrt{\frac{2}{\alpha(H)}}X,$$

$$Y_\alpha=\sqrt{\frac{2}{\alpha(H)}}Y.$$

Direct calculation (check!) then shows that these elements have the required commutation relations.

Note that, on the one hand, $[H_\alpha,X_\alpha]=\alpha(H_\alpha)X_\alpha$ and, on the other hand, $[H_\alpha,X_\alpha]=2X_\alpha$. So, evidently, $\alpha(H_\alpha)=2$. Note also that we have chosen X_α and Y_α in such a way that $Y_\alpha=X_\alpha^*$.

Step 4. If β is a root of the form $\beta=k\alpha$ for some constant k, then k is an integer multiple of $\frac{1}{2}$.

We let $\mathfrak{s}^\alpha$ be the subalgebra of $\mathfrak{g}$ given by

$$\mathfrak{s}^\alpha = \text{span}\{X_\alpha, Y_\alpha, H_\alpha\},$$

which is isomorphic to $\mathsf{sl}(2;\mathbb{C})$. We then let V^α be the subspace of $\mathfrak{g}$ spanned by (1) the subspace $(\ker\alpha)^\perp \subset \mathfrak{h}$ and (2) the root spaces $\mathfrak{g}_\beta$ for which β is a multiple of α. Recall that $(\ker\alpha)^\perp$ is one dimensional and that H_α is a nonzero element of this space.

I claim that V^α is invariant under the adjoint action of $\mathfrak{s}^\alpha$. To see this, we first show that V^α is invariant under ad_{H_α}, which is clear since both $(\ker\alpha)^\perp$ and the $\mathfrak{g}_\beta$'s are eigenspaces for ad_{H_α}. Then, we show that V^α is invariant under ad_{X_α}. By Proposition 6.18, for Y in $\mathfrak{g}_\beta$, $\text{ad}_{X_\alpha}Y$ will be in $\mathfrak{g}_{\alpha+\beta}$. If $\alpha+\beta$ is not zero, then $\mathfrak{g}_{\alpha+\beta}$ is either zero or another root space with root a multiple of α. If $\alpha+\beta$ is zero, then $\text{ad}_{X_\alpha}Y$ will be in $(\ker\alpha)^\perp$, by Step 1, and, thus, $\text{ad}_{X_\alpha}Y$ is in V^α. A similar argument shows that V^α is invariant under ad_{Y_α}.

Thus, V^α is invariant under the adjoint action of $\mathfrak{s}^\alpha$, which means that V^α is a representation of $\mathfrak{s}^\alpha$. Now, we know that in any finite-dimensional representation of $\mathfrak{s}^\alpha \cong \mathsf{sl}(2;\mathbb{C})$, the eigenvalues for H_α must be integers. What eigenvalues of ad_{H_α} arise in V^α? Recall that $\alpha(H_\alpha) = 2$. Then, for $Y \in \mathfrak{g}_\beta$ with $\beta = k\alpha$, we will have $\text{ad}_{H_\alpha}Y = \beta(H_\alpha)Y = k\alpha(H_\alpha)Y = 2kY$. Thus, $2k$ must be an integer, which is what we are trying to show.

Step 5. If α is a root, then 2α is not a root.

We use the complete reducibility of representations of $\mathfrak{s}^\alpha \cong \mathsf{sl}(2;\mathbb{C})$ (which comes from the compactness of $\mathsf{SU}(2)$). Note that $\mathfrak{s}^\alpha$ itself is an (irreducible) invariant subspace for the adjoint action of $\mathfrak{s}^\alpha$. By complete reducibility and Proposition 4.33, V^α decomposes as a direct sum of $\mathfrak{s}^\alpha$ and several other irreducible invariant subspaces $U_1, \ldots, U_m$.

Recall that $\alpha(H_\alpha) = 2$. If, then, there were a nonzero element of $\mathfrak{g}_{2\alpha}$, it would be an eigenvector for ad_{H_α} with eigenvalue $2\alpha(H_\alpha) = 4$. This means that the eigenvalue 4 for ad_{H_α} would have to arise in one of the U_k's. (Since $\mathfrak{s}^\alpha$ and each of the U_k's is invariant, if 4 is to be an eigenvalue of ad_{H_α} then it must be an eigenvalue in one of these spaces, and it is not an eigenvalue in $\mathfrak{s}^\alpha$.) However, by what we know of the representation theory of $\mathsf{sl}(2;\mathbb{C})$, if 4 is an eigenvalue of ad_{H_α} in U_k, then 0 is also an eigenvalue of ad_{H_α} in U_k. (After all, the eigenvalues in any irreducible representation go from some maximum value of m to $-m$ in increments of 2, and m must be even in order for 4 to occur, in which case 0 must also occur.)

This means that we must have a nonzero vector Z in some $U_k \subset V^\alpha$ with $\text{ad}_{H_\alpha}Z = 0$. The vector Z must be in $(\ker\alpha)^\perp \subset \mathfrak{h}$, because V^α is the direct sum of $(\ker\alpha)^\perp$ and eigenspaces for ad_{H_α} with nonzero eigenvalues (namely the $\mathfrak{g}_\beta$'s with β a multiple of α). However, we have already established that $(\ker\alpha)^\perp$ is one dimensional and that H_α is a nonzero element of this space. Thus, Z is a nonzero multiple of H_α, which means that $\mathfrak{s}^\alpha$ and U_k have a nonzero intersection. This is a contradiction, since V^α is the direct sum of $\mathfrak{s}^\alpha$ and the various U_k's.

Step 6. The only multiples of α that are roots are α and $-\alpha$.

Suppose $\beta = k\alpha$ is a root (with $k \neq 0$, since roots by definition are nonzero). By Step 4, k must be an integer multiple of $\frac{1}{2}$, and by Step 5, $k \neq 2$. We may assume $k > 0$ (if not, replace β by $-\beta$), in which case, $k = 1/2$, 1, 3/2, 5/2, 6/2, ... However, now $\beta = \frac{1}{k}\alpha$, and precisely the same arguments apply with α replaced by β, and, so, $1/k$ must also be an integer multiple of $1/2$, with $1/k \neq 2$. Of the possible values for k, the only one for which $1/k$ has these properties is $k = 1$.

Step 7. The root spaces $\mathfrak{g}_\alpha$ are one dimensional.

Suppose not. Then there exists another element X' in $\mathfrak{g}_\alpha$ that is linearly independent of X_α. In that case, X' is an eigenvector for ad_{H_α} with eigenvalue $\alpha(H_\alpha) = 2$. However, reasoning as in Step 5, if there is another eigenvector in V^α for ad_{H_α} with eigenvalue 2 (independent of X_α), then there must also be another eigenvector in V^α for ad_{H_α} with eigenvalue 0, which we have seen is impossible since the intersection of V^α with $\mathfrak{h}$ is one dimensional.

It remains only to show the uniqueness of the elements H_α. Since the root spaces $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are one dimensional, H_α is certainly unique up to a constant. However, this constant is determined by the condition that $[H_\alpha, X_\alpha] = 2X_\alpha$, which is independent of the normalization of X_α since both sides are linear in X_α. To say the same thing in a different way, the normalization of H_α is determined by the condition $\alpha(H_\alpha) = 2$. This concludes the proof of Theorem 6.20. □

6.5 Inner Products of Roots and Co-roots

We continue with the setting of the previous section: $\mathfrak{g}$ is a complex semisimple Lie algebra, $\mathfrak{k}$ is a compact real form of $\mathfrak{g}$, $\mathfrak{t}$ is a maximal commutative subalgebra of $\mathfrak{k}$, and $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$ is the associated Cartan subalgebra. We choose, once and for all, an inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$ that is invariant under the adjoint action of K and takes real values on $\mathfrak{k}$, and we consider the restriction of this inner product to $\mathfrak{h}$.

This section will give a geometric picture of the roots, which were treated algebraically in the previous section.

Proposition 6.22. *Suppose α and β are roots and H_α is the co-root associated to root α. Then, $\beta(H_\alpha)$ is an integer.*

Proof. We once again let $\mathfrak{s}^\alpha = \mathrm{span}\{X_\alpha, Y_\alpha, H_\alpha\}$ be the subalgebra of $\mathfrak{g}$ (isomorphic to $\mathfrak{sl}(2;\mathbb{C})$) given by Point 3 of Theorem 6.20. Then, $\mathfrak{s}^\alpha$ acts on $\mathfrak{g}$ by the adjoint action, so that $\mathfrak{g}$ becomes a finite-dimensional representation of $\mathfrak{s}^\alpha$. From our knowledge of the representations of $\mathfrak{sl}(2;\mathbb{C})$, we know that any eigenvalue of ad_{H_α} must be an integer. If X_β is any nonzero element of $\mathfrak{g}_\beta$, then $[H_\alpha, X_\beta] = \beta(H_\alpha)X_\beta$, so $\beta(H_\alpha)$ is an eigenvalue for ad_{H_α}, which must then be an integer. □

Recall that the roots α are elements of the dual space $\mathfrak{h}^*$, whereas the co-roots H_α (defined in Theorem 6.20) are elements of $\mathfrak{h}$ itself. We are going to use the inner product on $\mathfrak{h}$ to identify $\mathfrak{h}^*$ with $\mathfrak{h}$ and thus putting the roots and the co-roots into the same space and giving a more geometric picture of the roots and of the integrality condition in Proposition 6.22. We make use of the following elementary result from Section B.7.

Proposition 6.23. *Given any linear functional $\alpha \in \mathfrak{h}^*$ (not necessarily a root), there exists a unique element H^α in $\mathfrak{h}$ such that*

$$\alpha(H) = \langle H^\alpha, H \rangle$$

for all H in $\mathfrak{h}$.

Observe that the notation is H^α (not H_α). Recall that we take the inner product to be linear in the second factor. The map $\alpha \to H^\alpha$ is a one-to-one and onto correspondence between $\mathfrak{h}^*$ and $\mathfrak{h}$. However, this correspondence is not linear but rather conjugate-linear, since the inner product is conjugate-linear in the first factor (where H^α is).

It is convenient to permanently identify each root $\alpha \in \mathfrak{h}^*$ with the corresponding element $H^\alpha \in \mathfrak{h}$. Having done this, we then omit the H^α notation and denote that element of $\mathfrak{h}$ simply as α.

Notation 6.24 *From now on, we identify each root with the corresponding element of $\mathfrak{h}$ given by Proposition 6.23. Thus, we now regard a root α as a nonzero element of $\mathfrak{h}$ (not $\mathfrak{h}^*$) with the property that there exists a nonzero X in $\mathfrak{g}$ with*

$$[H, X] = \langle \alpha, H \rangle X$$

for all $H \in \mathfrak{h}$.

This notational change means that we now write $\langle \alpha, H \rangle$ every time that $\alpha(H)$ occurs in the previous section. So, for example, the assertion that $\beta(H_\alpha)$ is an integer now becomes the assertion that $\langle \beta, H_\alpha \rangle$ is an integer. The basic properties of the roots (Propositions 6.15 and 6.19) can now be translated into our new notation as follows.

Proposition 6.25. *Let $R \subset \mathfrak{h}$ be the set of roots in the sense of Notation 6.24.*

1. *Each root $\alpha \in R$ is contained in it $\subset \mathfrak{h}$.*
2. *The roots span $\mathfrak{h}$.*
3. *If α is a root, then so is $-\alpha$.*

Proof. Points 2 and 3 follow from Proposition 6.19 under the identification of $\mathfrak{h}^*$ with $\mathfrak{h}$. (Although the correspondence between $\mathfrak{h}^*$ and $\mathfrak{h}$ is conjugate-linear rather than linear, it still takes spanning sets in $\mathfrak{h}^*$ to spanning sets in $\mathfrak{h}$.) For Point 1, we note that Proposition 6.15, translated into our new notation,

says that $\langle \alpha, H \rangle$ is imaginary for all roots α and all H in $\mathfrak{t}$. We now write $\alpha = \alpha_1 + i\alpha_2$ with α_1 and α_2 in $\mathfrak{t}$. Then, $\langle \alpha, \alpha_1 \rangle$ must be imaginary, but also $\langle \alpha, \alpha_1 \rangle = \langle \alpha_1, \alpha_1 \rangle - i \langle \alpha_2, \alpha_1 \rangle$. Since the inner product $\langle \cdot, \cdot \rangle$ is real on $\mathfrak{k}$ and, hence, on $\mathfrak{t}$, we must then have $\alpha_1 = 0$ or else $\langle \alpha, \alpha_1 \rangle$ would not be imaginary. □

Proposition 6.26. *Let α be a root in the sense of Notation 6.24 and let H_α be the corresponding co-root. Then, α and H_α are related by the formulas*

$$H_\alpha = 2\frac{\alpha}{\langle \alpha, \alpha \rangle}, \tag{6.5}$$

$$\alpha = 2\frac{H_\alpha}{\langle H_\alpha, H_\alpha \rangle}. \tag{6.6}$$

The real content of this proposition is that once we use the inner product to identity $\mathfrak{h}^*$ with $\mathfrak{h}$ (so that the roots and co-roots now live in the same space), α and H_α are multiples of one another. Once this is known, the normalization is determined by the condition that $\langle \alpha, H_\alpha \rangle = 2$, which reflects that $[H_\alpha, X_\alpha] = 2X_\alpha$. Observe that both (6.5) and (6.6) are consistent with the relation $\langle \alpha, H_\alpha \rangle = 2$.

Proof. In the previous section, we established that H_α belongs to the one-dimensional subspace $(\ker \alpha)^\perp$ of $\mathfrak{h}$. Now that we are thinking of α as an element of $\mathfrak{h}$ instead of $\mathfrak{h}^*$, $\ker \alpha$ is equal to $\{H \in \mathfrak{h} | \langle \alpha, H \rangle = 0\}$, which is just the orthogonal complement of the span of α. Thus $(\ker \alpha)^\perp$ is equal to $((\operatorname{span} \alpha)^\perp)^\perp = \operatorname{span} \alpha$. This means that α and H_α are multiples of one another. Then, as remarked above, the constants for expressing α in terms of H_α and vice versa are determined by the normalization condition $\langle \alpha, H_\alpha \rangle = 2$. □

Note that the expression (6.5) for H_α in terms of α is exactly parallel to the expression (6.6) for α in terms of H_α. If we substitute (6.6) into (6.5) we obtain $H_\alpha = 4H_\alpha / \langle \alpha, \alpha \rangle \langle H_\alpha, H_\alpha \rangle$. Thus

$$\langle \alpha, \alpha \rangle \langle H_\alpha, H_\alpha \rangle = 4. \tag{6.7}$$

A more symmetric way of expressing the relationship between α and H_α is to say that they are multiples of one another and their lengths are related by (6.7).

If we restate Proposition 6.22 using our new point of view on the roots and also using (6.5), we obtain that

$$\langle \beta, H_\alpha \rangle = 2\frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} \tag{6.8}$$

must be an integer for all roots α and β. This implies that $\langle \beta, \alpha \rangle$ is a real number, and so we can just as well say that

$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}$$

is an integer. If we use (6.6) (applied to β) instead of (6.5), we obtain that

$$\langle \beta, H_\alpha \rangle = 2\frac{\langle H_\beta, H_\alpha \rangle}{\langle H_\beta, H_\beta \rangle}$$

is also an integer. We have obtained, then, the following result.

Theorem 6.27. *Consider the roots in the sense of Notation 6.24 and the coroots defined by Theorem 6.20 and satisfying Proposition 6.26. Then for all roots α and β, the quantities*

$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \tag{6.9}$$

and

$$2\frac{\langle H_\beta, H_\alpha \rangle}{\langle H_\beta, H_\beta \rangle} \tag{6.10}$$

are integers and, furthermore,

$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = 2\frac{\langle H_\beta, H_\alpha \rangle}{\langle H_\beta, H_\beta \rangle}.$$

Note that the expressions (6.9) and (6.10) are both equal to $\langle \beta, H_\alpha \rangle$, and $\langle \beta, H_\alpha \rangle$ is the eigenvalue of ad_{H_α} in the root space $\mathfrak{g}_\beta$. This is the reason that these quantities must be integers. Recall from elementary linear algebra that if α and β are elements of some inner-product space, then the orthogonal projection of β onto α is given by

$$\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha.$$

The first quantity in Theorem 6.27 is thus twice the coefficient of α in the projection of β onto α. We may therefore interpret the integrality result in the following geometric way: *If α and β are roots, then the orthogonal projection of α onto β must be an integer or half-integer multiple of β, and vice versa.* This condition severely restricts the possible angles between α and β and (if α and β are not orthogonal) the possible ratios of their lengths. See Proposition 8.6.

6.6 The Weyl Group

We use here the compact group approach to defining the Weyl group, as opposed to the Lie algebra approach. The compact-group approach makes certain aspects of the Weyl group more transparent. Nevertheless, the two

approaches are equivalent. See the comments following Theorem 6.33 at the end of this section.

We continue with the setting of the previous section. Thus, $\mathfrak{g}$ is a complex semisimple Lie algebra given to us as a subalgebra of some $\mathfrak{gl}(n;\mathbb{C})$. We have chosen a compact real form $\mathfrak{k}$ of $\mathfrak{g}$ and we let K be the compact subgroup of $\mathsf{GL}(n;\mathbb{C})$ whose Lie algebra is $\mathfrak{k}$. We have chosen a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$, and we work with the associated Cartan subalgebra $\mathfrak{h} = \mathfrak{t}+i\mathfrak{t}$. We have chosen an inner product on $\mathfrak{g}$ that is invariant under the adjoint action of K and that takes real values on $\mathfrak{k}$.

Consider the following two subgroups of K:

$$Z(\mathfrak{t}) = \{A \in K \,|\, \mathrm{Ad}_A(H) = H \text{ for all } H \text{ in } \mathfrak{t}\},$$
$$N(\mathfrak{t}) = \{A \in K \,|\, \mathrm{Ad}_A(H) \subset \mathfrak{t} \text{ for all } H \text{ in } \mathfrak{t}\}.$$

Clearly, $Z(\mathfrak{t})$ is a subgroup of $N(\mathfrak{t})$, and it is easily seen that $Z(\mathfrak{t})$ is a *normal* subgroup of $N(\mathfrak{t})$. See Exercise 10 for an explanation of the notation. If T is the connected Lie subgroup of K with Lie algebra $\mathfrak{t}$, then $T \subset Z(\mathfrak{t})$, since T is generated by elements of the form $\exp H$ with H in $\mathfrak{t}$. It turns out that, in fact, $Z(\mathfrak{t}) = T$. See Bröcker and tom Dieck (1985).

Definition 6.28. *The Weyl group for* $\mathfrak{g}$ *is the quotient group* $W = N(\mathfrak{t})/Z(\mathfrak{t})$.

We can define an action of W on $\mathfrak{t}$ as follows. For each element w of W, choose an element A of the corresponding equivalence class in $N(\mathfrak{t})$. Then for H in $\mathfrak{t}$ we define the action $w \cdot H$ of w on H by

$$w \cdot H = \mathrm{Ad}_A(H).$$

As in the $\mathsf{SU}(3)$ case (Section 5.6), it is easy to verify that this action is well defined (i.e., independent of the choice of A in a given equivalence class). Since $\mathfrak{h} = \mathfrak{t}+i\mathfrak{t}$, each linear transformation of $\mathfrak{t}$ extends uniquely to a complex-linear transformation of $\mathfrak{h}$. Thus, we also think of W as acting on $\mathfrak{h}$. If w is an element of the Weyl group, then we write $w \cdot H$ for the action of w on an element H of $\mathfrak{h}$. It is easily seen that W is isomorphic to the group of linear transformations of $\mathfrak{h}$ that can be expressed as Ad_A for some $A \in N(\mathfrak{t})$.

Proposition 6.29.

1. *The inner product* $\langle \cdot, \cdot \rangle$ *on* $\mathfrak{h}$ *is invariant under the action of* W.
2. *The set* $R \subset \mathfrak{h}$ *of roots is invariant under the action of* W.
3. *The set of co-roots is invariant under the action of* W, *and* $w \cdot H_\alpha = H_{w \cdot \alpha}$ *for all* $w \in W$ *and* $\alpha \in R$.
4. *The Weyl group is a finite group.*

Proof. The action of W on $\mathfrak{h}$ is nothing but the adjoint action of $N(\mathfrak{t}) \subset K$ on $\mathfrak{g}$, restricted to $\mathfrak{h}$. Since the inner product on $\mathfrak{g}$ is invariant under the adjoint action of K, the restriction of this inner product to $\mathfrak{h}$ is invariant under the adjoint action of $N(\mathfrak{t})$. This establishes Point 1.

For Point 2, given an element w of the Weyl group, let $A \in K$ be an element of $N(\mathfrak{t})$ that represents w. Now, suppose that $\alpha \in \mathfrak{h}$ is a root. Then (Notation 6.24), there exists a nonzero element X of $\mathfrak{g}$ such that

$$[H, X] = \langle \alpha, H \rangle X$$

for all H in $\mathfrak{h}$. Now, let us consider the element $\mathrm{Ad}_A(X)$ of $\mathfrak{g}$ and compute how $\mathfrak{h}$ acts on it. For H in $\mathfrak{h}$, we have

$$[H, \mathrm{Ad}_A(X)] = \mathrm{Ad}_A([\mathrm{Ad}_{A^{-1}}(H), X]), \tag{6.11}$$

because Ad_A is a Lie algebra automorphism. Since A is in $N(\mathfrak{t})$, $\mathrm{Ad}_{A^{-1}}(H)$ is again in $\mathfrak{h}$ and so

$$[\mathrm{Ad}_{A^{-1}}(H), X] = \langle \alpha, \mathrm{Ad}_{A^{-1}}(H) \rangle X.$$

Thus, (6.11) becomes

$$[H, \mathrm{Ad}_A(X)] = \langle \alpha, \mathrm{Ad}_{A^{-1}}(H) \rangle \mathrm{Ad}_A(X). \tag{6.12}$$

Since the inner product on $\mathfrak{h}$ is invariant under Ad_A, we have

$$\langle \alpha, \mathrm{Ad}_{A^{-1}}(H) \rangle = \langle \mathrm{Ad}_A(\alpha), H \rangle = \langle w \cdot \alpha, H \rangle .$$

Thus, (6.12) becomes

$$[H, \mathrm{Ad}_A(X)] = \langle w \cdot \alpha, H \rangle \mathrm{Ad}_A(X).$$

This shows that $w \cdot \alpha$ is a root with root vector $\mathrm{Ad}_A(X)$ $(= w \cdot X)$, establishing Point 2.

For Point 3, we write

$$H_\alpha = 2 \frac{\alpha}{\langle \alpha, \alpha \rangle},$$

as in Proposition 6.26. Then

$$w \cdot H_\alpha = 2 \frac{w \cdot \alpha}{\langle \alpha, \alpha \rangle} = 2 \frac{w \cdot \alpha}{\langle w \cdot \alpha, w \cdot \alpha \rangle} = H_{w \cdot \alpha},$$

where in the second equality we have used the invariance of the inner product under the action of W.

Finally, we note that since the roots span $\mathfrak{h}$, the action of an element w on $\mathfrak{h}$ is determined by what w does to the roots. However, each w preserves the set of roots and thus we may think of each w as a permutation of the set of roots, and there are only finitely many of these. □

Let us now compute the Weyl group of $\mathsf{sl}(2;\mathbb{C})$. This is a worthwhile exercise in its own right, and, as with other $\mathsf{sl}(2;\mathbb{C})$ calculations, it will aid us in our analysis of general semisimple Lie algebras. We consider the compact real form $\mathfrak{k} = \mathsf{su}(2)$ of $\mathsf{sl}(2;\mathbb{C})$ and the maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$ given by

$$\mathfrak{t} = \left\{ \begin{pmatrix} ia & 0 \\ 0 & -ia \end{pmatrix} \middle| \, a \in \mathbb{R} \right\}.$$

Theorem 6.30. *The subgroup $Z(\mathfrak{t})$ of $\mathsf{SU}(2)$ is given by*

$$Z(\mathfrak{t}) = \left\{ \begin{pmatrix} e^{ia} & 0 \\ 0 & e^{-ia} \end{pmatrix} \middle| \, a \in \mathbb{R} \right\}$$

and $N(\mathfrak{t})$ is the set of matrices A in $\mathsf{SU}(2)$ that are either in $Z(\mathfrak{t})$ or of the form

$$A = \begin{pmatrix} 0 & e^{ia} \\ -e^{-ia} & 0 \end{pmatrix} \tag{6.13}$$

for a in $\mathbb{R}$. The Weyl group $N(\mathfrak{t})/Z(\mathfrak{t})$ has two elements. For any H in $\mathfrak{h} = \mathfrak{t}+i\mathfrak{t}$ and for any A of the form (6.13), we have

$$AHA^{-1} = -H.$$

Proof. Let H_0 be the element of $\mathfrak{t}$ given by

$$H_0 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Now, suppose that $A \in \mathsf{SU}(2)$ commutes with each element of $\mathfrak{t}$ and thus, in particular, with H_0. Since A commutes with H_0, it must preserve each of the eigenspaces for H_0, which are $\mathbb{C}e_1$ and $\mathbb{C}e_2$, where e_1 and e_2 are the standard basis elements for $\mathbb{C}^2$. Thus, Ae_1 must equal c_1e_1 and Ae_2 must equal c_2e_2, for some constants c_1 and c_2. Thus,

$$A = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}.$$

We also require that A be in $\mathsf{SU}(2)$, which means that $|c_1| = |c_2| = 1$ and that $c_2 = 1/c_1$. Thus,

$$A = \begin{pmatrix} e^{ia} & 0 \\ 0 & e^{-ia} \end{pmatrix} \tag{6.14}$$

for some a in $\mathbb{R}$. Clearly, also, every matrix of the form (6.14) does, in fact, commute with every element of $\mathfrak{t}$, so $Z(\mathfrak{t})$ is precisely the group of matrices of this form.

We now compute $N(\mathfrak{t})$. If $A \in N(\mathfrak{t})$, then, in particular, AH_0A^{-1} must be in $\mathfrak{t}$. Now, the eigenvectors of H_0 are e_1 and e_2, with eigenvalues i and $-i$, respectively. The eigenvectors of AH_0A^{-1} are then (check!) Ae_1 and Ae_2, with the same eigenvalues i and $-i$. However, every element of $\mathfrak{t}$ has eigenvectors e_1 and e_2, so we must have either $Ae_1 = c_1e_1$ and $Ae_2 = c_2e_2$ (in which case, $A \in Z(\mathfrak{t})$) or $Ae_1 = c_1e_2$ and $Ae_2 = c_2e_1$, in which case,

$$A = \begin{pmatrix} 0 & c_2 \\ c_1 & 0 \end{pmatrix}.$$

Then, for A to be in $\mathsf{SU}(2)$, we must have $|c_1| = |c_2| = 1$ and $c_2 = -1/c_1$, so

$$A = \begin{pmatrix} 0 & e^{ia} \\ e^{-ia} & 0 \end{pmatrix}.$$

If A is of this form, then we compute that

$$A \begin{pmatrix} ia & 0 \\ 0 & -ia \end{pmatrix} A^{-1} = \begin{pmatrix} -ia & 0 \\ 0 & ia \end{pmatrix} \in \mathfrak{t}, \tag{6.15}$$

and, indeed, A is in $N(\mathfrak{t})$. Therefore, the elements of $N(\mathfrak{t})$ are precisely matrices of the form (6.14) or (6.13).

Now, a matrix of the form (6.13) is not in $Z(\mathfrak{t})$. However, if A and B are two matrices of the form (6.13), then direct calculation shows that $AB^{-1} \in Z(\mathfrak{t})$. This shows that the quotient group $N(\mathfrak{t})/Z(\mathfrak{t})$ has precisely two elements. Let us see how the Weyl group $N(\mathfrak{t})/Z(\mathfrak{t})$ acts on $\mathfrak{t}$. For A of the form (6.13), the restriction of Ad_A to $\mathfrak{t}$ will be $-I$ (by (6.15)), and for A of the form (6.14), the restriction of Ad_A to $\mathfrak{t}$ is I. So, the Weyl group may be identified with the two-element group $\{I, -I\}$ inside $\mathsf{GL}(\mathfrak{t}) \cong \mathsf{GL}(1; \mathbb{R})$. □

For application to general semisimple Lie algebras, it is useful to describe the Weyl group in Lie algebraic terms. So, let X, Y, and H be the usual basis elements for $\mathsf{sl}(2; \mathbb{C})$. Then, the matrix

$$\frac{\pi}{2}(X - Y) = \begin{pmatrix} 0 & \pi/2 \\ -\pi/2 & 0 \end{pmatrix}$$

is in $\mathsf{su}(2)$ and by the calculation in Section 2.2, we have

$$\exp\left[\frac{\pi}{2}(X - Y)\right] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{6.16}$$

which represents the one nontrivial element of the Weyl group of $\mathsf{sl}(2; \mathbb{C})$. (Here, π is the number $\pi = 3.14...$, *not* a representation.) Then, for any H in $\mathfrak{h}$, we have

$$\mathrm{Ad}_{\exp[(\pi/2)(X-Y)]} H = \exp\left[\frac{\pi}{2}(\mathrm{ad}_X - \mathrm{ad}_Y)\right](H) = -H. \tag{6.17}$$

The element on the right-hand side of (6.16) can also be computed as

$$e^X e^{-Y} e^X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{6.18}$$

We are now ready to apply our knowledge of the Weyl group of $\mathsf{sl}(2; \mathbb{C})$ to obtain information about the Weyl group of a general complex semisimple Lie algebra $\mathfrak{g}$. We now resume the context that $\mathfrak{g}$ is a complex semisimple Lie algebra, $\mathfrak{k}$ is a fixed compact real form, $\mathfrak{t}$ is a fixed maximal commutative subalgebra of $\mathfrak{k}$, and $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$.

Theorem 6.31. *For each root α, there exists an element w_α of W such that*

$$w_\alpha \cdot \alpha = -\alpha$$

and such that

$$w_\alpha \cdot H = H$$

for all H in $\mathfrak{h}$ with $\langle \alpha, H \rangle = 0$.

Note that since H_α is a multiple of α, saying $w_\alpha \cdot \alpha = -\alpha$ is equivalent to saying that $w_\alpha \cdot H_\alpha = -H_\alpha$.

The linear operator corresponding to the action of w_α on $\mathfrak{h}$ is "the reflection about the hyperplane perpendicular to α." This means that w_α acts as the identity on the hyperplane (codimension-one subspace) perpendicular to α and as minus the identity on the span of α. We can work out a formula for w_α as follows. Any vector β can be decomposed uniquely as a multiple of α plus a vector orthogonal to α. This decomposition is given explicitly by

$$\beta = \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \alpha + \left(\beta - \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \alpha \right), \tag{6.19}$$

where direct calculation shows that the second term is indeed orthogonal to α. (The projection of β onto α is a linear function of β and so β must go in the "linear" side of the inner product, which for us is the right-hand side.) Now, to obtain $w_\alpha \cdot \beta$, we should change the sign of the part of β parallel to α and leave alone the part of β that is orthogonal to α. This means that we change the sign of the first term on the right-hand side of (6.19), giving

$$w_\alpha \cdot \beta = \beta - 2 \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \alpha. \tag{6.20}$$

We now have another way of thinking about the quantity $2\langle \alpha, \beta \rangle / \langle \alpha, \alpha \rangle$ in Theorem 6.27: It is the coefficient of α in the expression for $w_\alpha \cdot \beta$. So, we can re-express Theorem 6.27 as follows.

Corollary 6.32. *If α and β are roots, then $\beta - w_\alpha \cdot \beta$ is an integer multiple of α.*

Note that for any β in $\mathfrak{h}$, whether a root or not, $\beta - w_\alpha \cdot \beta$ will be a multiple of α, as a consequence of the formula (6.20) for w_α. The content of the corollary is that if β happens to be a root, then $\beta - w_\alpha \cdot \beta$ is an *integer* multiple of α.

We now turn to the proof of Theorem 6.31.

Proof. We choose elements X_α, Y_α, and H_α as in Point 3 of Theorem 6.20. As shown in Step 3 of the proof of that theorem, it is possible to choose X_α and Y_α so that $Y_\alpha = X_\alpha^*$, in which case, $X_\alpha - Y_\alpha$ will be contained in $\mathfrak{k}$. (Recall the definition (6.2) of X^*.) We then let A_α be the element of K given by

$$A_\alpha = \exp\left[\frac{\pi}{2}(X_\alpha - Y_\alpha)\right]. \tag{6.21}$$

We want to show that A_α is in $N(\mathfrak{t})$ and that Ad_{A_α} acts in the indicated way on $\mathfrak{h}$.

Suppose, first, that H is in $\mathfrak{h}$ and that $\langle \alpha, H\rangle = 0$. Then, $[H, X_\alpha] = \langle \alpha, H\rangle X_\alpha = 0$ and similarly for Y_α; that is, X_α and Y_α commute with H. Therefore, using the relationship between Ad and ad (Proposition 2.25), we have

$$\begin{aligned}\mathrm{Ad}_{A_\alpha}(H) &= \exp\left[\frac{\pi}{2}(\mathrm{ad}_{X_\alpha} - \mathrm{ad}_{Y_\alpha})\right](H)\\ &= H.\end{aligned}$$

Now, let us consider the action of Ad_{A_α} on the one-dimensional subspace of $\mathfrak{g}$ spanned by H_α (or by α). We have, as above,

$$\mathrm{Ad}_{A_\alpha}(H_\alpha) = \exp\left[\frac{\pi}{2}(\mathrm{ad}_{X_\alpha} - \mathrm{ad}_{Y_\alpha})\right](H_\alpha). \tag{6.22}$$

The right-hand side of (6.22) involves only Lie algebra quantities. Thus, since the Lie algebra spanned by X_α, Y_α, and H_α is isomorphic to the one spanned by the usual $\mathfrak{sl}(2;\mathbb{C})$ basis elements, the result of computing the right-hand side of (6.22) will be the same as in $\mathfrak{sl}(2;\mathbb{C})$, which we have computed above in (6.17). We obtain, then, that

$$\mathrm{Ad}_{A_\alpha}(H_\alpha) = -H_\alpha. \tag{6.23}$$

See Exercise 8 for an alternative calculation of this result.

So, Ad_{A_α} acts as the identity on elements H with $\langle \alpha, H\rangle = 0$, and Ad_{A_α} acts as minus the identity on the span of H_α, which is the same as the span of α. This shows that A_α represents an element of the Weyl group that acts in the indicated way on $\mathfrak{h}$. □

The element A_α in (6.21) can also be computed as

$$A_\alpha = e^{X_\alpha} e^{-Y_\alpha} e^{X_\alpha}. \tag{6.24}$$

(Compare (6.18).) Direct computation shows that the elements in (6.21) and (6.24) are equal in the $\mathfrak{sl}(2;\mathbb{C})$ case and, then, the argument in the proof of Theorem 6.31 shows that they are equal in general. The description of the A_α's given in (6.24) will be useful in the Verma module construction of the representations of $\mathfrak{g}$. See Section 7.3.

Theorem 6.33. *The Weyl group W is generated by the elements w_α as α ranges over all roots.*

That is to say, the smallest subgroup of W that contains all of the w_α's is W itself. This is somewhat involved to prove and I will not do so here; see Bröcker and tom Dieck (1985). In the Lie algebra approach to the Weyl group, the Weyl group is *defined* as the set of linear transformations of $\mathfrak{h}$ generated by the reflections w_α. Theorem 6.33 shows that the Lie algebra definition of the Weyl group gives the same group as the compact-group approach.

6.7 Root Systems

In the previous sections, we have established several properties of the roots. From Proposition 6.15, we know that the roots are imaginary on $\mathfrak{t}$, which, after transferring the roots from $\mathfrak{h}^*$ to $\mathfrak{h}$ (as in Notation 6.24), means that the roots live in $i\mathfrak{t} \subset \mathfrak{h}$. The inner product $\langle \cdot, \cdot \rangle$ was constructed to take real values on $\mathfrak{k}$ and, hence, on $\mathfrak{t}$. The inner product then also takes real values on $i\mathfrak{t}$, since $\langle iX, iY \rangle = (-i)i \langle X, Y \rangle = \langle X, Y \rangle$. So, the roots live in the real inner-product space $E = i\mathfrak{t}$.

From Proposition 6.19 and Theorem 6.20 we know that the roots span $i\mathfrak{t}$ and that if α is a root, then $-\alpha$ is a root but no other multiples of α are roots. Theorem 6.27 tells us that for any roots α and β, $2\langle \alpha, \beta \rangle / \langle \alpha, \alpha \rangle$ is an integer. Proposition 6.29 tells us that the roots are invariant under the action of the Weyl group, and Theorem 6.31 tells us that the Weyl group contains the reflection about the hyperplane orthogonal to each root α. We summarize these results in the following theorem.

Theorem 6.34. *The roots form a finite set of nonzero elements of a real inner-product space E and have the following properties:*

1. *The roots span E.*
2. *If α is a root, then $-\alpha$ is a root and the only multiples of α that are roots are α and $-\alpha$.*
3. *If α is a root, let w_α denote the linear transformation of E given by*

$$w_\alpha \cdot \beta = \beta - 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha.$$

   Then, for all roots α and β, $w_\alpha \cdot \beta$ is also a root.
4. *If α and β are roots, then the quantity*

$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}$$

   is an integer.

Any collection of vectors in a finite-dimensional real inner-product space having these properties is called a **root system**. The **Weyl group** for a root system R is the group of linear transformations of E generated by the w_α's. We will look more closely at the properties of root systems in Chapter 8. Note that Point 4 is equivalent to saying that $\beta - w_\alpha \cdot \beta$ must be an integer multiple of α for all roots α and β.

We have also established certain important properties of the root *spaces* that are not properties of the roots themselves, namely that each root space $\mathfrak{g}_\alpha$ is one dimensional and that out of $\mathfrak{g}_\alpha$, $\mathfrak{g}_{-\alpha}$, and $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$, we can form a subalgebra isomorphic to $\mathfrak{sl}(2; \mathbb{C})$.

Finally, I claim that the co-roots H_α (as in Point 3 of Theorem 6.20) themselves form a root system. Theorem 6.27 tells us that the co-roots satisfy

Property 4 and Proposition 6.29 tells us that the set of co-roots is invariant under the Weyl group and hence, in particular, under the reflections w_α. However, note that since H_α is a multiple of α, the reflection generated by H_α is the same as the reflection generated by α. Thus, the set of co-roots satisfies Property 3. Properties 1 and 2 for the co-roots follow from the corresponding properties for the roots, since each H_α is a multiple of α. The set of co-roots is called the "dual root system" to the set of roots. See Chapter 8 for more information on root systems, including many pictures.

6.8 Positive Roots

In the next chapter, we will classify the irreducible representations of $\mathfrak{g}$ in terms of a "highest weight." In the $\mathfrak{sl}(3;\mathbb{C})$ case, we defined "highest" in terms of the two roots α_1 and α_2. There is nothing sacred about those particular two roots. What we need is simply some consistent notion of higher and lower that will allow us to divide the root vectors X_α into "raising operators" and "lowering operators." This should be done in such a way that the commutator of two raising operators is, again, a raising operator and not a lowering operator. This means that we want to divide the roots into two groups, one of which will be called "positive" and the other "negative." This should be done is such a way that if the sum of positive roots is again a root, that root should be positive. There is no unique way to make the division into positive and negative; any consistent division will do. The uniqueness theorems of the next section show that it does not really matter which choice we make.

The following definition and theorem shows that it is possible to make a good choice.

Definition 6.35. *Suppose that E is a finite-dimensional real inner-product space and that $R \subset E$ is a root system. Then, a* ***base*** *for R is a subset $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ of R such that Δ forms a basis for E as a vector space and such that for each $\alpha \in R$, we have*

$$\alpha = n_1\alpha_1 + n_2\alpha_2 + \cdots + n_r\alpha_r,$$

where the n_j's are integers and either all greater than or equal to zero or all less than or equal to zero.

Once a base Δ has been chosen, the α's for which $n_j \geq 0$ are called the ***positive roots*** *(with respect to the given choice of Δ) and the α's with $n_j \leq 0$ are called the* ***negative roots****. The elements of Δ are called the* ***positive simple roots****.*

Therefore, to be a base (in the sense of root systems), $\Delta \subset R$ must, in particular, be a basis for E in the vector space sense. In addition, the expansion of any $\alpha \in R$ in terms of the elements of Δ must have integer coefficients and all of the nonzero coefficients (for a given α) must be of the same sign.

Theorem 6.36. *For any root system, a base exists.*

The proof of Theorem 6.36 is given in Section 8.3. In the case of $\mathsf{sl}(3;\mathbb{C})$, one should verify that any pair of roots with a 120° angle constitutes a base, but that a pair of roots with a 60° angle does not constitute a base. See Chapter 8 for additional examples and pictures.

Proposition 6.37. *If R is the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$ and if $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ is a base for R, then $\{H_{\alpha_1}, \ldots, H_{\alpha_r}\}$ is a base for the system of co-roots.*

We will prove this in Chapter 8.

6.9 The $\mathsf{sl}(n;\mathbb{C})$ Case

Let us see how all of the structures described in this chapter work out in the case $\mathfrak{g} = \mathsf{sl}(n;\mathbb{C})$. For calculations in the case of other classical Lie algebras, see Exercises 12, 13, and 14 in this chapter and Section 8.8.

6.9.1 The Cartan subalgebra

We work with the compact real form $\mathfrak{k} = \mathsf{su}(n)$ and the maximal commutative subalgebra $\mathfrak{t}$ which is the intersection of the set of diagonal matrices with $\mathsf{su}(n)$; that is,

$$\mathfrak{t} = \left\{ \begin{pmatrix} ia_1 & & \\ & \ddots & \\ & & ia_n \end{pmatrix} \middle| \, a_j \in \mathbb{R}, \quad a_1 + \cdots + a_n = 0 \right\}. \tag{6.25}$$

It is clear that $\mathfrak{t}$ is a commutative subalgebra of $\mathfrak{k}$; that $\mathfrak{t}$ is *maximal* commutative will be evident once we compute the roots. The associated Cartan subalgebra is then $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$, so that $\mathfrak{h}$ is the set of all diagonal matrices with trace zero:

$$\mathfrak{h} = \left\{ \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \middle| \, \lambda_j \in \mathbb{C}, \quad \lambda_1 + \cdots + \lambda_n = 0 \right\}. \tag{6.26}$$

6.9.2 The roots

Now, let E_{kl} denote the matrix that has a one in the k^{th} row and l^{th} column and that has zeros elsewhere. A simple calculation shows that if $H \in \mathfrak{t}$ is as in (6.26), then $HE_{kl} = \lambda_k E_{kl}$ and $E_{kl}H = \lambda_l E_{kl}$. Thus,

$$[H, E_{kl}] = (\lambda_k - \lambda_l) E_{kl}. \tag{6.27}$$

If $k = l$, then E_{kl} does not have trace zero and so is not in $\mathsf{sl}(n;\mathbb{C})$. If $k \neq l$, then E_{kl} is in $\mathsf{sl}(n;\mathbb{C})$ and (6.27) shows that E_{kl} is a simultaneous eigenvector for each ad_H with H in $\mathfrak{h}$, with eigenvalue $\lambda_k - \lambda_l$. Note that every element X of $\mathsf{sl}(n;\mathbb{C})$ can be written uniquely as an element of the Cartan subalgebra (the diagonal entries of X) plus a linear combination of the E_{kl}'s with $k \neq l$ (the off-diagonal entries of X). From this it is not hard to see that $\mathfrak{t}$ is actually maximal commutative and, therefore, that $\mathfrak{h}$ is actually a Cartan subalgebra (Exercise 15).

If we think at first of the roots as elements of $\mathfrak{h}^*$, then (according to (6.27)) the roots are the linear functionals α_{kl} that associate to each $H \in \mathfrak{h}$, as in (6.26), the quantity $\lambda_k - \lambda_l$. Note that $\alpha_{lk} = -\alpha_{kl}$ but that no other multiple of α_{kl} is a root. Also, each root space is one dimensional—the span of E_{kl}. For each root $\alpha = \alpha_{kl}$, we may take $X_\alpha = E_{kl}$, $Y_\alpha = E_{lk}$, and $H_\alpha = [X_\alpha, Y_\alpha] = E_{kk} - E_{ll}$. The subalgebra spanned by X_α, Y_α, and H_α is just the copy of $\mathsf{sl}(2;\mathbb{C})$ inside $\mathsf{sl}(n;\mathbb{C})$ in which all the action is in the k^{th} and l^{th} coordinates.

The roots of $\mathsf{sl}(n;\mathbb{C})$ form a root system that is conventionally called A_{n-1}, with the subscript $n-1$ indicating that the rank of $\mathsf{sl}(n;\mathbb{C})$ (i.e., the dimension of $\mathfrak{h}$) is $n-1$.

6.9.3 Inner products of roots

We use the Hilbert–Schmidt inner product on $\mathsf{sl}(n;\mathbb{C})$, namely

$$\langle X, Y \rangle = \mathrm{trace}(X^* Y),$$

where X^* is the usual matrix adjoint of X. This inner product is invariant under the adjoint action of $\mathsf{SU}(n)$, as is easily verified (Exercise 9). When we restrict this inner product to $\mathfrak{h}$ we get the "obvious" inner product on $\mathfrak{h}$, in which the inner product of $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ is equal to $\bar{\lambda}_1\sigma_1 + \cdots + \bar{\lambda}_n\sigma_n$. (Here $\mathrm{diag}(\cdot)$ is the diagonal matrix with the indicated diagonal entries.) If we use this inner product as in Notation 6.24 to transfer the roots from $\mathfrak{h}^*$ to $\mathfrak{h}$, we obtain

$$\alpha_{kl} = E_{kk} - E_{ll} \quad (k \neq l).$$

We see, then, that each root satisfies

$$\langle \alpha_{kl}, \alpha_{kl} \rangle = 2.$$

Furthermore,

$$\langle \alpha_{kl}, \alpha_{k'l'} \rangle$$

has the value 0, ± 1, or ± 2, depending on whether $\{k, l\}$ and $\{k', l'\}$ have zero, one, or two elements in common. Thus

$$2\frac{\langle \alpha_{kl}, \alpha_{k'l'} \rangle}{\langle \alpha_{kl}, \alpha_{kl} \rangle} \in \{0, \pm 1, \pm 2\}.$$

Note that all the roots have length $\sqrt{2}$. If α and β are roots and $\alpha \neq \beta$ and $\alpha \neq -\beta$, then the angle between α and β is either $60°$ (if $\langle \alpha, \beta \rangle = 1$), $90°$ (if $\langle \alpha, \beta \rangle = 0$), or $120°$ (if $\langle \alpha, \beta \rangle = -1$). (In other root systems, the roots may not all have the same length, and other angles can arise.) Since each root for $\mathfrak{sl}(n;\mathbb{C})$ has length $\sqrt{2}$ (with this choice of inner product), we see that the co-roots $H_\alpha = 2\alpha/\langle \alpha, \alpha \rangle$ simply coincide with the roots α. (In other root systems, the system of co-roots may be inequivalent to the system of roots.)

6.9.4 The Weyl group

Following the argument in the computation of the Weyl group of $\mathfrak{sl}(2;\mathbb{C})$, we see that $Z(\mathfrak{t})$ is the subgroup of $K = \mathrm{SU}(n)$ consisting of diagonal matrices, and $N(\mathfrak{t})$ is the set of matrices A in $\mathrm{SU}(n)$ such that for each k there exists an l such that Ae_k is a multiple of e_l. This means that associated to each A in $N(\mathfrak{t})$ there is a permutation $(k \to l)$. So, the Weyl group is isomorphic to the permutation group S_n and it acts on $\mathfrak{h}$ by permuting the diagonal entries. For each $\alpha = \alpha_{kl}$, the Weyl group element w_α acts by interchanging the k^{th} and l^{th} diagonal entries, and the Weyl group is generated by such interchanges.

6.9.5 Positive roots

Finally, we can find a base as follows. We take as our base the roots α_{kl} with $l = k + 1$ (i.e., the roots $\alpha_{k,k+1}$). Recall that $\alpha_{kl} = E_{kk} - E_{ll}$. If $k < l$, we note that

$$E_{kk} - E_{ll} = (E_{kk} - E_{k-1,k-1}) + (E_{k-1,k-1} - E_{k-2,k-2}) + \cdots + (E_{l+1,l+1} - E_{ll})$$

and, so,

$$\alpha_{kl} = \alpha_{k-1,k} + \alpha_{k-1,k-2} + \cdots + \alpha_{l,l+1}.$$

Thus, every root α_{kl} with $k < l$ can be written as a linear combination of the simple roots with non-negative integer coefficients. (In fact, the coefficients are either 0 or 1.) So, the α_{kl}'s with $k < l$ are the positive roots and the α_{kl}'s with $k > l$ are the negative roots. Every negative root is just the negative of a positive root and so can be written as a linear combination of the positive simple roots with nonpositive integer coefficients.

6.10 Uniqueness Results

Given a complex semisimple Lie algebra $\mathfrak{g}$, we have made three choices: a compact real form $\mathfrak{k}$ of $\mathfrak{g}$, a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$, and a system of positive roots. None of these objects is unique, and so it is important to think about how different possibilities are related. Furthermore, we have only considered Cartan subalgebras that arise as the complexification of maximal

commutative subalgebras of a compact real form. It is not obvious that every Cartan subalgebra arises in this way.

The following theorems take care of all such worries. The following theorems tell us that each structure is unique up to the adjoint action of relevant group (G, K, or W).

We assume that $\mathfrak{g}$ is given to us as a subalgebra of some $\mathfrak{gl}(n;\mathbb{C})$, and we let G denote the connected Lie subgroup of $\mathrm{GL}(n;\mathbb{C})$ with Lie algebra $\mathfrak{g}$.

Theorem 6.38. *Suppose that $\mathfrak{k}_1$ and $\mathfrak{k}_2$ are two compact real forms of $\mathfrak{g}$. Then, there is an element A of G such that $\mathrm{Ad}_A(\mathfrak{k}_1) = \mathfrak{k}_2$.*

Theorem 6.39. *Suppose that $\mathfrak{k}$ is a compact real form of $\mathfrak{g}$ and that K is the compact subgroup of G whose Lie algebra is $\mathfrak{k}$. Suppose that $\mathfrak{t}_1$ and $\mathfrak{t}_2$ are two maximal commutative subalgebras of $\mathfrak{k}$. Then, there is an element A of K such that $\mathrm{Ad}_A(\mathfrak{t}_1) = \mathfrak{t}_2$.*

Theorem 6.40. *If $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, then there exists a compact real form $\mathfrak{k}$ of $\mathfrak{g}$ and a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$ such that $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$. If $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two Cartan subalgebras of $\mathfrak{g}$, then there exists $A \in G$ such that $\mathrm{Ad}_A(\mathfrak{h}_1) = \mathfrak{h}_2$.*

Theorem 6.41. *Any two systems of positive simple roots can be mapped into one another by the action of the Weyl group.*

In the last theorem, it should be understood that different orderings of $\alpha_1, \ldots, \alpha_r$ count as the same system of positive simple roots.

I will not prove these results. Exercises 16, 18, and 19 provide proofs of some of these results in the case $\mathfrak{g} = \mathfrak{sl}(n;\mathbb{C})$.

A related issue is the extent to which one can recover a semisimple Lie algebra from its root system. This will be discussed in Chapter 8. The result is that if $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are complex semisimple Lie algebras and the root system for $\mathfrak{g}_1$ is isomorphic (in the appropriate sense) to the root system for $\mathfrak{g}_2$, then $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are isomorphic Lie algebras. See Theorem 8.28.

6.11 Exercises

1. Show that the center of any semisimple Lie algebra $\mathfrak{g}$ is trivial. Show that the adjoint representation of $\mathfrak{g}$ is faithful.
2. Suppose that $\mathfrak{g}$ is a complex Lie algebra with the complete reducibility property. Show that $\mathfrak{g}$ is semisimple.
   Hint: First show that a one-dimensional commutative Lie algebra does not have the complete reducibility property.
3. Let $\mathfrak{h}_{\mathbb{C}}$ denote the complexification of the Lie algebra of the Heisenberg group, namely the space of all complex 3×3 upper triangular matrices with zeros on the diagonal.

(a) Show that every maximal commutative subalgebra of $\mathfrak{h}_{\mathbb{C}}$ is two dimensional and contains the center of $\mathfrak{h}_{\mathbb{C}}$.
(b) Show that $\mathfrak{h}_{\mathbb{C}}$ does not have any Cartan subalgebras.

4. Give an example of a maximal commutative subalgebra of $\mathfrak{sl}(2;\mathbb{C})$ that is not a Cartan subalgebra.
5. Verify Proposition 6.18 by direct calculation in the case $\mathfrak{g} = \mathfrak{sl}(3;\mathbb{C})$, using the Cartan subalgebra $\mathfrak{h} = \text{span}\{H_1, H_2\}$.
6. Let $\mathfrak{g}$ denote the vector space of 3×3 complex matrices of the form

$$\begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix},$$

where A is a 2×2 matrix with trace zero and B is an arbitrary 2×1 matrix.
(a) Show that $\mathfrak{g}$ is a subalgebra of $M_3(\mathbb{C})$.
(b) Let X, Y, H, e_1, and e_2 be the following basis for $\mathfrak{g}$. We let X, Y, and H be the usual $\mathfrak{sl}(2;\mathbb{C})$ basis in the "A" slot, with $B = 0$. We let e_1 be the matrix with a 1 in the first slot in B and zeros everywhere else, and we let e_2 be the matrix with a 1 in the second slot of B and zeros everywhere else. Compute the commutation relations among these basis elements.
(c) Show that $\mathfrak{g}$ has precisely one nontrivial ideal, namely the span of e_1 and e_2.
Hint: First, determine the subspaces of $\mathfrak{g}$ that are invariant under the adjoint action of the $\mathfrak{sl}(2;\mathbb{C})$ algebra spanned by X, Y, and H, and then determine which of these subspaces are also invariant under the adjoint action of e_1 and e_2. In determining the $\mathfrak{sl}(2;\mathbb{C})$-invariant subspaces, use Exercise 15 of Chapter 4.
(d) Show that the one-dimensional subspace of $\mathfrak{g}$ spanned by the element H is a Cartan subalgebra of $\mathfrak{g}$ and determine the associated roots.
(e) Is $\mathfrak{g}$ semisimple?

7. Let $\mathfrak{g}$, $\mathfrak{k}$, K, $\mathfrak{t}$, and $\mathfrak{h}$ be as usual in this chapter. Consider a root α and the associated subalgebra $\mathfrak{s}^\alpha = \text{span}\{X_\alpha, Y_\alpha, H_\alpha\}$, where X_α and Y_α are chosen so that $Y_\alpha = X_\alpha^*$ (where X^* is defined by (6.2)). We use, as usual, an inner product on $\mathfrak{g}$ that is invariant under the adjoint action of K. Suppose that V is any subspace of $\mathfrak{g}$ that is invariant under the adjoint action of $\mathfrak{s}^\alpha$. Show that the orthogonal complement $V^\perp$ of V is also invariant under $\mathfrak{s}^\alpha$.
8. Let $\mathfrak{s}^\alpha$ be the complex Lie algebra spanned by X_α, Y_α, and H_α, which satisfy the usual $\mathfrak{sl}(2;\mathbb{C})$ commutation. Show that the elements

$$\begin{aligned} &X_\alpha - Y_\alpha, \\ &X_\alpha + Y_\alpha + iH_\alpha, \\ &X_\alpha + Y_\alpha - iH_\alpha \end{aligned}$$

are eigenvectors for the operator $\text{ad}_{X_\alpha} - \text{ad}_{Y_\alpha}$. Compute directly the quantity

$$\exp\left[\frac{\pi}{2}(\mathrm{ad}_{X_\alpha} - \mathrm{ad}_{Y_\alpha})\right](H_\alpha).$$

Compare (6.23) in the proof of Theorem 6.31. (Here π is the number 3.14..., not a representation.)

9. Show that the Hilbert–Schmidt inner product on $\mathsf{sl}(n;\mathbb{C})$, $\langle X, Y\rangle = \mathrm{trace}(X^*Y)$, is invariant under the adjoint action of $\mathsf{SU}(n)$.
10. Let $\mathfrak{g}$ be a complex semisimple Lie algebra contained in $\mathfrak{gl}(n;\mathbb{C})$, let $\mathfrak{k}$ be a compact real form of $\mathfrak{g}$, and let K be the compact subgroup of $\mathsf{GL}(n;\mathbb{C})$ whose Lie algebra is $\mathfrak{k}$. Now, let $\mathfrak{t}$ be a maximal commutative subalgebra of $\mathfrak{k}$ and let T be the connected Lie subgroup of K whose Lie algebra is $\mathfrak{t}$.
(a) Prove that T is closed in K.
Hint: Show that the closure of T is connected and commutative.
(b) Now, let $Z(T)$ be the *centralizer* of T in K, (i.e., the set of all A in K such that $AtA^{-1} = t$ for all $t \in T$). (The centralizer of T is the largest subgroup H of K that contains T and such that T is in the center of H.) Show that $Z(T)$ coincides with $Z(\mathfrak{t})$ as defined in Section 6.6.
(c) Let $N(T)$ denote the *normalizer* of T in K (i.e., the group of all A in K such that $AtA^{-1} \in T$ for all $t \in T$). (The normalizer of T is the largest subgroup H of K that contains T and such that T is normal in H.) Show that $N(T)$ coincides with $N(\mathfrak{t})$ as defined in Section 6.6.
Note: The group $T \subset K$ is a "maximal torus" and the customary definition of the Weyl group (from the compact group point of view) is $W = N(T)/Z(T)$. See Bröcker and tom Dieck (1985).
11. Continue with the notation of Exercise 10. Suppose that $\mathfrak{g} = \mathsf{sl}(n;\mathbb{C})$, $\mathfrak{k} = \mathsf{su}(n)$, and $\mathfrak{t}$ is the diagonal subalgebra of $\mathsf{su}(n)$. Show that T is indeed a torus (i.e., isomorphic to $S^1 \times S^1 \times \cdots \times S^1$). Show that $Z(\mathfrak{t}) = T$.
12. Consider the complex semisimple Lie algebra $\mathsf{so}(4;\mathbb{C})$ with compact real form $\mathsf{so}(4)$. Consider the space $\mathfrak{t}$ of matrices of the form

$$\begin{pmatrix} 0 & a & & \\ -a & 0 & & \\ & & 0 & b \\ & & -b & 0 \end{pmatrix} \tag{6.28}$$

with $a, b \in \mathbb{R}$. This is a maximal commutative subalgebra of $\mathsf{so}(4)$ (assume that this is so). Thus, the space $\mathfrak{h}$ of such matrices with $a, b \in \mathbb{C}$ is a Cartan subalgebra of $\mathsf{so}(4;\mathbb{C})$.
Now, consider the matrices of the form

$$\begin{pmatrix} 0 & C \\ -C^{tr} & 0 \end{pmatrix},$$

where C is one of the following 2×2 matrices:

$$C_1 = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix},$$
$$C_3 = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}, \quad C_4 = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

Show that each of the resulting elements of $\mathsf{so}(4;\mathbb{C})$ is a root vector and show that the corresponding roots are given by $\alpha_1 = i(a+b)$, $\alpha_2 = -i(a+b)$, $\alpha_3 = i(a-b)$, and $\alpha_4 = -i(a-b)$. Here, we are thinking of the roots as elements of $\mathfrak{h}^*$ and, for example, $i(a+b)$ means the linear functional that associates to the matrix (6.28) the number $i(a+b)$. Show that the roots $i(a+b)$ and $i(a-b)$ form a base for this root system.

Now, consider on $\mathsf{so}(4;\mathbb{C})$ the inner product $\langle X, Y\rangle = \mathrm{trace}(X^*Y)$, which is invariant under the adjoint action of $\mathsf{SO}(4)$. Use this inner product to identify $\mathfrak{h}^*$ with $\mathfrak{h}$ (as in Section 6.9) and compute the elements of $\mathfrak{h}$ that represent $\alpha_1, \alpha_2, \alpha_3$, and α_4 under this identification. Show that the elements of the base in the previous paragraph are orthogonal with respect to the given inner product.

13. Consider the complex semisimple Lie algebra $\mathsf{so}(5;\mathbb{C})$ with compact real for $\mathsf{so}(5)$. Consider the space $\mathfrak{t}$ of matrices of the form

$$\begin{pmatrix} 0 & a & & & \\ -a & 0 & & & \\ & & 0 & b & \\ & & -b & 0 & \\ & & & & 0 \end{pmatrix} \tag{6.29}$$

with $a, b \in \mathbb{R}$. This is a maximal commutative subalgebra of $\mathsf{so}(5)$ (assume that this is so). Thus, the space $\mathfrak{h}$ of such matrices with $a, b \in \mathbb{C}$ is a Cartan subalgebra of $\mathsf{so}(5;\mathbb{C})$.

Show that the matrices of the form

$$\begin{pmatrix} 0 & C & \\ -C^{tr} & 0 & \\ & & 0 \end{pmatrix}, \tag{6.30}$$

where C is one of the matrices in the previous problem, are root vectors, with roots given by the same formulas as in the previous problem. Show that matrices of the form

$$\begin{pmatrix} & & 1 & \\ & & \pm i & \\ & & 0 & \\ & & 0 & \\ -1 & \mp i & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} & & 0 & \\ & & 0 & \\ & & 1 & \\ & & \pm i & \\ 0 & 0 & -1 & \mp i \end{pmatrix} \tag{6.31}$$

are also root vectors with roots $\pm ia$ and $\pm ib$, respectively. Show that the roots $i(a-b)$ and ib form a base for this root system.

Now, as in the previous problem, use the inner product given by $\langle X, Y\rangle = \text{trace}(X^*Y)$ to identify $\mathfrak{h}^*$ with $\mathfrak{h}$. Show that the roots associated to the root vectors (6.30) have length $\sqrt{2}$ longer than the root vectors in (6.31). Show that the angle between the two elements of the base in the previous paragraph is 135°.

14. Consider the complex semisimple Lie algebra $\mathsf{sp}(2;\mathbb{C}) \subset M_4(\mathbb{C})$ and the compact real form $\mathsf{sp}(2) = \mathsf{sp}(2;\mathbb{C}) \cap \mathsf{u}(4)$. Consider the space $\mathfrak{t}$ of matrices of the form

$$\begin{pmatrix} ia & 0 & & \\ 0 & ib & & \\ & & -ia & 0 \\ & & 0 & -ib \end{pmatrix}$$

($a, b \in \mathbb{R}$). This is a maximal commutative subalgebra of $\mathsf{sp}(2)$ (assume that this is so). Thus, the space $\mathfrak{h}$ of such matrices with $a, b \in \mathbb{C}$ is a Cartan subalgebra of $\mathsf{sp}(2;\mathbb{C})$.

Show that the following matrices are root vectors for $\mathfrak{h}$ with roots $i(a+b)$, $-i(a+b)$, $i(a-b)$, and $-i(a-b)$, respectively:

$$\begin{pmatrix} & & 0 & 1 \\ & & 1 & 0 \\ 0 & 0 & & \\ 0 & 0 & & \end{pmatrix}, \quad \begin{pmatrix} & & 0 & 0 \\ & & 0 & 0 \\ 0 & 1 & & \\ 1 & 0 & & \end{pmatrix},$$
$$\begin{pmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 0 & 0 \\ & & -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & & \\ 1 & 0 & & \\ & & 0 & -1 \\ & & 0 & 0 \end{pmatrix}. \tag{6.32}$$

Show that the following matrices are root vectors for $\mathfrak{h}$ with roots $2ia$, $-2ia$, $2ib$, and $-2ib$, respectively:

$$\begin{pmatrix} & & 1 & 0 \\ & & 0 & 0 \\ 0 & 0 & & \\ 0 & 0 & & \end{pmatrix}, \quad \begin{pmatrix} & & 0 & 0 \\ & & 0 & 0 \\ 1 & 0 & & \\ 0 & 0 & & \end{pmatrix},$$
$$\begin{pmatrix} & & 0 & 0 \\ & & 0 & 1 \\ 0 & 0 & & \\ 0 & 0 & & \end{pmatrix}, \quad \begin{pmatrix} & & 0 & 0 \\ & & 0 & 0 \\ 0 & 0 & & \\ 0 & 1 & & \end{pmatrix}. \tag{6.33}$$

Note that the roots for $\mathsf{sp}(2;\mathbb{C})$ are given by the same formulas as for $\mathsf{so}(5;\mathbb{C})$, except that for $\mathsf{sp}(2;\mathbb{C})$, we have $\pm 2ia$ and $\pm 2ib$, whereas for $\mathsf{so}(5;\mathbb{C})$, we have $\pm ia$ and $\pm ib$. Show that the roots $i(a-b)$ and $2ib$ form a base for this root system.

Now, as in the previous two problems, use the inner product $\langle X, Y\rangle = \text{trace}(X^*Y)$ to identify $\mathfrak{h}^*$ with $\mathfrak{h}$. Show that the roots in (6.32) are $\sqrt{2}$

shorter than the roots in (6.33). Show that the angle between the two elements of the base in the previous paragraph is 135°.

15. Show that the subalgebra $\mathfrak{t}$ of $\mathfrak{su}(n)$ given in Section 6.9 is maximal commutative.
Hint: If X is any matrix in $\mathfrak{su}(n)$ that commutes with every H in $\mathfrak{t}$, write X as an element of $\mathfrak{t}$ plus a linear combination of the E_{kl}'s with $k \neq l$.
16. Suppose that $\mathfrak{k}$ is a compact real form of the complex semisimple Lie algebra $\mathfrak{sl}(n;\mathbb{C})$. Let K be the compact subgroup (Proposition 6.8) of $\mathsf{SL}(n;\mathbb{C})$ whose Lie algebra is $\mathfrak{k}$.
(a) Show that there exists an inner product on $\mathbb{C}^n$ that is invariant under the action of K.
(b) Show that $\mathfrak{k}$ consists precisely of those matrices that are skew self-adjoint with respect to this inner product and have trace zero.
(c) Show that there exists an element A of $\mathsf{SL}(n;\mathbb{C})$ such that $\mathrm{Ad}_A(\mathfrak{k}) = \mathfrak{su}(n)$.
This establishes the uniqueness result in Theorem 6.38 for the case $\mathfrak{g} = \mathfrak{sl}(n;\mathbb{C})$.
17. (a) Suppose that $X \in \mathfrak{sl}(n;\mathbb{C})$ is diagonalizable. Show that $\mathrm{ad}_X : \mathfrak{sl}(n;\mathbb{C}) \to \mathfrak{sl}(n;\mathbb{C})$ is diagonalizable.
(b) Suppose that $N \in \mathfrak{sl}(n;\mathbb{C})$ is nilpotent. Show that $\mathrm{ad}_N : \mathfrak{sl}(n;\mathbb{C}) \to \mathfrak{sl}(n;\mathbb{C})$ is nilpotent.
(c) Suppose that $X \in \mathfrak{sl}(n;\mathbb{C})$ is such that $\mathrm{ad}_X : \mathfrak{sl}(n;\mathbb{C}) \to \mathfrak{sl}(n;\mathbb{C})$ is diagonalizable. Show that X is diagonalizable.
Hint: What is the SN decomposition of X?
18. Suppose that $\mathfrak{h}$ is an arbitrary Cartan subalgebra of $\mathfrak{sl}(n;\mathbb{C})$.
(a) Show that the elements of $\mathfrak{h}$ are simultaneously diagonalizable.
(b) Show that there exists $g \in \mathsf{SL}(n;\mathbb{C})$ such that $g\mathfrak{h}g^{-1} = \mathfrak{h}_0$, where $\mathfrak{h}_0$ denotes the diagonal subalgebra of $\mathfrak{sl}(n;\mathbb{C})$.
(c) Show that there exists a compact real form $\mathfrak{k}$ of $\mathfrak{sl}(n;\mathbb{C})$ and a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$ such that $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$.
Use Exercise 17. This establishes the uniqueness result in Theorem 6.40 for the case $\mathfrak{g} = \mathfrak{sl}(n;\mathbb{C})$.
19. Let $\mathfrak{t}$ be an arbitrary maximal commutative subalgebra of $\mathfrak{su}(n)$.
(a) Show that the elements of $\mathfrak{t}$ are simultaneously diagonalizable.
(b) Show that there exists an element A of $\mathsf{SU}(n)$ such that $\mathrm{Ad}_A(\mathfrak{t})$ is the space of diagonal matrices in $\mathfrak{su}(n)$.
This (together with Exercise 16) establishes the uniqueness result in Theorem 6.39 for the case $\mathfrak{g} = \mathfrak{sl}(n;\mathbb{C})$.

7

Representations of Complex Semisimple Lie Algebras

In this chapter, we will study the finite-dimensional irreducible representations of a complex semisimple Lie algebra $\mathfrak{g}$. These will be classified by means of a "theorem of the highest weight." The theorem states that every irreducible representation has a (unique) highest weight, that two irreducible representations with the same highest weight are equivalent, and that the elements that actually arise as highest weights of irreducible representations are precisely the "dominant integral" elements.

Now that we have developed (in the previous chapter) the relevant structures for semisimple Lie algebras, most of the proof of the theorem of the highest weight goes precisely as in the case of $\mathfrak{sl}(3;\mathbb{C})$. Nevertheless, there is one part of the proof that cannot be done the way we did the $\mathfrak{sl}(3;\mathbb{C})$ case, namely showing that every dominant integral element actually arises as the highest weight of some irreducible representation. For this step, we need some method of constructing representations, in contrast to the rest of the proof, in which we assume that we are given a representation and we start analyzing it.

In the $\mathfrak{sl}(3;\mathbb{C})$ case, we constructed the representations by starting with the standard representation and the dual of the standard representation and then taking subspaces of tensor products of these two representations. Although similar methods can be used (as in Fulton and Harris (1991)) for other classical groups, this method does not work in general.

For a general semisimple Lie algebra, there are three standard methods of constructing the representations. The first is a purely Lie-algebraic approach using Verma modules. The second method constructs the representations as representations of the associated simply-connected compact group and uses the Peter–Weyl theorem and the Weyl character formula. The third method constructs the representations as representations of the complex Lie group G whose Lie algebra is $\mathfrak{g}$. In this approach, G acts on the space of holomorphic functions that transform in a certain way under the action of a certain subgroup B of G. This approach is called Borel–Weil theory.

We will give an essentially complete treatment of the Verma module approach. For the other two approaches, I provide a detailed outline with references for further reading.

In Chapter 8 we will work out examples for semisimple Lie algebras of rank 2 and rank 3.

We continue with the setting of the previous chapter. We consider a complex semisimple Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(n;\mathbb{C})$. We choose, once and for all, a compact real form $\mathfrak{k}$ of $\mathfrak{g}$ and a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$, and we consider the Cartan subalgebra $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$ of $\mathfrak{g}$. We also choose an inner product on $\mathfrak{g}$ that is invariant under the adjoint action of $K \subset \mathsf{GL}(n;\mathbb{C})$ and that takes real values on $\mathfrak{k}$.

We let $R \subset i\mathfrak{t} \subset \mathfrak{h}$ be the set of roots in the sense of Notation 6.24. We choose, once and for all, a base Δ for R (in the sense of Definition 6.35), the elements of which are called the positive simple roots. Every root is then either positive or negative (with respect to Δ) in the sense of Definition 6.35. We let W denote the Weyl group, which may be thought of (Theorem 6.33) as the group of linear transformations of $\mathfrak{h}$ generated by the reflections w_α, $\alpha \in R$. The set of roots has all the properties of a "root system," listed in Theorem 6.34.

We consider also the co-roots. For each α, there exist (Theorem 6.20) $X_\alpha \in \mathfrak{g}_\alpha$, $Y_\alpha \in \mathfrak{g}_{-\alpha}$, and $H_\alpha \in \mathfrak{h}$ such that $[H_\alpha, X_\alpha] = 2X_\alpha$, $[H_\alpha, Y_\alpha] = -2Y_\alpha$, and $[X_\alpha, Y_\alpha] = H_\alpha$. The element H_α is unique (independent of the choice of X_α and Y_α) and is called the co-root associated to the root α. According to Section 6.5, the roots and co-roots are related by the formulas

$$H_\alpha = 2\frac{\alpha}{\langle \alpha, \alpha \rangle} \tag{7.1}$$

and

$$\alpha = 2\frac{H_\alpha}{\langle H_\alpha, H_\alpha \rangle}. \tag{7.2}$$

In particular, $\langle \alpha, H_\alpha \rangle = 2$. The set of co-roots also constitutes a root system, and the set of H_α, $\alpha \in \Delta$, forms a base for the system of co-roots.

7.1 Integral and Dominant Integral Elements

Definition 7.1. *An element μ of $\mathfrak{h}$ is called an **integral element** if $\langle \mu, H_\alpha \rangle$ is an integer for each root α.*

As explained in the next section, the integral elements are precisely the elements of $\mathfrak{h}$ that arise as weights of finite-dimensional representations of $\mathfrak{g}$.

Proposition 7.2. *The set of integral elements is invariant under the action of the Weyl group.*

Proof. Suppose that $\mu \in \mathfrak{h}$ is an integral element and that w is an element of the Weyl group. Then, since the inner product on $\mathfrak{h}$ is invariant under the action of the Weyl group, we have for any root α, $\langle w \cdot \mu, H_\alpha \rangle = \langle \mu, w^{-1} \cdot H_\alpha \rangle$. Since the set of co-roots is invariant under the Weyl group, $w^{-1} \cdot H_\alpha$ is another co-root (namely $H_{w^{-1} \cdot \alpha}$) and, so, $\langle \mu, w^{-1} \cdot H_\alpha \rangle$ is an integer. This shows that $\langle w \cdot \mu, H_\alpha \rangle$ is an integer and thus that $w \cdot \mu$ is an integral element. □

Checking that $\langle \mu, H_\alpha \rangle$ is an integer for *every* root α is a rather tiresome process. Fortunately, it suffices to check just for the positive simple roots.

Theorem 7.3. *If μ is an element of $\mathfrak{h}$ for which $\langle \mu, H_\alpha \rangle$ is an integer for all positive simple roots α, then $\langle \mu, H_\alpha \rangle$ is an integer for all roots α and, thus, μ is an integral element.*

Proof. Suppose $\alpha_1, \ldots, \alpha_r$ are the positive simple roots. According to Proposition 6.37 (which is proved in Chapter 8), $H_{\alpha_1}, \ldots, H_{\alpha_r}$ form a base for the system of co-roots. This means that for any root α, the co-root H_α can be expressed as a linear combination of $H_{\alpha_1}, \ldots, H_{\alpha_r}$ with integer coefficients. Thus, if $\langle \mu, H_{\alpha_j} \rangle$ is an integer for each $j = 1, \ldots, r$, then $\langle \mu, H_\alpha \rangle$ is an integer for all roots α. □

Recalling the expression (7.1) for H_α in terms of α, we may restate Theorem 7.3 as follows.

Theorem 7.4. *An element μ of $\mathfrak{h}$ is integral if and only if*

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle}$$

is an integer for each positive simple root α.

Corollary 7.5. *Every root is an integral element.*

Recall now from elementary linear algebra that if μ and α are any two elements of an inner-product space, then the orthogonal projection of μ onto α is given by

$$\frac{\langle \alpha, \mu \rangle}{\langle \alpha, \alpha \rangle} \alpha.$$

Thus, we may reformulate the notion of an integral element yet again as follows.

Theorem 7.6. *An element μ of $\mathfrak{h}$ is integral if and only if the orthogonal projection of μ onto each positive simple root α is an integer or half-integer multiple of α.*

This characterization of the integral elements will help us visualize graphically what the set of integral elements looks like in examples. (See Sections 8.5 and 8.6.) All of these reformulations of the notion of an integral element

should not cause us to lose sight of the "real" definition of integrality, which is that $\langle \mu, H_\alpha \rangle$ be an integer for each positive simple root α and, therefore, by Proposition 6.37 for every root α. It is this form of integrality that explains why the weights of a finite-dimensional representation of $\mathfrak{g}$ must be integral, as we will see in the next section.

We now turn to the elements that will arise as the *highest* weights of finite-dimensional irreducible representations of $\mathfrak{g}$.

Definition 7.7. *An element μ of $\mathfrak{h}$ is called a* ***dominant integral element*** *if $\langle \mu, H_\alpha \rangle$ is a non-negative integer for each positive simple root α. Equivalently μ is a dominant integral element if*

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle}$$

is a non-negative integer for each positive simple root α.

If μ is dominant integral, then $\langle \mu, H_\alpha \rangle$ will automatically be a non-negative integer for each positive root α, not just the positive simple ones.

Definition 7.8. *The set of $\mu \in it \subset \mathfrak{h}$ such that $\langle \mu, \alpha \rangle \geq 0$ for all positive simple roots α is called the* ***closed fundamental Weyl chamber*** *relative to the given set of positive simple roots.*

The dominant integral elements are precisely those integral elements contained in the closed fundamental Weyl chamber. In the case of $\mathfrak{sl}(3;\mathbb{C})$, the fundamental Weyl chamber is a 60° sector—see Figure 5.2.

We have observed that every root is an integral element. It follows that any linear combination of roots with integer coefficients is also an integral element. We can ask whether the reverse holds: Is every integral element expressible as a linear combination of roots with integer coefficients? The answer in general is no. This matter is discussed further in Section 8.10.

7.2 The Theorem of the Highest Weight

We continue with the notation established at the beginning of this chapter. We begin with elementary properties of the representations of $\mathfrak{g}$.

Definition 7.9. *Suppose π is a finite-dimensional representation of $\mathfrak{g}$ on a vector space V. Then, $\mu \in \mathfrak{h}$ is called a* ***weight*** *for π if there exists a nonzero vector v in V such that*

$$\pi(H)v = \langle \mu, H \rangle v \tag{7.3}$$

for all $H \in \mathfrak{h}$. A nonzero vector v satisfying (7.3) is called a ***weight vector*** *for the weight μ, and the set of all vectors satisfying (7.3) (zero or nonzero) is called the* ***weight space*** *with weight μ. The dimension of the weight space is called the* ***multiplicity*** *of the weight.*

To understand this definition, suppose that $v \in V$ is a simultaneous eigenvector for each $\pi(H)$, $H \in \mathfrak{h}$. This means that for each $H \in \mathfrak{h}$, there is a number λ_H such that $\pi(H)v = \lambda_H v$. Since the representation $\pi(H)$ is linear in H, the λ_H's must depend linearly on H as well; that is, the map $H \to \lambda_H$ is a linear functional on $\mathfrak{h}$. Then (Section B.7), there is a unique element μ of $\mathfrak{h}$ such that $\lambda_H = \langle \mu, H \rangle$. Thus, a weight vector is nothing but a simultaneous eigenvector for all the $\pi(H)$'s and the vector μ is simply a convenient way of encoding the eigenvalues. Note that the roots (in the sense of Notation 6.24) are precisely the nonzero weights of the adjoint representation of $\mathfrak{g}$.

It is easily shown that two equivalent representations have the same weights and multiplicities.

Proposition 7.10. *If $\mu \in \mathfrak{h}$ is a weight of some finite-dimensional representation (π, V) of $\mathfrak{g}$, then μ is an integral element in the sense of Definition 7.1.*

Proof. Each co-root H_α is part of an $\mathfrak{sl}(2;\mathbb{C})$-subalgebra $\{X_\alpha, Y_\alpha, H_\alpha\}$ (Theorem 6.20). The restriction of π to this subalgebra is a finite-dimensional representation of $\mathfrak{sl}(2;\mathbb{C})$, and in any such representation, the eigenvalues of H_α must be integers (Theorem 4.12). If μ is a weight of π, then, by (7.3), $\langle \mu, H_\alpha \rangle$ is an eigenvalue for H_α in π, and, so, $\langle \mu, H_\alpha \rangle$ must be an integer. □

It is true, although by no means obvious, that every integral element actually arises as a weight of some finite-dimensional representation of $\mathfrak{g}$. See the discussion following Theorem 7.15.

We now observe, as in the $\mathfrak{sl}(3;\mathbb{C})$ case, that the root vectors shift the weights by the corresponding root.

Proposition 7.11. *Suppose that v is a weight vector with weight μ and suppose that X_α is an element of the root space $\mathfrak{g}_\alpha$. Then, for all H in $\mathfrak{h}$, we have*

$$\pi(H)\pi(X_\alpha)v = (\langle \mu, H \rangle + \langle \alpha, H \rangle)\pi(X_\alpha)v;$$

that is, either $\pi(X_\alpha)v$ is zero or $\pi(X_\alpha)v$ is a weight vector with weight $\mu + \alpha$.

Proof. This is proved in the same way as for the case of $\mathfrak{sl}(3;\mathbb{C})$. Since $[H, X_\alpha] = \langle \alpha, H \rangle X_\alpha$, we have

$$\begin{aligned} \pi(H)\pi(X_\alpha)v &= [\pi(X_\alpha)\pi(H) + \pi([H, X_\alpha])]v \\ &= [\pi(X_\alpha)\pi(H) + \langle \alpha, H \rangle \pi(X_\alpha)]v \\ &= [\langle \mu, H \rangle + \langle \alpha, H \rangle]\pi(X_\alpha)v. \end{aligned}$$

In the first equality, we have used that $[\pi(H), \pi(X_\alpha)] = \pi([H, X_\alpha])$. □

Proposition 7.12. *Every finite-dimensional representation (π, V) is the direct sum of its weight spaces; that is, the set of operators of the form $\pi(H)$, $H \in \mathfrak{h}$, are simultaneously diagonalizable in every finite-dimensional representation.*

Proof. By complete reducibility, the representation decomposes as a direct sum of irreducible representations, and so it suffices to prove the result in the irreducible case. Thus, we assume now that π is irreducible. Let U be the span of all the weight spaces in V; that is, U is the space of all vectors $u \in V$ such that u can be written as a linear combination of weight vectors. In light of Proposition B.14, U is actually the direct sum of all the weight spaces in V. Proposition B.10 tells us that any commuting family of operators on a finite-dimensional complex vector space has at least one simultaneous eigenvector. Applying this to the operators $\pi(H)$, $H \in \mathfrak{h}$, we see that V has at least one weight vector, which means that $U \neq \{0\}$.

I claim now that U is invariant under the action of $\mathfrak{g}$. Clearly, U is invariant under $\pi(H)$, $H \in \mathfrak{h}$, and by Proposition 7.11, U is invariant under each of the root spaces $\mathfrak{g}_\alpha$. Since $\mathfrak{g}$ is the direct sum of $\mathfrak{h}$ and the root spaces, U is invariant under $\mathfrak{g}$. Then, since we are assuming V is irreducible and since $U \neq \{0\}$, we must have $U = V$. □

Proposition 7.13. *For any finite-dimensional representation π of $\mathfrak{g}$, the weights of π and their multiplicity are invariant under the action of the Weyl group.*

Proof. Recall that we are thinking of the complex semisimple Lie algebra $\mathfrak{g}$ as sitting inside some $\mathfrak{gl}(n;\mathbb{C})$, that we have chosen a compact real form $\mathfrak{k}$ of $\mathfrak{g}$, and that K denotes the connected Lie subgroup of $\mathsf{GL}(n;\mathbb{C})$ whose Lie algebra is $\mathfrak{k}$. Saying that $\mathfrak{k}$ is a compact real form of $\mathfrak{g}$ means that there exists a simply-connected compact matrix Lie group K_1 whose Lie algebra $\mathfrak{k}_1$ is isomorphic to $\mathfrak{k}$. Since $\mathfrak{k}_1$ and $\mathfrak{k}$ are isomorphic, we can think of $\mathfrak{k}$ as being the Lie algebra of K or as being the Lie algebra of K_1. However, the groups K and K_1 need not be isomorphic. (For example, we may have $K = \mathsf{SO}(3)$ and $K_1 = \mathsf{SU}(2)$.) On the surface of things, it appears that the notion of the Weyl group might depend on whether we think of $\mathfrak{k}$ as the Lie algebra of K or of K_1. However, Theorem 6.33 tells us that we do get the same Weyl group either way, namely the group generated by the reflections w_α. With this in mind, we choose to think of $\mathfrak{k}$ as the Lie algebra of the simply-connected group K_1. Then, there is a representation Π of K_1 such that $\Pi(\exp X) = \exp \pi(X)$ for all X in $\mathfrak{k}$ (which we identify with $\mathfrak{k}_1$).

Now, let w be an element of W and let A be an element of $N(\mathfrak{t})$ that represents it. If v is a weight vector with weight μ, consider $\Pi(A)v$. We have

$$\begin{aligned}\pi(H)\Pi(A)v &= \Pi(A)\Pi(A)^{-1}\pi(H)\Pi(A)v \\ &= \Pi(A)\pi(\mathrm{Ad}_{A^{-1}}(H))v \\ &= \langle \mu, \mathrm{Ad}_{A^{-1}}(H)\rangle\, \Pi(A)v \\ &= \langle \mathrm{Ad}_A(\mu), H\rangle\, \Pi(A)v.\end{aligned}$$

In the second equality we have used Point 1 of Theorem 2.21 and in the last equality we have used the invariance of the inner product under the adjoint

action of K. This calculation shows that $\Pi(A)v$ is a weight vector with weight $\mathrm{Ad}_A(\mu) = w \cdot \mu$. The same line of reasoning shows that $\Pi(A)$ is an isomorphism between the weight space with weight μ and the weight space with weight $w \cdot \mu$, and, so, $w \cdot \mu$ is, again, a weight for π with the same multiplicity as μ. □

Definition 7.14. *Let μ_1 and μ_2 be two elements of $\mathfrak{h}$. Then, μ_1 is **higher** than μ_2 (or, equivalently, μ_2 is **lower** than μ_1) if there exist non-negative real numbers $a_1, \ldots, a_r$ such that*

$$\mu_1 - \mu_2 = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_r\alpha_r,$$

where $\{\alpha_1, \ldots, \alpha_r\} = \Delta$ is the set of positive simple roots. This relationship is written as $\mu_1 \succeq \mu_2$ or $\mu_2 \preceq \mu_1$.

*If π is a representation of $\mathfrak{g}$, then a weight μ_0 for π is said to be a **highest weight** if for all weights μ of π, $\mu \preceq \mu_0$.*

Theorem 7.15 (Theorem of the Highest Weight).

1. *Every irreducible representation has a highest weight.*
2. *Two irreducible representations with the same highest weight are equivalent.*
3. *The highest weight of every irreducible representation is a dominant integral element.*
4. *Every dominant integral element occurs as the highest weight of an irreducible representation.*

The proof of the first three points of the theorem is almost precisely as in the case of $\mathfrak{sl}(3;\mathbb{C})$. The proof of Point 4 is substantially more complicated than in the $\mathfrak{sl}(3;\mathbb{C})$ case and is discussed at length in the following sections.

It follows from Theorem 7.15 and properties of the Weyl group that every integral element occurs as a weight of some finite-dimensional irreducible representation of $\mathfrak{g}$. Specifically, if μ is an integral element then (Section 8.7) there exists $w \in W$ such that $\mu_0 := w \cdot \mu$ is a dominant integral element. Suppose V is the irreducible representation with highest weight μ_0. Then by Proposition 7.13, $\mu = w^{-1} \cdot \mu_0$ will be a weight of V.

Definition 7.16. *A representation (π, V) of $\mathfrak{g}$ is said to be a **highest weight cyclic representation with weight** μ_0 if there exists $v \neq 0$ in V such that*

1. *v is a weight vector with weight μ_0*
2. *$\pi(X_\alpha)v = 0$ for all positive roots α*
3. *the smallest invariant subspace of V containing v is all of V.*

*The vector v is called a **cyclic vector** for π.*

Proposition 7.17. *Let (π, V) be a highest weight cyclic representation of $\mathfrak{g}$ with weight μ_0. Then:*

1. *π has highest weight μ_0.*

2. *The weight space corresponding to the highest weight μ_0 is one dimensional.*

Proof. Let v be as in the definition. Let $\alpha_1, \ldots, \alpha_r$ be the positive simple roots and let $\alpha_{r+1}, \ldots, \alpha_m$ be the remaining positive roots (in any order). For each l, choose nonzero elements X_l in the root space $\mathfrak{g}_{\alpha_l}$ and Y_l in the root space $\mathfrak{g}_{-\alpha_l}$. Consider, then, the subspace U of V spanned by elements of the form

$$u = \pi(Y_{l_1})\pi(Y_{l_2}) \cdots \pi(Y_{l_N})v, \tag{7.4}$$

We want to show that U is invariant under the action of $\mathfrak{g}$. To show this, it suffices to show that U is invariant under each $\pi(H)$, $H \in \mathfrak{h}$, and that for each positive root α_l, U is invariant under $\pi(X_l)$ and $\pi(Y_l)$. We consider the basis for $\mathfrak{g}$ (as a vector space) consisting of $H_{\alpha_1}, \ldots, H_{\alpha_r}$ (where $\alpha_1, \ldots, \alpha_r$ are the positive simple roots) together with the elements $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_m$.

Now, we apply to an element u of the form (7.4) an operator of the form $\pi(H_{\alpha_l})$, $\pi(X_l)$, or $\pi(Y_l)$. We apply Lemma 5.14 to rewrite the resulting vector as a linear combination of terms, each of which has all of the factors of $\pi(X_l)$ to the right (acting first on v), followed by the factors of $\pi(H_{\alpha_l})$, followed by the factors of $\pi(Y_l)$. As in the $\mathfrak{sl}(3;\mathbb{C})$ case, any term that actually has any factors of $\pi(X_l)$ acting on v will be zero. In the remaining terms, each factor of $\pi(H_{\alpha_l})$ will hit v first and will give back just a constant times v. Thus, only the factors of $\pi(Y_l)$ will remain and we obtain a linear combination of factors of the form (7.4). Thus, the vector that we obtain by applying $\pi(H_{\alpha_l})$, $\pi(X_l)$, or $\pi(Y_l)$ to u is, again, in the space U.

The space U is invariant, and, by definition, it contains the vector v (taking $N = 0$ in (7.4)). So, by the definition of a highest weight cyclic representation, U must be all of V. Thus, every element of V is a linear combination of vectors of the form (7.4). However, by Proposition 7.11, each vector of the form (7.4) is either zero or a weight vector with weight $\mu_0 - \alpha_{l_1} - \cdots - \alpha_{l_N}$, which is lower than or equal to μ_0. So, μ_0 is the highest weight for V.

Furthermore, every element of V is a linear combination of v itself (the terms with $N = 0$) and weight vectors with weight strictly lower than μ_0 (the terms with $N > 0$). It then follows from Proposition B.14 that the only weight vectors with weight μ_0 are multiples of v, and, so, the weight space with weight μ_0 is one dimensional. □

Proposition 7.18. *Every irreducible representation of $\mathfrak{g}$ is a highest weight cyclic representation, with a unique highest weight μ_0.*

Proof. Uniqueness is immediate, since by the previous proposition, μ_0 is the highest weight, and two distinct weights cannot both be highest.

We have already shown that every irreducible representation is the direct sum of its weight spaces. Since the representation is finite dimensional, there can be only finitely many weights. It follows that there must be a maximal weight (i.e., a weight μ_0 such that there is no weight $\mu \neq \mu_0$ with $\mu \succeq \mu_0$). That being the case, we must have

$$\pi(X_\alpha)v = 0$$

for each element X_α of a root space $\mathfrak{g}_\alpha$ corresponding to a positive root α. (If not, then $\pi(X_\alpha)v$ would be a weight vector with weight $\mu_0 + \alpha \succeq \mu_0$.)

Since π is assumed irreducible, the smallest invariant subspace containing v must be the whole space; therefore, the representation is highest weight cyclic. □

Proposition 7.19. *Every highest weight cyclic representation of* $\mathfrak{g}$ *is irreducible.*

Proof. Let (π, V) be a highest weight cyclic representation with highest weight μ_0 and cyclic vector v. By complete reducibility, V decomposes as a direct sum of irreducible representations

$$V \cong \bigoplus_i V_i.$$

By Proposition 7.12, each of the V_i's is the direct sum of its weight spaces. Thus, since the weight μ_0 occurs in V, it must occur in some V_i. On the other hand, Proposition 7.17 says that the weight space corresponding to μ_0 is one dimensional; that is, v is (up to a constant) the *only* vector in V with weight μ_0. Thus, V_i must contain v. However, then, V_i is an invariant subspace containing v, so $V_i = V$. Thus, there is only one term in the sum (5.9), and V is irreducible. □

Proposition 7.20. *Two irreducible representations of* $\mathfrak{g}$ *with the same highest weight are equivalent.*

Proof. We now know that a representation is irreducible if and only if it is highest weight cyclic. Suppose that (π, V) and (σ, X) are two such representations with the same highest weight μ_0. Let v and w be highest weight cyclic vectors for V and X, respectively. Now, consider the representation $V \oplus X$, and let U be smallest invariant subspace of $V \oplus X$ that contains the vector (v, w).

The weights occurring in $V \oplus X$ are simply the weights of V together with the weights of X. This means that μ_0 is the highest weight occurring in $V \oplus X$. Since (v, w) is a weight vector with weight μ_0 and since this vector generates U (by definition), U is a highest weight cyclic representation, and, therefore, irreducible by Proposition 7.19. Consider the two "projection" maps $P_1 : V \oplus X \to V$, $P_1(v, w) = v$ and $P_2 : V \oplus X \to X$, $P_1(v, w) = w$. It is easy to check that P_1 and P_2 are intertwining maps of representations. Therefore, the restrictions of P_1 and P_2 to $U \subset V \oplus X$ will also be intertwining maps.

Clearly, neither $P_1|_U$ nor $P_2|_U$ is the zero map (since both are nonzero on (v, w)). Moreover, U, V, and X are all irreducible. Therefore, by Schur's Lemma, $P_1|_U$ is an isomorphism of U with V, and $P_2|_U$ is an isomorphism of U with X. Thus, $V \cong U \cong X$. □

Proposition 7.21. *If π is an irreducible representation of $\mathfrak{g}$, then the highest weight μ_0 of π is a dominant integral element.*

Proof. We know already that *all* of the weights of π (not just the highest weight) must be integral. Now, suppose that μ_0 is the highest weight of π and that v is a nonzero weight vector with weight μ_0. Then, $\pi(X_\alpha)v = 0$ for all positive simple roots α (otherwise, $\pi(X_\alpha)v$ would be a weight vector with weight higher than μ_0). Now, consider the subalgebra $\mathfrak{s}^\alpha = \{X_\alpha, Y_\alpha, H_\alpha\}$ of $\mathfrak{g}$, which is isomorphic to $\mathfrak{sl}(2;\mathbb{C})$. Then, v is an eigenvector for $\pi(H_\alpha)$ with eigenvalue $\mu_0(H_\alpha)$, and v is annihilated by $\pi(X_\alpha)$. By Theorem 4.12, this can occur only if $\mu_0(H_\alpha)$ is a non-negative integer. Thus, μ_0 is dominant integral. □

We have now completed the proof of Theorem 7.15, except for Point 4, namely that every dominant integral element arises as the highest weight of some irreducible representation. The strategy we used to prove this in the case of $\mathfrak{sl}(3;\mathbb{C})$ does not work in general. We devote the next three sections to a discussion of three different proofs of Point 4.

7.3 Constructing the Representations I: Verma Modules

We now turn to a discussion of the three standard methods of constructing an irreducible finite-dimensional representation having a given dominant integral element as its highest weight, namely Verma modules, the Peter–Weyl theory, and the Borel–Weil theory. We begin, in this section, with the Verma module approach. Given any μ in $\mathfrak{h}$, we will construct a representation V_μ called a Verma module. ("Module" is just another word for a representation.) Here, μ can truly be *any* element of $\mathfrak{h}$, not necessarily dominant and not necessarily integral. The catch is that the Verma module V_μ is always infinite dimensional, even if μ is dominant integral. We will see eventually that if μ is dominant integral, then V_μ contains an invariant subspace U_μ such that the quotient space V_μ/U_μ is finite dimensional and irreducible and has highest weight μ.

7.3.1 Verma modules

The Verma module is constructed as follows. Let $\mathfrak{n}^+$ be the subspace of $\mathfrak{g}$ spanned by the root spaces $\mathfrak{g}_\alpha$ where α is a positive root. This is a subalgebra of $\mathfrak{g}$ since $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$ and if α and β are positive roots, then $\mathfrak{g}_{\alpha+\beta}$ is either zero or is a root space corresponding to the positive root $\alpha+\beta$. Similarly, let $\mathfrak{n}^-$ be the span of the root spaces corresponding to negative roots, which is also a subalgebra. Then, $\mathfrak{g}$ decomposes as a direct sum (in the vector space sense) of $\mathfrak{h}$, $\mathfrak{n}^+$, and $\mathfrak{n}^-$. Now, let μ be any element of $\mathfrak{h}$. We want to construct an (infinite-dimensional) representation V_μ of $\mathfrak{g}$ having highest weight μ. We first describe V_μ as a vector space and then describe the action of $\mathfrak{g}$ on it.

As in the previous section, we let $\alpha_1, \ldots, \alpha_r$ be the positive simple roots and we let $\alpha_{r+1}, \ldots, \alpha_m$ be the remaining positive roots. We choose nonzero elements $X_l \in \mathfrak{g}_{\alpha_l}$ and $Y_l \in \mathfrak{g}_{-\alpha_l}$, $1 \leq l \leq m$. We begin with a vector v_0 that will be our highest weight vector. Then, the rest of V_μ will be finite linear combinations of vectors of the form

$$\pi_\mu(Y_{l_1})\pi_\mu(Y_{l_2}) \cdots \pi_\mu(Y_{l_N})v_0. \tag{7.5}$$

There are certain dependence relations among such vectors that are forced on us by the commutation relations in the subalgebra $\mathfrak{n}^-$. For example, if π_μ is going to be a representation, then we must have

$$\pi_\mu(Y_k)\pi_\mu(Y_l) = \pi_\mu(Y_l)\pi_\mu(Y_k) + \pi_\mu([Y_k, Y_l]),$$

where $[Y_k, Y_l]$ is again in $\mathfrak{n}^-$ and, therefore, expressible as a linear combination of the Y_j's. Thus,

$$\pi_\mu(Y_k)\pi_\mu(Y_l)v_0 = \pi_\mu(Y_l)\pi_\mu(Y_k)v_0 + \sum_{j=1}^{m} c_{klj}\, \pi_\mu(Y_j)v_0.$$

We construct V_μ as a vector space by imposing *only* those dependence relations among vectors of the form (7.5) that are forced on us by the commutation relations of $\mathfrak{n}^-$. As a vector space, V_μ is isomorphic to the "universal enveloping algebra" $U(\mathfrak{n}^-)$ of $\mathfrak{n}^-$. (See Section 17.2 of Humphreys (1972).) It follows from this that V_μ is always infinite-dimensional.

We now want to describe an action of $\mathfrak{g}$ on this space. For $Y \in \mathfrak{n}^-$, $\pi_\mu(Y)$ acts in the only possible way, namely by adding on one more factor of $\pi_\mu(Y)$ on the left. For the action of $\mathfrak{h}$, we decree that v_0 be a weight vector with weight μ:

$$\pi_\mu(H)v_0 = \langle \mu, H \rangle\, v_0, \quad H \in \mathfrak{h}. \tag{7.6}$$

By Proposition 7.11 (the proof of which is perfectly valid even for infinite-dimensional representations), each $\pi_\mu(Y_l)$ lowers the weight of v_0 by α_l, and so we must have

$$\begin{aligned} &\pi_\mu(H)\pi_\mu(Y_{l_1})\pi_\mu(Y_{l_2}) \cdots \pi_\mu(Y_{l_N})v_0 \\ &= (\mu(H) - \alpha_{l_1}(H) - \cdots - \alpha_{l_N}(H))\pi_\mu(Y_{l_1})\pi_\mu(Y_{l_2}) \cdots \pi_\mu(Y_{l_N})v_0. \end{aligned} \tag{7.7}$$

It remains then to describe the action of $\mathfrak{n}^+$ on $\mathfrak{g}$. If μ is actually going to be the highest weight occurring in V_μ, then we must have

$$\pi_\mu(X)v_0 = 0 \tag{7.8}$$

for all $X \in \mathfrak{n}^+$. Then, if we want to apply $\pi_\mu(X)$, $X \in \mathfrak{n}^+$, to an element of the form (7.5), we apply Lemma 5.14. The lemma allows us to rewrite

$$\pi_\mu(X)\pi_\mu(Y_{l_1})\pi_\mu(Y_{l_2}) \cdots \pi_\mu(Y_{l_N}) \tag{7.9}$$

as a linear combination of terms in which the elements of $\mathfrak{n}^+$ are to the right (acting first), the elements of $\mathfrak{h}$ are next, and the elements of $\mathfrak{n}^-$ are to the left (acting last). If we apply all of the resulting terms to v_0, any terms which actually contain any factors from $\mathfrak{n}^+$ must be zero, by (7.8). All of the other terms involve only factors from $\mathfrak{h}$ and $\mathfrak{n}^-$, with the elements of $\mathfrak{h}$ acting first. When we apply such a term to v_0, the factors from $\mathfrak{h}$ simply give constants, because of (7.6). This leaves us again with a linear combination of terms of the form (7.5), which means that we have a constructive procedure for determining how $\pi_\mu(X)$ acts on V_μ.

It is not completely clear that this procedure really yields a well-defined representation of $\mathfrak{g}$. The action of $\mathfrak{n}^+$ is the most problematic in this regard; there are many different ways to commute the factors in (7.9) into the desired form and one needs to know that the value of

$$\pi_\mu(X)\pi_\mu(Y_{l_1})\pi_\mu(Y_{l_2})\cdots\pi_\mu(Y_{l_N})v_0$$

is the same no matter which way is used. Nevertheless, fairly elementary algebraic means (Section 20.3 of Humphreys (1972)) can be used to show that the Verma module is well defined.

It is important to note that the Verma module is a representation of the Lie algebra $\mathfrak{g}$ only—there is no associated representation of the simply-connected compact group K. Although in the finite-dimensional case that every representation of $\mathfrak{g}$ comes from a representation of K, this result does not generalize to the infinite-dimensional case. What goes wrong is that in the infinite-dimensional case, the exponential of an operator may not be defined, because the series defining the exponential may not converge in any reasonable sense. We will revisit this issue in the next subsection.

7.3.2 Irreducible quotient modules

The good thing about Verma modules is that it is fairly easy to prove they exist. The bad thing is that they are always infinite dimensional, even when the highest weight μ is dominant integral. The strategy for constructing finite-dimensional representations is first to show that every Verma module has a largest proper invariant subspace U_μ and that the quotient space V_μ/U_μ is irreducible with highest weight μ. This much is true for any μ in $\mathfrak{h}$. Then, one shows that in the case that μ is dominant integral, the quotient space is finite dimensional.

Let us look into this strategy in greater detail. The invariant subspace U_μ is defined as follows. It follows from (7.7) and Proposition B.14 that V_μ is a direct sum of its weight spaces. Thus, for any vector v in V_μ, it makes sense to talk about the component of v in the one-dimensional subspace spanned by v_0, which we refer to as the v_0-component of v.

Definition 7.22. *Given a Verma module V_μ let U_μ be the subspace of V_μ consisting of all vectors v such that the v_0-component of v is zero and such that the v_0-component of*

$$\pi_\mu(X^1)\cdots\pi_\mu(X^l)v$$

is also zero for any collection of vectors $X^1,\ldots,X^l$ *in* $\mathfrak{n}^+$.

Certainly the zero vector is in U_μ; for some $\mathfrak{g}$'s and μ's, it happens that $U_\mu = \{0\}$.

Proposition 7.23. *The space* U_μ *is an invariant subspace for the action of* $\mathfrak{g}$.

Proof. Suppose that v is in U_μ and that Z is some element of $\mathfrak{g}$. We want to show that $\pi_\mu(Z)v$ is also in U_μ. Thus, we consider

$$\pi_\mu(X^1)\cdots\pi_\mu(X^l)\pi_\mu(Z)v \tag{7.10}$$

and we must show that the v_0-component of this vector is zero. Using Lemma 5.14, we may rewrite the vector in (7.10) as a linear combination of vectors of the form

$$\pi_\mu(Y^1)\cdots\pi_\mu(Y^j)\pi_\mu(H^1)\cdots\pi_\mu(H^k)\pi_\mu(\tilde{X}^1)\cdots\pi_\mu(\tilde{X}^m)v, \tag{7.11}$$

where the Y's are in $\mathfrak{n}^-$, the H's are in $\mathfrak{h}$, and the $\tilde{X}$'s are in $\mathfrak{n}^+$. However, since v is in U_μ, the v_0-component of

$$\pi_\mu(\tilde{X}^1)\cdots\pi_\mu(\tilde{X}^m)v \tag{7.12}$$

is zero, and thus this vector is a linear combination of weight vectors with weight lower than μ. Then, applying elements of $\mathfrak{h}$ and $\mathfrak{n}^-$ to the vector in (7.12) will only keep the weights the same or lower them. Thus, the v_0-component of the vector in (7.11), and hence also the v_0-component of the vector in (7.10), is zero. This shows that $\pi_\mu(Z)v$ is, again, in U_μ. □

If V is any vector space and U is a subspace of V, then one can form the **quotient space** V/U. The construction of V/U is analogous to the construction of quotient groups, as described in Appendix A. We define two elements of V to be equivalent if their difference is an element of U and then V/U is defined to be the set of equivalence classes. Because U is a subspace, the vector space operations on V (addition and scalar multiplication) "descend" unambiguously to equivalence classes and make V/U into a vector space. If V carries a representation of some Lie algebra $\mathfrak{g}$ and if U is an invariant subspace of V, then the action of $\mathfrak{g}$ on V descends to an action on V/U and, thus, the space V/U carries a representation of $\mathfrak{g}$, called the **quotient representation**. We apply the quotient construction to the Verma module V_μ and the invariant subspace U_μ.

Proposition 7.24. *The quotient space* V_μ/U_μ *is an irreducible representation of* $\mathfrak{g}$.

Proof. A simple argument shows that the invariant subspaces of the representation V_μ/U_μ are in one-to-one correspondence with the invariant subspaces of V_μ that contain U_μ. So, to prove that V_μ/U_μ is irreducible is equivalent to showing that any invariant subspace of V_μ that contains U_μ is either U_μ or V_μ. Suppose, then, that W is an invariant subspace that contains U_μ and suppose that $W \neq U_\mu$ (i.e., that W contains at least one vector v not contained in U_μ). This means that W contains a vector $u = \pi_\mu(X^1)\cdots\pi_\mu(X^l)v$ whose v_0-component is nonzero.

I claim then that W must contain v_0 itself. To see this, we decompose u as a nonzero multiple of v_0 plus a sum of weight vectors corresponding to weights $\lambda \neq \mu$. Since $\lambda \neq \mu$, we can find H in $\mathfrak{h}$ with $\langle \lambda, H\rangle \neq \langle \mu, H\rangle$ and then we may apply to u the operator $\pi_\mu(H) - \langle \lambda, H\rangle I$. This operator will keep us in W and will "kill" the component of u that is in the weight space corresponding to the weight λ while leaving the v_0-component of u nonzero. We then continue applying operators of this form until we have killed all the components of u in weight spaces different from μ, giving us a nonzero multiple of v_0.

This means that W contains v_0 and, therefore (in light of (7.5)), all of V_μ. So, any invariant subspace of V_μ that properly contains U_μ must be equal to V_μ. This shows that V_μ/U_μ is irreducible. □

Since for each $u \in U_\mu$ the v_0-component of u is zero, it is not hard to see that the quotient space V_μ/U_μ still has highest weight μ. So, for any μ in $\mathfrak{h}$ (not necessarily dominant or integral), we have a method of constructing an irreducible representation of $\mathfrak{g}$ with highest weight μ. Of course, we do not know that this representation is finite dimensional; indeed, it cannot be finite dimensional unless μ is dominant integral. Therefore, the crucial last step in the argument is to show that in the dominant integral case, the quotient space V_μ/U_μ is finite dimensional.

7.3.3 Finite-dimensional quotient modules

The way we will prove finite dimensionality is to show that there is an action of the Weyl group on V_μ/U_μ that transforms the weights in the same way as in the finite-dimensional case. This will show that the set of weights for V_μ/U_μ is invariant under the action of the Weyl group on $\mathfrak{h}$. Meanwhile, if μ is dominant integral, then all of the weights of V_μ/U_μ must be integral (since they are of the form in the right-hand side of (7.7)), and all the weights must be lower than μ. However, standard Weyl group theory implies that there are only finitely many integral elements λ with the property that $w \cdot \lambda$ is lower than μ for all $w \in W$. So, if we can show that the Weyl group acts on V_μ/U_μ, then we will conclude that there are only finitely many weights in V_μ/U_μ. Since (even in the Verma module) each weight has finite multiplicity, this will show that V_μ/U_μ is finite dimensional. (Note that the set of weights for the Verma module itself is never invariant under the action of W on $\mathfrak{h}$.)

How, then, do we construct an action of the Weyl group on V_μ/U_μ? Recall that in the finite-dimensional case we exponentiate each representation π of

$\mathfrak{g}$ to get a representation Π of the simply-connected compact group K. The action of the Weyl group is then obtained by restricting Π to the subgroup $N(\mathfrak{t}) \subset K$. In the infinite-dimensional case, the exponential of an operator is not necessarily well defined (since the series for the exponential may not converge) and, so, we cannot, in general, obtain a representation of the group Π. If μ is dominant integral, then we will eventually conclude that V_μ/U_μ is finite dimensional, but, of course, we are not allowed to assume that at this stage.

This means that we need a method of exponentiating operators that can be used in a possibly infinite-dimensional space. To do this, we introduce the concept of a **locally nilpotent** operator. A linear operator X on an arbitrary vector space V is said to be locally nilpotent if for each $v \in V$, there exists a positive integer k such that $X^k v = 0$. If V is finite dimensional, then a locally nilpotent operator must actually be nilpotent, that is, there must exist a single k such that $X^k v = 0$ for all v. In the infinite-dimensional case, the value of k depends on v and there may be no single value of k that works for all v. If X is locally nilpotent, then we define e^X to be the operator satisfying

$$e^X v = \sum_{k=0}^{\infty} \frac{X^k}{k!} v,$$

where for each $v \in V$ the series on the right terminates.

Proposition 7.25. *For each positive simple root $\alpha \in \Delta$, let X_α be an element of $\mathfrak{g}_\alpha$ and let Y_α be an element of $\mathfrak{g}_{-\alpha}$. If μ is dominant integral, then X_α and Y_α act in a locally nilpotent fashion on the quotient space V_μ/U_μ.*

We will give the proof of this result at the end of this subsection. Let us now continue with the argument for the finite dimensionality of V_μ/U_μ.

Proposition 7.26. *If μ is dominant integral, then the set of weights for V_μ/U_μ is invariant under the action of the Weyl group on $\mathfrak{h}$.*

Proof. We make use of a result from Weyl group theory, namely that W is generated by the reflections w_α, where α ranges over the set Δ of positive simple roots. (See Section 8.7.) So, it suffices to show that the set of weights is invariant under each w_α, $\alpha \in \Delta$.

Let $\tilde{\pi}_\mu$ denote the action of $\mathfrak{g}$ on the quotient space V_μ/U_μ. For each positive simple root α, let $X_\alpha \in \mathfrak{g}_\alpha$ and $Y_\alpha \in \mathfrak{g}_{-\alpha}$ be as in Theorem 6.20. Since, by Proposition 7.25, $\tilde{\pi}_\mu(X_\alpha)$ and $\tilde{\pi}_\mu(Y_\alpha)$ are locally nilpotent, it makes sense to exponentiate these operators. Define, then, operators B_α on V_μ/U_μ by

$$B_\alpha = e^{\tilde{\pi}_\mu(X_\alpha)} e^{-\tilde{\pi}_\mu(Y_\alpha)} e^{\tilde{\pi}_\mu(X_\alpha)}.$$

This operator is going to describe the action of the Weyl group element w_α on V_μ/U_μ. (Compare the expression (6.24) for the B_α's following Theorem 6.31.)

Assume that v is a weight vector in V_μ/U_μ with weight λ. We want to prove, then, that $B_\alpha v$ is a weight vector with weight $w_\alpha \cdot \lambda$, where, as usual, w_α is the reflection about the hyperplane perpendicular to α. To do this, it suffices to show that

$$\tilde{\pi}_\mu(H)B_\alpha = B_\alpha \tilde{\pi}_\mu(w_\alpha \cdot H).$$

This is really just a Lie algebra calculation; if it is true in $\mathsf{sl}(2;\mathbb{C})$, then it is true here as well.

To be a bit more precise about this, let $\tilde{X}_\alpha$, $\tilde{Y}_\alpha$, and $\tilde{H}$ stand for $\tilde{\pi}_\mu(X_\alpha)$, $\tilde{\pi}_\mu(Y_\alpha)$, and $\tilde{\pi}_\mu(H)$, respectively. Then, we have

$$\tilde{H}e^{\tilde{X}_\alpha}e^{-\tilde{Y}_\alpha}e^{\tilde{X}_\alpha} = e^{\tilde{X}_\alpha}e^{-\tilde{Y}_\alpha}e^{\tilde{X}_\alpha}\mathrm{Ad}_{e^{-\tilde{X}_\alpha}}\mathrm{Ad}_{e^{\tilde{Y}_\alpha}}\mathrm{Ad}_{e^{-\tilde{X}_\alpha}}(\tilde{H}).$$

Now, the relationship between Ad and ad still holds for locally nilpotent operators in the infinite-dimensional case (think of the power series argument in Exercise 19 of Chapter 2) and, so,

$$\mathrm{Ad}_{e^{-\tilde{X}_\alpha}}\mathrm{Ad}_{e^{\tilde{Y}_\alpha}}\mathrm{Ad}_{e^{-\tilde{X}_\alpha}}(\tilde{H}) = e^{-\mathrm{ad}_{\tilde{X}_\alpha}}e^{\mathrm{ad}_{\tilde{Y}_\alpha}}e^{-\mathrm{ad}_{\tilde{X}_\alpha}}(\tilde{H}). \tag{7.13}$$

If H is such that $\langle \alpha, H\rangle = 0$, then (7.13) is simply equal to $\tilde{H}$. If $H = H_\alpha$, then all of (7.13) is taking place in a three-dimensional Lie algebra isomorphic to $\mathsf{sl}(2;\mathbb{C})$; thus, the answer is the same as in the $\mathsf{sl}(2;\mathbb{C})$ case, namely $-\tilde{H}_\alpha$. In either case (check), (7.13) is equal to $\tilde{\pi}_\mu(w_\alpha \cdot H)$. (Compare Exercise 8 in Chapter 6.) □

In the dominant integral case, the weights for V_μ/U_μ are invariant under the action of the Weyl group and all of the weights are integral (since the weights that occur differ from μ by an integer linear combination of roots). A standard result from Weyl group theory (see Section 8.7) says that there are only finitely many integral elements λ such that $w \cdot \lambda$ is lower than μ for all $w \in W$. So, we conclude that if μ is dominant integral, then there are only finitely many weights in V_μ/U_μ.

Meanwhile, for any $\mu \in \mathfrak{h}$, we know that V_μ/U_μ has at least one weight space, namely the one with weight μ. (This weight space survives the passage from V_μ to V_μ/U_μ because the elements of U_μ have no v_0-component.) Since, as we have shown, V_μ/U_μ is irreducible, the same argument as in the finite-dimensional case shows that V_μ/U_μ is the direct sum of its weight spaces. Furthermore, all of the weights for V_μ, and so also for V_μ/U_μ, have finite multiplicity.

We conclude, then, that in the dominant integral case, V_μ/U_μ is the direct sum of its weight spaces, there are only finitely many of these weight spaces, and each of the weight spaces has finite dimension. This shows that V_μ/U_μ is finite dimensional and establishes that each dominant integral element arises as the highest weight of a finite-dimensional irreducible representation of $\mathfrak{g}$.

It now remains only to provide the proof of Proposition 7.25.

Proof. As in the previous proof, we use $\tilde{X}$ as an abbreviation for $\tilde{\pi}_\mu(X)$, for any $X \in \mathfrak{g}$. We also make use of the following standard result from the theory of semisimple Lie algebras: The subalgebra of $\mathfrak{g}$ generated by the spaces $\mathfrak{g}_\alpha$, with α in Δ, is equal to the span of the spaces $\mathfrak{g}_\alpha$, where α ranges over all of R^+. This follows from Proposition 8.4(d) and the corollary to Lemma 10.2A of Humphreys (1972). (See also Proposition 14.2 in Humphreys (1972).) For each root α, we let $\mathfrak{s}^\alpha$ denote the three-dimensional subalgebra $\{X_\alpha, Y_\alpha, H_\alpha\}$ given by Theorem 6.20.

Step 1. For each $\alpha \in \Delta$, $\tilde{X}_\alpha$ is locally nilpotent. Every vector $v \in V_\mu$, and so also every vector in V_μ/U_μ, is a finite linear combination of vectors of the form (7.5) and (by (7.7)) these vectors are weight vectors. Applying $\tilde{X}_\alpha$ repeatedly will raise all the weights until they are no longer lower than μ and, at that point, $\tilde{X}_\alpha^k v$ must be zero.

Step 2. For each positive simple root α, there exists a nonzero finite-dimensional subspace of V_μ/U_μ that is invariant under $\mathfrak{s}^\alpha$. Let

$$m = \langle \mu, H_\alpha \rangle,$$

which is a non-negative integer because μ is dominant integral. Now, consider the vectors $\tilde{Y}_\alpha^k v_0$, $k = 0, 1, 2, \ldots$. Then, from the calculations in Chapter 4, we have

$$\tilde{H}_\alpha \tilde{Y}_\alpha^k v_0 = (m - 2k)\tilde{Y}_\alpha^k v_0, \tag{7.14}$$

$$\tilde{X}_\alpha \tilde{Y}_\alpha^k v_0 = k(m + 1 - k)\tilde{Y}_\alpha^{k-1} v_0. \tag{7.15}$$

In particular, $\tilde{X}_\alpha \tilde{Y}_\alpha^{m+1} v_0 = 0$.

Now, consider $\beta \in \Delta$ with $\beta \neq \alpha$. Then, I claim that $[\tilde{X}_\beta, \tilde{Y}_\alpha] = 0$. To see this, note that $Y_\alpha \in \mathfrak{g}_{-\alpha}$ and $X_\beta \in \mathfrak{g}_\beta$, and, therefore, $[X_\beta, Y_\alpha] \in \mathfrak{g}_{\beta-\alpha}$. However, $\beta - \alpha$ is nonzero and cannot be a root. After all, every root has a unique expansion in terms of elements of Δ and in this expansion all nonzero coefficients have the same sign, whereas $\beta - \alpha$ has one positive coefficient and one negative coefficient. It follows that $\mathfrak{g}_{\beta-\alpha} = \{0\}$; thus $[X_\beta, Y_\alpha] = 0$ and so, also, $[\tilde{X}_\beta, \tilde{Y}_\alpha] = 0$. This being the case, we have $\tilde{X}_\beta \tilde{Y}_\alpha^{m+1} v_0 = \tilde{Y}_\alpha^{m+1} \tilde{X}_\beta v_0 = 0$. Thus, $\tilde{Y}_\alpha^{m+1} v_0$ is annihilated by all of the $\tilde{X}_\beta$'s, $\beta \in \Delta$, and so also by all of the $\tilde{X}_\beta$'s, $\beta \in R^+$ (by the result at the beginning of this proof).

Now, if $\tilde{Y}_\alpha^{m+1} v_0$ were nonzero, then since $\tilde{Y}_\alpha^{m+1} v_0$ is annihilated by all of the $\tilde{X}_\beta$'s, with β in R^+, the proof of Proposition 7.17 would tell us that the smallest $\mathfrak{g}$-invariant subspace containing $\tilde{Y}_\alpha^{m+1} v_0$ would have highest weight $\mu - (m+1)\alpha$. Since, on the contrary, V_μ/U_μ is irreducible with highest weight μ, we must have $\tilde{Y}_\alpha^{m+1} v_0 = 0$. This (together with (7.14) and (7.15)) tells us that the space spanned by $v_0, \tilde{Y}_\alpha v_0, \ldots, \tilde{Y}_\alpha^m v_0$ is invariant under $\mathfrak{s}^\alpha$.

Step 3. Given any $\alpha \in \Delta$, every vector in V_μ/U is contained in a finite-dimensional s^α-invariant subspace. Call a vector $v \in V_\mu/U_\mu$ **$\mathfrak{s}^\alpha$-finite** if v is

contained in a finite-dimensional $\mathfrak{s}^\alpha$-invariant subspace. Then, define a subspace T_α of V_μ/U_μ by

$$T_\alpha = \{\, v \in V_\mu/U_\mu \,|\, v \text{ is } \mathfrak{s}^\alpha\text{-finite}\}\,.$$

Step 2 shows that $T_\alpha \neq \{0\}$. I claim that T_α is invariant under $\mathfrak{g}$. To see this, consider $v \in T_\alpha$ and let T be a finite-dimensional $\mathfrak{s}^\alpha$-invariant subspace containing v. Now, let $\{Z_k\}_{k=1}^{\dim \mathfrak{g}}$ be a basis for $\mathfrak{g}$ and let T' be the sum of the spaces $Z_k T$. Since $\mathfrak{g}$ is finite dimensional, T' is also finite dimensional. However, now we can see that T' is invariant under $\mathfrak{s}^\alpha$, since for Z^α any element of $\mathfrak{s}^\alpha$,

$$Z^\alpha Z_k T = Z_k Z^\alpha T + [Z^\alpha, Z_k]\, T,$$

and $Z^\alpha T \subset T$ and $[Z^\alpha, Z_k]$ is a linear combination of the Z_l's. So, T' is a finite-dimensional $\mathfrak{s}^\alpha$-invariant subspace that contains Zv for all $Z \in \mathfrak{g}$. This shows that Zv is again in T_α, and, so, T_α is invariant under $\mathfrak{g}$.

Since V_μ/U_μ is irreducible and T_α is nonzero and $\mathfrak{g}$-invariant, we conclude that $T_\alpha = V_\mu/U_\mu$.

Step 4. For each $\alpha \in \Delta$, $\tilde{Y}_\alpha$ is locally nilpotent. Given any $v \in V_\mu/U_\mu$, v is contained in a finite-dimensional $\mathfrak{s}^\alpha$-invariant subspace T. By complete reducibility for $\mathfrak{sl}(2;\mathbb{C})$, T decomposes as a direct sum of (finitely many) irreducible $\mathfrak{s}^\alpha$-invariant subspaces. In each of these irreducible spaces, we have completely worked out (in Chapter 4) the action of Y_α and this action is nilpotent. So, $\tilde{Y}_\alpha^k v = 0$, where k is the maximum of the dimensions of the irreducible summands in T.

This concludes the proof of Proposition 7.25. □

7.3.4 The $\mathfrak{sl}(2;\mathbb{C})$ case

Let us see how this all works out in the case of $\mathfrak{sl}(2;\mathbb{C})$. If X, Y, and H are the usual basis elements, we work with the Cartan subalgebra $\mathfrak{h} = \text{span}(H)$. Weights may then be thought of simply as eigenvalues for $\pi_\mu(H)$. We build a vector space containing linearly independent vectors v_0, v_1, $v_2, \ldots$, and we define V_μ as the space of *finite* linear combinations of these vectors. Here, μ is an arbitrary complex number and the vector space itself does not depend on μ.

We now describe an action of $\mathfrak{sl}(2;\mathbb{C})$ on this space as follows. For H, we define

$$\begin{aligned}\pi_\mu(H)v_0 &= \mu v_0,\\ \pi_\mu(H)v_k &= (\mu - 2k)v_k.\end{aligned}$$

For Y, we define

$$\pi_\mu(Y)v_k = v_{k+1}.$$

Finally, for X, we define

$$\pi_\mu(X)v_0 = 0, \tag{7.16}$$

$$\pi_\mu(X)v_k = k(\mu + 1 - k)v_{k-1}, \quad k \geq 1. \tag{7.17}$$

The calculations of Chapter 4 show that if we define $\pi_\mu(H)$ and $\pi_\mu(Y)$ as above, then we must define $\pi_\mu(X)$ as above, if $\pi_\mu(X)$, $\pi_\mu(Y)$, and $\pi_\mu(H)$ are to satisfy the $\mathfrak{sl}(2;\mathbb{C})$ commutation relations. Direct calculation then shows that with these definitions the operators really do satisfy these commutation relations.

Let us now compute the invariant subspace U_μ. If we begin with the vector v_k, then using (7.17) repeatedly, we obtain

$$\pi_\mu(X)^k v_k = \left(\prod_{l=1}^{k} l(\mu + 1 - l)\right) v_0. \tag{7.18}$$

If $\mu = l - 1$ for any l in the range $1, \ldots, k$, then the coefficient of v_0 will be nonzero. So, if μ is anything other than a non-negative integer, the coefficient of v_0 will always be nonzero, and from this it follows easily that $U_\mu = \{0\}$. In that case, $V_\mu / U_\mu \cong V_\mu$ will be infinite dimensional.

On the other hand, if $\mu = m$, where m is a non-negative integer, then the coefficient of v_0 in (7.18) will be zero for all $k > m$. In this case, $U_\mu = U_m$ will consist of all linear combinations of the v_k's with $k > m$. The quotient space V_m / U_m can then be identified with the span of $v_0, \ldots, v_m$ and is finite dimensional.

Note that the Weyl group for this problem is simply $\{I, -I\}$. The eigenvalues λ of H occurring in V_μ are definitely *not* invariant under $\lambda \to -\lambda$. Nevertheless, in the case $\mu = m$, the eigenvalues for H occurring in V_m / U_m are $m, m-2, \ldots, -m$ and this set of eigenvalues is invariant under $\lambda \to -\lambda$.

7.4 Constructing the Representations II: The Peter–Weyl Theorem

In this approach, we construct the representations as representations of the simply-connected compact group K whose complexified Lie algebra is $\mathfrak{g}$. We make use of the *Haar measure* on K, used already in the proof of complete reducibility. (See Section C.4.) The Haar measure is a finite measure on K that is invariant under the left and right actions of K. We normalize the measure so that $\mu(K) = 1$. We then consider the Hilbert space $L^2(K, \mu)$ consisting of measurable complex-valued functions f on K with the property that

$$\int_K |f(x)|^2 \, d\mu(x) < \infty.$$

The finite-dimensional irreducible representations of K will ultimately be realized as certain finite-dimensional subspaces of $L^2(K, \mu)$. The construction of

the representations is based on three main results: the Peter–Weyl theorem, the Weyl character formula, and the Weyl integral formula. (See Section 7.6 for more information on some of these results.)

7.4.1 The Peter–Weyl theorem

In this subsection, it is not necessary that K be a compact *Lie* group; any compact topological group will do. We consider the Hilbert space $L^2(K, \mu)$, the space of measurable functions on K that are square-integrable with respect to the Haar measure on K.

The Peter–Weyl theorem gives a decomposition of $L^2(K, \mu)$ into finite-dimensional subspaces that are invariant under the left and right actions of K. Specifically, if Σ is a finite-dimensional irreducible representation of K acting on a vector space V, then we consider the space of **matrix entries** of Σ. Suppose we choose a basis $\{u_k\}$ for V. Then, for each $x \in K$, the linear operator $\Sigma(x)$ can be expressed as a matrix with respect to this basis; we denote the entries of this matrix as $\Sigma(x)_{kl}$. Then, a matrix entry for Σ is a function on K that can be expressed in the form

$$f(x) = \sum_{k,l=1}^{\dim V} \alpha_{kl} \Sigma(x)_{kl} \tag{7.19}$$

for some set of constants α_{kl}.

We can describe the space of matrix entries in a basis-independent way as the space of functions that can be expressed in the form

$$f(x) = \text{trace}(\Sigma(x)A) \tag{7.20}$$

for some linear operator A on V. To see the equivalence of these two forms, let A_{kl} be the matrix for the operator A in the basis $\{u_k\}$. Then, the matrix for $\Sigma(x)A$ is given by the matrix product $(\Sigma(x)A)_{kl} = \sum_m \Sigma(x)_{km} A_{ml}$ and, so,

$$\text{trace}(\Sigma(x)A) = \sum_{k,m=1}^{\dim V} \Sigma(x)_{km} A_{mk}.$$

Thus, every function of the form (7.20) can be expressed in form (7.19) with $\alpha_{kl} = A_{lk}$, and vice versa.

The significance of the space of matrix entries is that it is a finite-dimensional space of functions on K that is invariant under both left and right translations by K. To understand what this means, suppose f is a matrix entry for a representation Σ, given, say, as in (7.20). Now, suppose we define a new function f_{y_1,y_2} by shifting f on the left by y_1 and on the right by y_2; that is, we set $f_{y_1,y_2}(x) = f(y_1 x y_2)$. Then, we have that

$$\begin{aligned} f_{y_1,y_2}(x) &= \text{trace}(\Sigma(y_1)\Sigma(x)\Sigma(y_2)A) \\ &= \text{trace}(\Sigma(x)[\Sigma(y_2)A\Sigma(y_1)]). \end{aligned}$$

Thus, f_{y_1,y_2} is, again, a matrix entry for Σ, with A replaced by $\Sigma(y_2)A\Sigma(y_1)$.

Of particular importance among the matrix entries is the **character** of the representation Σ, denoted χ_Σ, which is the function on K given by

$$\chi_\Sigma(x) = \text{trace}(\Sigma(x)).$$

This function is a matrix entry, obtained by taking $A = I$ in (7.20) or taking $\alpha_{kl} = \delta_{kl}$ in (7.19). The character is special because it satisfies

$$\chi_\Sigma(xyx^{-1}) = \text{trace}(\Sigma(x)\Sigma(y)\Sigma(x)^{-1}) = \text{trace}(\Sigma(y)) = \chi_\Sigma(y) \tag{7.21}$$

for all x and y in K. Any function f on K satisfying $f(xyx^{-1}) = f(y)$ for all x and y in K is called a **class function**. The reason for this terminology is that the set of group elements of the form xyx^{-1}, with $y \in K$ fixed and x ranging over K, is called the **conjugacy class** of y. A class function is then a function that is constant on each conjugacy class.

It is easily verified that two equivalent representations have the same character. The converse of this is much less obvious but also true: If two finite-dimensional representations of K have the same character, they are equivalent.

The Peter–Weyl theorem gives a way of expressing any function $f \in L^2(K, \mu)$ in a series expansion in terms of matrix entries of the irreducible representations of K. We are interested primarily in the case in which f is a class function. In that case, the expansion involves only the characters. The Peter–Weyl theorem, specialized to the case of class functions, is as follows.

Theorem 7.27 (Peter–Weyl). *Let $L^2(K, \mu)^K$ denote the subspace of the Hilbert space $L^2(K, \mu)$ consisting of square-integrable class functions. Then, the functions*

$$\chi_\Sigma$$

form an orthonormal basis for $L^2(K, \mu)^K$, where Σ ranges over the equivalence classes of irreducible finite-dimensional representations of K.

Proving that the characters form an orthonormal *set* of functions in $L^2(K, \mu)^K$ is a fairly elementary calculation using little more than Schur's Lemma. (See Section II.4 of Bröcker and tom Dieck (1985).) Proving that the characters form an orthonormal *basis* for $L^2(K, \mu)^K$ requires some analytical argument. (See Section III.3 of Bröcker and tom Dieck (1985).)

7.4.2 The Weyl character formula

We now assume that K is a simply-connected compact Lie group. (There is also a version of the result for connected compact Lie groups that are not simply connected.) We choose, as usual, a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$ and we let T be the connected Lie subgroup of K whose Lie algebra is $\mathfrak{t}$. It can be shown that T is a closed subgroup of K (called a "maximal torus"). It can further be shown that every element of K is conjugate to an element

of T. This means that the values of a class function on K are, in principle, determined by its values on T. The Weyl character formula is a formula for the restriction to T of the character of an irreducible representation of K.

We let $\mathfrak{g}$ denote the complexification of the Lie algebra $\mathfrak{k}$ of K, so that $\mathfrak{g}$ is a complex semisimple Lie algebra. Then, $\mathfrak{h} := \mathfrak{t} + i\mathfrak{t}$ is a Cartan subalgebra in $\mathfrak{g}$. We follow Notation 6.24 and regard the roots as elements of $\mathfrak{h}$ (not $\mathfrak{h}^*$). Proposition 6.15 (expressed in terms of Notation 6.24) tells us that if $\alpha \in \mathfrak{h}$ is a root, then $\langle \alpha, H \rangle$ is imaginary for all H in $\mathfrak{t}$, which means that α itself is in $i\mathfrak{t}$. It is then convenient to introduce the **real roots**, which are simply $1/i$ times the ordinary roots. This means that a real root is a nonzero element α of $\mathfrak{t}$ with the property that there exists a nonzero X in $\mathfrak{g}$ with

$$[H, X] = i\langle \alpha, H \rangle X$$

for all H in $\mathfrak{t}$ (or, equivalently, for all H in $\mathfrak{h}$). We can also introduce the real co-roots as the elements of $\mathfrak{t}$ of the form $H_\alpha = 2\alpha/\langle \alpha, \alpha \rangle$, where α is a real root.

In the same way, we will consider the **real weights**, which we think of as elements of $\mathfrak{t}$ in the same way as for the roots. So, if (Σ, V) is an irreducible representation, then an element μ of $\mathfrak{t}$ is called a real weight for Σ if there exists a nonzero vector $v \in V$ such that

$$\sigma(H)v = i\langle \mu, H \rangle v$$

for all H in $\mathfrak{t}$. (Here, σ is the Lie algebra representation associated to the group representation Σ.) An element μ of $\mathfrak{t}$ is said to be **integral** if $\langle \mu, H_\alpha \rangle$ is an integer for each real co-root H_α. (All of the "real" objects are simply $1/i$ times the corresponding objects without the qualifier "real.") The real weights of any finite-dimensional representation of $\mathfrak{g}$ must be integral.

For the rest of this section, all of roots and weights will be assumed real, even if this is not explicitly stated.

If α is an integral element, then it can be shown that there is a function f on T satisfying

$$f(e^H) = e^{i\langle \alpha, H \rangle} \tag{7.22}$$

for all $H \in \mathfrak{h}$. To understand this assertion, note that because T is connected and commutative, every element t of T can be expressed as $t = e^H$ (Exercise 25 from Chapter 2). However, a given t can be expressed as $t = e^H$ in many different ways; the content of the above assertion is that the right side of (7.22) is independent of the choice of H for a given t. This means that we want to say that the right-hand side of (7.22) defines a function on T, not just on $\mathfrak{t}$. We will discuss this point further in the next subsection.

Next, we introduce the element δ of $\mathfrak{t}$ defined to be half the sum of the positive roots:

$$\delta = \frac{1}{2} \sum_{\alpha \in R^+} \alpha.$$

It can be shown that δ is an integral element. (Clearly, 2δ is integral, but it is not obvious that δ itself is integral.) Finally, if w is any element of the Weyl group, we think of w as an orthogonal linear transformation of $\mathfrak{t}$—in which case, $\det(w) = \pm 1$.

We are now ready to state the Weyl character formula.

Theorem 7.28 (Weyl Character Formula). *If Σ is an irreducible representation of K with highest real weight μ, then we have*

$$\chi_\Sigma\left(e^H\right) = \frac{\sum_{w \in W} \det(w)\, e^{i\langle w \cdot (\mu+\delta), H\rangle}}{\sum_{w \in W} \det(w)\, e^{i\langle w \cdot \delta, H\rangle}} \tag{7.23}$$

for all H in $\mathfrak{t}$ for which the denominator of the right-hand side of (7.23) is nonzero. Here, δ denotes half the sum of the positive real roots.

The set of points H for which the denominator of the Weyl character formula (the so-called **Weyl denominator**) is nonzero is dense in $\mathfrak{t}$. At points where the denominator is zero, there is an apparent singularity in the formula for χ_Σ. However, actually at such points the numerator is also zero and the character itself remains finite (as must be the case since, from the definition of the character, it is well defined and finite at every point). Note that the character formula gives a formula for the restriction of χ_Σ to T. Since χ_Σ is a class function and since (as we have asserted but not proved) every element of K is conjugate to an element of T, knowing χ_Σ on T determines, in principle, χ_Σ on all of K.

A sketch of the proof of the Weyl character formula is given in Section 7.6.

7.4.3 Constructing the representations

Recall that our goal is to show that every dominant integral element μ actually arises as the highest weight of some irreducible representation of K. To do this, we consider an arbitrary dominant integral element μ, which, at the moment, we do not know to be the highest weight of any representation. However, whether or not μ is the highest weight of some representation, it can be shown that the right-hand side of (7.23) defines a function on T that is invariant under the action of the Weyl group. Then, there exists a unique class function f_μ on K whose restriction to T is given by the right-hand side of (7.23). Using something called the **Weyl integral formula** (see Section 7.6), it can be shown that the functions f_μ, where μ ranges over all dominant integral elements, are orthonormal. It is essential here that we can prove that all of the f_μ's are orthonormal by direct computation, without appealing to the Peter–Weyl theorem and without knowing that every μ is the highest weight of a representation.

Let us take stock of the situation. We have the following results. First, the Peter–Weyl theorem tells us that the characters for the (equivalence classes of) irreducible representations form an orthonormal basis for the space of L^2 class

functions. Second, the Weyl character formula tells us that for an irreducible representation having highest weight μ, the character of the representation is given by (7.23). Third, the Weyl integral formula tells us that if for *every* dominant integral element μ we define f_μ to be the unique class function whose restriction to T is given by (7.23), then the f_μ's are orthonormal. This holds even though we do not know at the moment that every dominant integral element is the highest weight of a representation.

These results together imply that every dominant integral element is actually the highest weight of some representation. To see this, note that the Peter–Weyl theorem and the Weyl character formula tell us that the set of f_μ's, where μ ranges over all the highest weights of representations, form an orthonormal *basis* for $L^2(K,\mu)^K$. On the other hand, the Weyl integral formula says that the set of f_μ's, where μ ranges over all dominant integral elements, forms an orthonormal set. This second orthonormal set contains the first one, since the highest weight of an irreducible representation must be dominant integral. However, an orthonormal *basis* cannot be contained in a strictly larger orthonormal set—if it were, it would not be a basis. So, the only possibility is that the set of highest weights of irreducible representations is equal to the set of dominant integral elements, which is what we are trying to prove.

To say the same thing a different way, suppose there were some dominant integral element μ that was not the highest weight of any representation, and consider the function f_μ. Since (we are assuming) μ is not the highest weight of a representation, the set of characters is a certain set of f_σ's, where σ ranges over some subset of the dominant integral elements not including μ. However, the Weyl integral formula tells us that f_μ is orthogonal to f_σ, for all $\sigma \neq \mu$. This means that f_μ is a nonzero class function that is is orthogonal to all characters, since the characters are all f_σ's with $\sigma \neq \mu$. This, however, is impossible: The Peter–Weyl theorem implies that any class function that is orthogonal to all the characters must be zero. So, μ must, after all, be the highest weight of some representation.

To see this argument spelled out in greater detail, see Bröcker and tom Dieck (1985) or Simon (1996). The argument in those books is slightly more complicated than the one described here because those books consider arbitrary connected compact groups, not necessarily simply connected.

This "construction" of the representations of K is not very constructive; that is, we have proved that a representation with each dominant integral element exists, but we have not given a very explicit description of the representation. If one looks at the proof of the Peter–Weyl theorem, one will see that the representations are realized as certain finite-dimensional, translation-invariant spaces of functions on K, but it is not especially easy to see precisely which functions one gets. The Borel–Weil construction, described in the next section, gives a more explicit realization of the representations. Thus, the Borel–Weil construction is useful even if one has already proved that every dominant integral element is the highest weight of some representation.

7.4.4 Analytically integral versus algebraically integral elements

One step of the argument in the previous subsection deserves elaboration. We asserted (see (7.22)) that if K is simply connected and μ is a (real) integral element, then there exists a function f on T satisfying

$$f\left(e^H\right) = e^{i\langle \mu, H\rangle} \tag{7.24}$$

for all H in $\mathfrak{t}$. Let us think about what is entailed in this statement. Since the Lie algebra $\mathfrak{t}$ of the connected group T is commutative, T itself must also be commutative. It follows that the exponential map $\exp : \mathfrak{t} \to T$ is a homomorphism. The image of this homomorphism contains a neighborhood of the identity in T (by the local surjectivity of the exponential mapping for arbitrary Lie groups). Since, also, T is connected, it follows that the exponential mapping for T is surjective.

Now, let $\Phi \subset \mathfrak{t}$ be the kernel of the exponential mapping; that is,

$$\Phi = \left\{ H \in \mathfrak{t} \middle| e^H = I \right\}.$$

Since the exponential mapping for T is surjective, every $t \in T$ can be written as $t = e^H$ for some H in $\mathfrak{t}$. If H_1 and H_2 are in $\mathfrak{t}$ and $e^{H_1} = e^{H_2}$, then (since $\mathfrak{t}$ is commutative) $e^{H_1 - H_2} = I$ and, so, $H_1 - H_2$ is in Φ. This means that every $t \in T$ can be written as $t = e^H$, and the H is unique up to adding on an element of Φ. So, now suppose we try to define a function f on T by defining (as in (7.24)) $f(t) = \exp i\langle \mu, H\rangle$, where H is chosen so that $e^H = t$. When is this well defined (i.e., independent of the choice of H)? Well, if H is such that $e^H = t$, then any H' with $e^{H'} = t$ must be of the form $H' = H + \phi$, with $\phi \in \Phi$. Therefore, we need that $\exp i\langle \mu, H\rangle = \exp i\langle \mu, H'\rangle = \exp i\langle \mu, H\rangle \exp i\langle \mu, \phi\rangle$. This will hold precisely if $\langle \mu, \phi\rangle$ is an integer multiple of 2π. We conclude, then, that the function in (7.24) is well defined precisely if μ has the property that $\langle \mu, \phi\rangle$ is an integer multiple of 2π for all ϕ in the kernel of the exponential mapping. An element μ of $\mathfrak{t}$ having this property is called an analytically integral element.

For reasons of consistency with Appendix E, it is convenient to introduce the set

$$\Lambda = \left\{ H \in \mathfrak{t} \middle| e^{2\pi H} = I \right\},$$

so that the elements λ of Λ are precisely those for which $2\pi\lambda \in \Phi$ (i.e., those λ of the form $\lambda = \phi/2\pi$ for some $\phi \in \Phi$). Saying that $\langle \mu, \phi\rangle$ is an integer multiple of 2π is the same as saying that $\langle \mu, \lambda\rangle$ is an integer. So, we have the following definition.

Definition 7.29. *An element μ of $\mathfrak{t}$ is called an* ***analytically integral element*** *if $\langle \mu, \lambda\rangle$ is an integer for all λ in Λ.*

The following result summarizes the conclusion of the previous paragraphs.

Proposition 7.30. *For $\mu \in \mathfrak{t}$, there exists a function f on T satisfying*

$$f\left(e^H\right) = e^{i\langle \mu, H\rangle}$$

if and only if μ is an analytically integral element.

Meanwhile, we have another notion of integral elements, namely that $\mu \in \mathfrak{h}$ is an integral element if $\langle \mu, H_\alpha \rangle$ is an integer for each co-root H_α. To distinguish this condition from the condition for an analytically integral element, we call μ an **algebraically integral element** if $\langle \mu, H_\alpha \rangle$ is an integer for all co-roots. In the previous subsection, we asserted that (assuming K is simply connected) f in (7.24) is well defined provided that μ is an *algebraically* integral element. In light of Proposition 7.30, this amounts to asserting that *every algebraically integral element is analytically integral.* In fact, we have the following result.

Theorem 7.31. *If K is simply connected, then the set of algebraically integral elements and the set of analytically integral elements are the same.*

This theorem is not at all obvious. It is not hard to show that every analytically integral element is algebraically integral; indeed, this is true even if K is not simply connected. Showing (in the simply-connected case) that every algebraically integral element is analytically integral is more involved. See Section E.4 for more information.

Books such as Bröcker and tom Dieck (1985) and Simon (1996) are concerned with the representations of compact Lie *groups*, which are not assumed to be simply connected. These books do not address the issue of whether every Lie algebra representation comes from a group representation. Thus, in those books, the only relevant notion of integral element is that of an analytically integral element and one never needs to worry about the relationship between algebraically integral and analytically integral element. The theorem of the highest weight, as presented in those books, says that every dominant and *analytically* integral element is the highest weight of a representation of K. This result holds whether K is simply connected or not.

We, on the other hand, wish to connect the compact-group approach to the Lie algebra approach and, for this, it is necessary to know Theorem 7.31.

7.4.5 The SU(2) case

Let us see how the argument described in this section works out in the $\mathrm{SU}(2)$ case. We work with the maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{su}(2)$ given by

$$\mathfrak{t} = \left\{ \begin{pmatrix} ia & 0 \\ 0 & -ia \end{pmatrix} \middle| \, a \in \mathbb{R} \right\}.$$

Note that $\mathfrak{t}$ is the set of matrices of the form iaH, where, as usual,

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

The associated maximal torus $T \subset \mathsf{SU}(2)$ is then

$$T = \left\{ e^{iaH} = \begin{pmatrix} e^{ia} & 0 \\ 0 & e^{-ia} \end{pmatrix} \middle| \, a \in \mathbb{R} \right\}.$$

The Weyl group is the two-element group $\{I, -I\} \subset \mathsf{O}(\mathrm{t})$.

Let us now compute the characters of the irreducible representations of $\mathsf{SU}(2)$, or, more precisely, the restriction to T of the characters. We know that in each irreducible representation Σ, $\sigma(H)$ is diagonalizable with eigenvalues $m, m-2, \ldots, -m$. Thus, $\Sigma(\exp iaH) = \exp(ia\sigma(H))$ is also diagonalizable with eigenvalues $\exp(ima)$, $\exp(i(m-2)a)$, etc. The trace of $\Sigma(\exp iaH)$ is then the sum of the eigenvalues:

$$\chi_m\left(e^{iaH}\right) = \mathrm{trace}\left(\Sigma\left(e^{iaH}\right)\right) = \sum_{k=0}^{m} e^{i(m-2k)a}.$$

This is a finite geometric series, which we will sum using a slight variation of the usual approach. We multiply χ_m by $\exp(ia)$ and then by $\exp(-ia)$, and subtract the results. When we do this, all but two terms in the geometric series cancel and we get

$$\left(e^{ia} - e^{-ia}\right)\chi_m\left(e^{iaH}\right) = e^{i(m+1)a} - e^{-i(m+1)a} \tag{7.25}$$

so that

$$\chi_m\left(e^{iaH}\right) = \frac{e^{i(m+1)a} - e^{-i(m+1)a}}{e^{ia} - e^{-ia}} = \frac{\sin((m+1)a)}{\sin a}. \tag{7.26}$$

The first equality in (7.26) is nothing but the Weyl character formula for the $\mathsf{SU}(2)$ case. Note that $\sin((m+1)a)$ is zero at all points at which $\sin a$ is zero (namely all integer multiples of π) and so the expression for χ_m is nonsingular, even at points where the denominator is zero.

Meanwhile, suppose that f is any class function on $\mathsf{SU}(2)$ and let dA denote the normalized Haar measure on $\mathsf{SU}(2)$. The Weyl integral formula (Section 7.6) in this case states that

$$\int_{\mathsf{SU}(2)} f(A)\, dA = \int_0^{2\pi} f\left(e^{iaH}\right) 2\sin^2 a \, \frac{da}{2\pi}. \tag{7.27}$$

Here, $da/2\pi$ is the normalized Haar measure on $T = \left\{ e^{iaH} \middle| \, a \in \mathbb{R} \right\}$.

We know from Chapter 4 that for each non-negative integer m, there is an irreducible representation with highest weight m. Let us pretend that we do not know this and see how the result follows from the Peter–Weyl theorem, the Weyl character formula (7.26), and the Weyl integral formula (7.27). For any non-negative integer m, whether or not we know that m is the highest

weight of a representation, (7.26) gives a well-defined function on T that is invariant under the map $a \to -a$. It is not hard to show, then, that there exists a unique class function f_m on $\mathsf{SU}(2)$ whose restriction to T is given by (7.26). According to (7.27), we have for any distinct non-negative integers m and n,

$$\int_{\mathsf{SU}(2)} f_m(A) f_n(A)\, dA = \int_0^{2\pi} \frac{\sin((m+1)a)}{\sin a} \frac{\sin((n+1)a)}{\sin a} 2\sin^2 a \, \frac{da}{2\pi} = 0, \tag{7.28}$$

because of the usual orthogonality of trigonometric functions on $[0, 2\pi]$.

Suppose, now, that there were some m that was not the highest weight of a representation. Then, the set of characters for $\mathsf{SU}(2)$ would consist of f_n's with n ranging over some subset of the non-negative integers not including m. This would mean, by (7.28), that f_m is a nonzero class function that is orthogonal to all of the characters (since the characters are all f_n's with $n \neq m$). However, the Peter–Weyl theorem says that the characters form an orthonormal *basis* for the set of L^2 class functions on $\mathsf{SU}(2)$, and thus a nonzero class function cannot be orthogonal to all of the characters. This, then, is a contradiction and m must be, after all, the highest weight of a representation.

7.5 Constructing the Representations III: The Borel–Weil Construction

The Borel–Weil construction, described in this section, is often described *after* the theorem of the highest weight has been proved (using, say, Verma modules to prove that every dominant integral element arises as the highest weight of an irreducible finite-dimensional representation). In such approaches, the Borel–Weil construction is simply an illuminating way to "realize" representations whose existence has already been demonstrated. However, it is also possible to use the Borel–Weil construction to prove the existence of the representations, and that is the approach we will follow here.

7.5.1 The complex-group approach

We have seen already two approaches to constructing the representations: the Lie algebra point of view, using Verma modules, and the compact group-point of view, using the Peter–Weyl theorem and the Weyl character formula. We now consider the complex-group point of view, using something called the Borel–Weil construction. (There is also a variant of the Borel–Weil construction that uses algebraic groups instead of complex groups.)

Before getting into the details of the Borel–Weil construction, we need to establish some notation and at the same time make sure that we understand the relationships between the representations of the various objects (Lie algebra, compact group, complex group). Let $\mathfrak{g}$ be a complex semisimple Lie

algebra realized as a subalgebra of some $M_n(\mathbb{C})$. Let $\mathfrak{k}$ be a compact real form of $\mathfrak{g}$ and let K and G be the connected Lie subgroups of $\mathsf{GL}(n;\mathbb{C})$ whose Lie algebras are $\mathfrak{k}$ and $\mathfrak{g}$, respectively. It can be shown that both K and G are closed subgroups of $\mathsf{GL}(n;\mathbb{C})$ and, hence, matrix Lie groups. Let us assume for simplicity that K and G are both simply connected. (It turns out that G is simply connected if and only if K is—see the last section of Appendix E.) In this case, the representations of K (i.e., continuous homomorphisms of K into some $\mathsf{GL}(N;\mathbb{C})$) are in one-to-one correspondence with the representations of $\mathfrak{k}$, which, in turn, are in one-to-one correspondence with the complex-linear representations of $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$.

Now, since we assume that G is simply connected, every complex-linear representation of $\mathfrak{g}$ can be exponentiated to give a representation of G. However, not every representation of G arises in this way. After all, if Π is a representation of G (continuous homomorphism of G into some $\mathsf{GL}(N;\mathbb{C})$), there is no reason that the associated Lie algebra representation π should be *complex*-linear. For example, consider the representation of $\mathsf{SL}(2;\mathbb{C})$ given by entrywise complex conjugation:

$$\Pi\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}.$$

(This *is* a representation (i.e., a continuous homomorphism of $\mathsf{SL}(2;\mathbb{C})$ into $\mathsf{GL}(2;\mathbb{C})$) because the complex conjugate of the product of two matrices is the same as the product of the complex conjugates.) The associated representation π of $\mathsf{sl}(2;\mathbb{C})$ is given by the same formula as Π, since the complex conjugate of $\exp(tX)$ is $\exp(t\bar{X})$, where $\bar{X}$ is the complex conjugate (entrywise) of X. Clearly, then, π is not complex-linear but rather conjugate-linear.

We call a representation of G **holomorphic** if the associated representation of $\mathfrak{g}$ is complex-linear. Since $\mathfrak{g}$ is a complex subalgebra of $M_n(\mathbb{C})$, the group G is automatically a complex submanifold of $\mathsf{GL}(n;\mathbb{C})$ (Appendix C), and it can be shown that if Π is holomorphic in the sense of the previous sentence, then Π is a holomorphic mapping of the complex manifold G into $\mathsf{GL}(N;\mathbb{C})$.

Assuming still that K and G are simply connected, we conclude that the following objects are in one-to-one correspondence with each other:

continuous representations of K
real-linear representations of $\mathfrak{k}$
complex-linear representations of $\mathfrak{g}$
holomorphic representations of G.

Here, "continuous" in the representations of K is to emphasize that we allow *any* continuous homomorphism of K into $\mathsf{GL}(N;\mathbb{C})$. There is one other class of objects that is frequently studied: the "algebraic" representations of G. It is not hard to see that every algebraic representation of G is holomorphic;

it is less obvious but still true that every holomorphic representation of G is also algebraic. So, actually, the algebraic representations of G are also in one-to-one correspondence with the above-listed objects.

7.5.2 The setup

We continue to assume that $\mathfrak{g} \subset M_n(\mathbb{C})$ is a complex semisimple Lie algebra with compact real form $\mathfrak{k}$ and that G and K are the connected subgroups of $\mathsf{GL}(n;\mathbb{C})$ with Lie algebras $\mathfrak{g}$ and $\mathfrak{k}$, respectively. We will assume that K is contained in $\mathsf{U}(n)$. (This is a harmless assumption since, by the averaging method of Section 4.10, there is an inner product on $\mathbb{C}^n$ that is invariant under the action of K, and we can make a change of basis that converts this inner product into the usual one on $\mathbb{C}^n$.) Having made this assumption, we will have that each $X \in \mathfrak{k}$ will satisfy $X^* = -X$, where X^* is the usual matrix adjoint of X. It then follows that for any $Z = X_1 + iX_2 \in \mathfrak{g}$, we have that $Z^* = -X_1 + iX_2$ also belongs to $\mathfrak{g}$. From this it is not hard to show that for any A in G, A^* is also in G.

Now, suppose that π is a representation of $\mathfrak{g}$ acting on some space V, and Π is the associated representation of G. We can choose an inner product on V so that $\Pi(x)$ is unitary for all x in K. In that case, it is not hard to show that Π satisfies

$$\Pi(A)^* = \Pi(A^*)$$

for all $A \in G$. For $A \in K$, this means simply that $\Pi(A^*) = \Pi(A^{-1}) = \Pi(A)^{-1} = \Pi(A)^*$, since both A and $\Pi(A)$ are unitary.

We now choose a maximal commutative subalgebra $\mathfrak{t}$ in $\mathfrak{k}$ and we let $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$ be the associated Cartan subalgebra of $\mathfrak{g}$. We let R denote the set of roots for $\mathfrak{g}$ relative to $\mathfrak{h}$, we choose a base Δ for R, and we let R^+ denote the set of positive roots with respect to Δ. Now, consider the following subspaces of $\mathfrak{g}$,

$$\begin{aligned}
\mathfrak{n}^+ &= \bigoplus_{\alpha \in R^+} \mathfrak{g}_\alpha, \\
\mathfrak{n}^- &= \bigoplus_{\alpha \in R^-} \mathfrak{g}_\alpha, \\
\mathfrak{b}^+ &= \mathfrak{h} \oplus \mathfrak{n}^+, \\
\mathfrak{b}^- &= \mathfrak{h} \oplus \mathfrak{n}^-.
\end{aligned}$$

It is easily seen that each of these spaces is actually a subalgebra of $\mathfrak{g}$. It is also easily seen that $\mathfrak{n}^+$ is an ideal in $\mathfrak{b}^+$ and $\mathfrak{n}^-$ is an ideal in $\mathfrak{b}^-$ (i.e., $[\mathfrak{b}^+, \mathfrak{n}^+] \subset \mathfrak{n}^+$ and $[\mathfrak{b}^-, \mathfrak{n}^-] \subset \mathfrak{n}^-$). Neither $\mathfrak{n}^+$ nor $\mathfrak{n}^-$ is an ideal in $\mathfrak{g}$. From the proof of Proposition 6.19 we see that $(\mathfrak{g}_\alpha)^* = \mathfrak{g}_{-\alpha}$ and, thus, $(\mathfrak{b}^-)^* = \mathfrak{b}^+$ and $(\mathfrak{b}^+)^* = \mathfrak{b}^-$.

We now let B^+, B^-, N^+, and N^- be the connected Lie subgroups of G corresponding to $\mathfrak{b}^+$, $\mathfrak{b}^-$, $\mathfrak{n}^+$, and $\mathfrak{n}^-$, respectively. These are always closed subgroups of G and hence matrix Lie groups. We have

$$(B^+)^* = B^-,$$
$$(B^-)^* = B^+.$$

(This follows from the corresponding result for the Lie algebras $\mathfrak{b}^+$ and $\mathfrak{b}^-$.)

Now, suppose that μ is any element of $\mathfrak{h}$ and consider the linear map $\chi_\mu : \mathfrak{b}^+ \to \mathbb{C}$ given by

$$\chi_\mu(H + X) = \langle \mu, H \rangle, \quad H \in \mathfrak{h},\ X \in \mathfrak{b}^+. \tag{7.29}$$

It is easily checked (using that $\mathfrak{n}^+$ is an ideal in $\mathfrak{b}^+$) that χ_μ is a Lie algebra homomorphism of $\mathfrak{b}^+$ into the one-dimensional commutative Lie algebra $\mathbb{C}$. Now, $\mathfrak{b}^+$ is the Lie algebra of the group B^+, and we may think of $\mathbb{C}$ as the Lie algebra of the group $\mathbb{C}^*$ of nonzero complex numbers, with the exponential mapping from $\mathbb{C}$ to $\mathbb{C}^*$ being the usual exponential function. We may then ask whether or not there is an associated Lie group homomorphism

$$X_\mu : B^+ \to \mathbb{C}^*. \tag{7.30}$$

Note that even though we are assuming that G is simply connected, $B^+ \subset G$ is not necessarily simply connected. Indeed, it turns out that B^+ is *never* simply connected.

Proposition 7.32. *Let $\chi_\mu : \mathfrak{b}^+ \to \mathbb{C}$ be the Lie algebra homomorphism defined by (7.29). If G is simply connected, then an associated group homomorphism $X_\mu : B^+ \to \mathbb{C}^*$ exists if and only if $\mu \in \mathfrak{h}$ is an integral element.*

The motivation for the definition of χ_μ is that if u_0 is a highest weight vector for some representation π with highest weight μ, then

$$\pi(H + X)u_0 = \langle \mu, H \rangle u_0$$

for all $H \in \mathfrak{h}$ and $X \in \mathfrak{n}^+$. (Note that $\pi(X)u_0$ must be zero for X in $\mathfrak{n}^+$ since otherwise $\pi(X)u_0$ would be a linear combination of weight vectors with weight higher than μ.)

To make Proposition 7.32 plausible, suppose that μ is the highest weight of some representation π of $\mathfrak{g}$ (in which case, μ is certainly integral) and that u_0 is a weight vector with weight μ. If G is simply connected, then there will be an associated representation Π of G. Then, for all $X \in \mathfrak{b}^+$, we will have

$$\Pi\left(e^X\right) u_0 = e^{\pi(X)} u_0 = e^{\chi_\mu(X)} u_0. \tag{7.31}$$

Now, any $a \in B^+$ will be a finite product of elements of the form e^X, $X \in \mathfrak{b}^+$. It follows that for any $a \in B^+$, we will have $\Pi(a)u_0 = f(a)u_0$ for some nonzero complex number $f(a)$. The map $a \to f(a)$ will be a homomorphism of B^+ into $\mathbb{C}^*$, and by (7.31), the associated Lie algebra map is χ_μ, so, in fact, $f = X_\mu$. We conclude, then, that if μ is the highest weight of a representation π of $\mathfrak{g}$ and G is simply connected, then X_μ will exist. As a side benefit, we have shown that, in this case, the associated representation Π of G satisfies

$$\Pi(a)u_0 = X_\mu(a)u_0 \tag{7.32}$$

for all $a \in B^+$.

The argument in the preceding paragraph proves Proposition 7.32 in the case where μ is the highest weight of some representation, in which case, μ is certainly integral. However, of course, only *dominant* integral elements arise as highest weights and the proposition claims that X_μ exists for any integral μ, dominant or not. Furthermore, even in the dominant integral case, we are going to use the proposition in *proving* that every dominant integral element arises as the highest weight of some representation, in which case, we are not allowed to assume the existence of the representations in proving the proposition. Nevertheless, it is the dominant case that we are mainly interested in, and the reason that we are interested in the homomorphism X_μ is because of (7.32).

Let me now sketch briefly the proof of Proposition 7.32. It follows from the "polar decomposition" for G (see Section E.5) that G is simply connected if and only if K is simply connected. Consider the restriction of χ_μ to the maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$, which is the linear map $H \to \langle \mu, H \rangle$. Let T be the subgroup of K whose Lie algebra is $\mathfrak{t}$. Since $\mathfrak{t} \subset \mathfrak{b}^+$, it follows that $T \subset B^+$. If G is simply connected (and so also K), then Corollary E.8 implies that there exists a homomorphism $A_\mu : T \to \mathbb{C}^*$ such that $A_\mu(\exp H) = \exp\langle \mu, H \rangle$ if and only μ is an integral element. Now, if X_μ exists, then, certainly, A_μ must exist, since in that case, A_μ is simply the restriction of X_μ to $T \subset B^+$. Thus, if X_μ exists, then A_μ exists and, so (by Corollary E.8), μ must be integral. To go in the other direction, recall that B^+ is never simply connected, even when G and K are. (This is why not all of the homomorphisms of $\mathfrak{b}^+$ into $\mathbb{C}$ give rise to homomorphisms of B^+ into $\mathbb{C}^*$.) However, one can show that the fundamental group of B^+ is isomorphic to that of T in a natural way. This implies that all of the difficulty in passing from $\mathfrak{b}^+$ to B^+ is already present in passing from $\mathfrak{t}$ to T. So, if μ is integral, Corollary E.8 tells us that we can pass χ_μ from $\mathfrak{t}$ to T and, therefore, also from $\mathfrak{b}^+$ to B^+.

7.5.3 The strategy

The idea of the Borel–Weil construction is to realize each representation as the action of G on a certain space of functions on G itself. In order to understand the strategy, let us first assume that we have a representation Π of G and see what sort of functions on G we get from this. Then, the Borel–Weil construction will reverse this procedure: We will first construct a certain space of functions on G and then build the representation Π from these functions.

Thus, let (π, V) be a complex-linear representation of $\mathfrak{g}$ and let Π be the associated holomorphic representation of G. We continue to assume (as in the previous subsection) that Π satisfies $\Pi(g^*) = \Pi(g)^*$ for all g in G. Then, consider the **matrix entries** of Π. These are the functions on G of the form

$$F_{u,v}(g) = \langle u, \Pi(g)v \rangle\,, \tag{7.33}$$

where u and v are elements of V. Because Π is a holomorphic representation of G, these functions are holomorphic functions on G. For a fixed $u \in V$, let

$$\mathcal{F}^u = \text{the space of all functions of the form } F_{u,v}, \quad v \in V. \tag{7.34}$$

For any fixed u, the map $v \to F_{u,v}$ is a linear map, which (by the definition of $\mathcal{F}^u$) sends V onto $\mathcal{F}^u$. If Π is irreducible and $u \neq 0$, then it is not hard to show (Exercise 4) that the map $v \to F_{u,v}$ is injective. So, for $u \neq 0$, $\mathcal{F}^u$ is a finite-dimensional vector space that is naturally isomorphic (as a vector space) to V itself, by the map $v \to F_{u,v}$.

Now, if F is any function on G and h is an element of G, define a new function $R_h F$ by the formula

$$(R_h F)(g) = F(gh).$$

The operator R_h is the "right-translation by h" operator and it is a linear operator on the space of all functions on G. Furthermore, we compute that

$$(R_{h_1} R_{h_2} F)(g) = (R_{h_2} F)(gh_1) = F(gh_1 h_2) = (R_{h_1 h_2} F)(g).$$

So, the map $h \to R_h$ is a homomorphism and we may think of R as a representation of G, acting on the (infinite-dimensional) space of all functions on G. If we apply R_h to a matrix entry, then we have

$$R_h F_{u,v}(g) = \langle u, \Pi(gh)v \rangle = \langle u, \Pi(g)\Pi(h)v \rangle = F_{u,\Pi(h)v}(g). \tag{7.35}$$

So, evidently, the right action of G leaves $\mathcal{F}^u$ invariant and, thus, $\mathcal{F}^u$ is a finite-dimensional representation of G. Furthermore, if u is nonzero and Π is irreducible, then $\mathcal{F}^u$ is isomorphic as a representation to V, since the map $v \to F_{u,v}$ is an intertwining map, by (7.35).

The conclusion is this: We can realize any irreducible holomorphic representation of G as a space of holomorphic functions on G that is invariant under the right action of G.

Let us look as well at the left action of G on functions. Consider the "left-translation by h" operator, defined as

$$(L_h F)(g) = F(hg).$$

I leave it to the reader to check that $L_{h_1 h_2} = L_{h_2} L_{h_1}$. (For our purposes, it is not necessary that the map $h \to L_h$ be a homomorphism. If, however, one wants a homomorphism, then one merely needs to replace $F(hg)$ by $F(h^{-1}g)$ in the definition of L_h.) We compute that

$$\begin{aligned} L_h F_{u,v}(g) &= \langle u, \Pi(hg)v \rangle = \langle u, \Pi(h)\Pi(g)v \rangle \\ &= \langle \Pi(h)^* u, \Pi(g)v \rangle = F_{u',v}, \end{aligned}$$

where $u' = \Pi(h)^* u$. Thus, for typical h, L_h does *not* leave the space $\mathcal{F}^u$ invariant, since the value of u is changed. Recall that we have defined Π in such a way that $\Pi(h)^* = \Pi(h^*)$, and so we can rewrite the above result as

$$L_h F_{u,v}(g) = \langle \Pi(h^*)u, \Pi(g)v \rangle . \tag{7.36}$$

Let us consider the case that $u = u_0$, where u_0 is a highest weight vector with some highest weight μ. Suppose that we take $h = b$, where b is an element of B^-, so that b^* is in B^+. Recall the homomorphism $X_\mu : B^+ \to \mathbb{C}^*$ defined in (7.30) and recall also that (by (7.32))

$$\Pi(a)u_0 = X_\mu(a)u_0$$

for all $a \in B^+$. Applying this with $a = b^*$ and substituting into (7.36) gives

$$F_{u_0,v}(bg) = \overline{X_\mu(b^*)} F_{u_0,v}(g) \tag{7.37}$$

for all $b \in B^-$ and $g \in G$.

What do we conclude from all this? Given any finite-dimensional holomorphic representation Π of G, we can realize Π as a space $\mathcal{F}^{u_0}$ of holomorphic functions on G, where the *space* $\mathcal{F}^{u_0}$ is invariant under the right action of G and where each *element* F of $\mathcal{F}^{u_0}$ transforms under the left action of B^- according to (7.37). Now, it turns out that *every* holomorphic function on G satisfying (7.37) is actually an element of $\mathcal{F}^{u_0}$. (This is far from obvious at the moment.)

The idea, then, of the Borel–Weil construction is this. We start with an integral element μ and we want to show that μ is actually the highest weight of some representation Π. We construct the homomorphism $X_\mu : B^+ \to \mathbb{C}^*$ described in Proposition 7.32 and we define a space $\mathcal{F}_{\mu_0}$ as follows.

Definition 7.33. *If μ is an integral element, let $X_\mu : B^+ \to \mathbb{C}^*$ be the homomorphism given by Proposition 7.32. Then, we define $\mathcal{F}_\mu$ to be the space of all holomorphic functions on G satisfying*

$$F(bg) = \overline{X_\mu(b^*)} F(g)$$

for all $b \in B^-$ and all $g \in G$.

Recall that if b is in B^-, then b^* is in B^+. Although the definition of $\mathcal{F}_\mu$ makes sense for any integral element, we are interested primarily in the case in which μ is dominant integral. (It turns out that if μ is integral but not dominant, then $\mathcal{F}_\mu$ contains only the zero function.) In the case that μ is dominant integral, we want to establish the following results. (1) $\mathcal{F}_\mu$ is invariant under the right action of G. (2) $\mathcal{F}_\mu$ if finite dimensional. (3) $\mathcal{F}_\mu$ contains some nonzero elements. (4) $\mathcal{F}_\mu$, under the right action of G, is an irreducible holomorphic representation such that the associated Lie algebra representation has highest weight μ. This will show that the dominant integral element μ is indeed the highest weight of some irreducible representation and will give a "concrete" realization of that representation.

If we can prove all of this and we let u_0 be a highest weight vector inside $\mathcal{F}_\mu$, then we will have

$$\mathcal{F}_\mu = \mathcal{F}^{u_0},$$

where $\mathcal{F}^{u_0}$ is the space defined in (7.34). Note, however, that we are not allowed to *define* $\mathcal{F}_\mu$ to be equal to $\mathcal{F}^{u_0}$ since we do not know at the beginning that there is any representation Π with highest weight μ. So, instead, we define $\mathcal{F}_\mu$ to be the space of functions having the properties that we know the elements of $\mathcal{F}^{u_0}$ should have.

The hardest part of the Borel–Weil construction is to prove that the space $\mathcal{F}_\mu$ is nonzero (i.e., that there exists a nonzero holomorphic function F satisfying $F(bg) = \overline{X_{\mu_0}(b^*)}F(g)$, $b \in B^-$, $g \in G$). Note that if we knew that μ was the highest weight of some irreducible representation, then functions of the form (7.33) would be in $\mathcal{F}_\mu$ and, thus, $\mathcal{F}_\mu$ would be nonzero. However, we are trying to use the Borel–Weil construction to *prove* that such a representation exists, and so we are not allowed to assume this.

Since G is a complex group and B^- is a complex subgroup, the quotient manifold G/B^- is a complex manifold. (The group B^- is not a normal subgroup of G and, therefore, the quotient G/B^- is not a group but only a manifold.) The space $\mathcal{F}_\mu$ should really be thought of as the space of holomorphic sections of a certain holomorphic line bundle over G/B^-. This point of view allows a large amount of differential geometric machinery to be brought to bear, especially cohomology and the Riemann–Roch formula. For example, it can be shown that G/B^- is a *compact* complex manifold. (Specifically, G/B^- is identifiable with K/T, where T is the connected subgroup of K with Lie algebra $\mathfrak{t}$.) This implies that the space $\mathcal{F}_\mu$ of holomorphic sections is finite dimensional. This machinery is beyond the scope of this book and is not necessary if all one wants is to prove the existence of a representation with a given dominant integral element. For more information on the differential geometric side of things see Pressley and Segal (1986), Duistermaat and Kolk (2000), Knapp (1986), Knapp (1988), and the article by Eastwood and Sawon in Bridson and Salamon (2002).

7.5.4 The construction

We continue with the notation established in Subsection 7.5.2. We consider an integral element $\mu \in \mathfrak{h}$, which, at the moment, we do not assume to be dominant. We let $X_\mu : B^+ \to \mathbb{C}^*$ be the homomorphism given by Proposition 7.32 and we consider the space $\mathcal{F}_\mu$ of functions on G defined in Definition 7.33. We begin with one important but easy result.

Proposition 7.34. *For any integral element μ, the space $\mathcal{F}_\mu$ is invariant under the right action of G.*

Proof. Suppose that F is an element of $\mathcal{F}_\mu$ and h is an element of G. Then, for all $g \in G$ and $b \in B^-$, we have (by the associativity of the product on G)

$$(R_h F)(bg) = F(bgh) = \overline{X_\mu(b^*)}F(gh) = \overline{X_\mu(b^*)}(R_h F)(g).$$

This shows that $R_h F$ is again in $\mathcal{F}_\mu$. □

We are now ready to state the main result of this section.

Theorem 7.35. *Assume G is simply connected. Let μ be an integral element, let $X_\mu : B^+ \to \mathbb{C}^*$ be as in Proposition 7.32, and let $\mathcal{F}_\mu$ be as in Definition 7.33. Then, the following results hold:*

1. *If μ is not dominant, then $\mathcal{F}_\mu$ contains only the zero function.*
2. *If μ is dominant integral, then $\mathcal{F}_\mu$ is nonzero but finite dimensional. In this case $\mathcal{F}_\mu$, forms an irreducible holomorphic representation under the right action of G, and this representation has highest weight μ.*

This construction of the representations of G is called the Borel–Weil construction.

Let us elaborate slightly on the meaning of Point 2 of the theorem. The space $\mathcal{F}_\mu$ is invariant under R_h ($h \in G$) and is finite dimensional. It can be shown that the map $h \to R_h|_{\mathcal{F}_\mu}$ is a continuous map of G into $\mathsf{GL}(\mathcal{F}_\mu)$ and, thus, $\mathcal{F}_\mu$ constitutes a finite-dimensional representation of G. Point 2 of the theorem asserts that this representation is holomorphic, meaning that the associated representation of $\mathfrak{g}$ is complex-linear. The statement that "this representation has highest weight μ" then means more precisely that the associated complex-linear representation of $\mathfrak{g}$ has highest weight μ.

It is not feasible for me to give a complete proof of this theorem here. I outline the intermediate steps needed and prove the ones that can be proved easily. The main omission is that I do not prove that $\mathcal{F}_\mu$ is nonzero in the dominant integral case. (Recall that if one already knows that a dominant integral element μ is the highest weight of a finite-dimensional representation, then it is easy to show, as in the previous subsection, that $\mathcal{F}_\mu$ is nonzero. However, we are trying to use the Borel–Weil construction to prove the existence of such a representation; to do so we must prove directly that $\mathcal{F}_\mu$ is nonzero.)

Lemma 7.36. *If $\mathcal{F}_\mu$ is nonzero and finite dimensional, then $\mathcal{F}_\mu$ forms an irreducible holomorphic representation under the right action of G and this representation has highest weight μ.*

We know that in any finite-dimensional irreducible representation, the highest weight must be dominant integral. The proposition thus implies that if $\mathcal{F}_\mu$ is finite dimensional and nonzero, then μ must be dominant.

Proof. Assume that $\mathcal{F}_\mu$ is nonzero and finite dimensional. Then, $\mathcal{F}_\mu$ forms a representation under the right action of G. Because the elements of $\mathcal{F}_\mu$ are holomorphic and the right action of G on itself is holomorphic, it can be shown that $\mathcal{F}_\mu$ is a holomorphic representation of G and, thus, there is an associated complex-linear action of $\mathfrak{g}$ on $\mathcal{F}_\mu$. By complete reducibility of the representations of $\mathfrak{g}$, $\mathcal{F}_\mu$ decomposes as a direct sum of irreducible $\mathfrak{g}$-invariant subspaces. Each of these subspaces contains a nonzero highest weight vector

F_σ with some highest weight σ. Applying (7.32) to the representation Π given by $\Pi(h) = R_h|_{\mathcal{F}_\mu}$, we obtain that

$$R_a F_\sigma = X_\sigma(a) F_\sigma$$

for all $a \in B^+$. Meanwhile, since F_σ is an element of $\mathcal{F}_\mu$, we have that

$$L_b F_\sigma = \overline{X_\mu(b^*)} F_\sigma$$

for all $b \in B^-$. Thus, F_σ satisfies

$$F(bga) = \overline{X_\mu(b^*)} X_\sigma(a) F(g) \tag{7.38}$$

for all $g \in G$, $a \in B^+$, and $b \in B^-$.

If we take $g = I$ in (7.38), we obtain

$$F_\sigma(ba) = \overline{X_\mu(b^*)} X_\sigma(a) F_\sigma(I). \tag{7.39}$$

Now, every element of $\mathfrak{g}$ can be written as the sum of an element of $\mathfrak{b}^+$ and an element of $\mathfrak{b}^-$ (nonuniquely since $\mathfrak{b}^+ \cap \mathfrak{b}^- = \mathfrak{h}$). It follows (using the Inverse Function Theorem as in the proof of Theorem 2.27) that there is a neighborhood U of I in G with the property that every element g of U can be written (nonuniquely) as $g = ba$ with $a \in B^+$ and $b \in B^-$. This, together with (7.39), tells us that if $F_\sigma(I)$ were equal to zero, then F_σ would be zero on U and, hence (since F is holomorphic), that F_σ would be identically zero. Since we have chosen F_σ to be nonzero, we conclude that $F_\sigma(I) \neq 0$. We may, therefore, normalize F_σ so that $F_\sigma(I) = 1$ and we obtain that

$$F_\sigma(ba) = \overline{X_\mu(b^*)} X_\sigma(a). \tag{7.40}$$

Consider the connected subgroup T of K whose Lie algebra is $\mathfrak{t}$. Then, T is contained in both B^+ and B^-. This means that

$$F_\sigma(t) = \overline{X_\mu(t^*)} = X_\sigma(t). \tag{7.41}$$

However, since $T \subset K \subset \mathrm{U}(n)$ we have that $t^* = t^{-1}$ for all $t \in T$. Furthermore, we know that $\langle \mu, H \rangle$ is imaginary for all $H \in \mathfrak{t}$. This implies that $X_\mu(t)$ has absolute value 1 for all $t \in T$ and, thus, that $\overline{X_\mu(t)} = X_\mu(t)^{-1}$ for all $t \in T$. So, we see that

$$\overline{X_\mu(t^*)} = X_\mu(t^{-1})^{-1} = X_\mu(t)$$

for all $t \in T$. Thus, (7.41) becomes

$$X_\mu(t) = X_\sigma(t)$$

for all $t \in T$. This can occur only if $\sigma = \mu$.

Thus, the only highest weight that can occur in $\mathcal{F}_\mu$ is μ. Suppose now that $\mathcal{F}_\mu$ is reducible, so that its decomposition into irreducibles contains at least two terms, each of which (we have seen) must have highest weight μ. Let F_μ^1 and F_μ^2 be the highest weight vectors of these two subspaces. We may normalize each of them to be equal to 1 at the identity, and then both satisfy (7.40). This means that F_μ^1 is equal to F_μ^2 in a neighborhood of the identity and, thus, everywhere since the functions are holomorphic. So, the subspaces associated to F_μ^1 and F_μ^2 are equal after all and we conclude that $\mathcal{F}_\mu$ is irreducible. □

Lemma 7.37. *For all integral elements μ, $\mathcal{F}_\mu$ is finite dimensional.*

I will not attempt to prove this result here. The usual argument is to regard $\mathcal{F}_\mu$ as a space of holomorphic sections of a complex line bundle over the manifold G/B^-. Then, one shows that G/B^- is a compact complex manifold (by identifying G/B^- with K/T) and one makes use of a standard result from complex geometry that the space of holomorphic sections of a vector bundle over a compact complex manifold is finite dimensional.

Once Lemma 7.37 is established, it remains only to show that in the dominant integral case, the space $\mathcal{F}_\mu$ is nonzero. This is the most difficult step in the argument. Our strategy is to build an element F_μ in $\mathcal{F}_\mu$ that will be our highest weight vector. According to (7.40) (with $\sigma = \mu$), F_μ is determined on elements g of G that are of the form $g = ba$, with $a \in B^+$ and $b \in B^-$. Because B^+ and B^- have a nontrivial intersection, it is convenient to look at elements of the form na, with $a \in B^+$ and $n \in N^- \subset B^-$. Note that X_μ is identically equal to 1 on N^+, because (by definition) χ_μ is zero on $\mathfrak{n}^-$. However, for $n \in N^-$, we have that $n^* \in N^+$. Thus, F_μ should satisfy

$$F_\mu(na) = X_\mu(a), \quad a \in B^+, n \in N^-.$$

Note that as a vector space, $\mathfrak{g}$ decomposes as $\mathfrak{g} = \mathfrak{b}^+ \oplus \mathfrak{n}^-$. It can be shown that B^+ and N^- intersect only at the identity. It follows that each $g \in G$ can be decomposed as $g = na$, with $a \in B^+$ and $n \in N^-$, in at most one way. Thus, it makes sense to define a function on the set $N^-B^+ \subset G$ by defining how the function acts on the pair (a, n). Unfortunately, not every element of G is a product of an element of B^+ and an element of N^-; that is, N^-B^+ is a proper subset of G. So, we first define F_μ on N^-B^+ and then we must show that F_μ extends to a holomorphic function on G.

Lemma 7.38. *If μ is dominant integral, then the function F_μ on $N^-B^+ \subset G$ given by*

$$F(na) = X_\mu(a), \quad a \in B^+, n \in N^-,$$

has a unique holomorphic extension to all of G. The resulting holomorphic function on G is an element of $\mathcal{F}_\mu$.

I will give only a sketchy outline of the proof, taken from Jantzen (1987). (Since Jantzen works in the setting of algebraic group schemes, it is necessary

to translate some of the arguments into the language of complex Lie groups.) For each root α let A_α be the element of $K \subset G$ given by (6.21) in Section 6.6. (For each α, A_α is an element of $N(\mathfrak{t})$ that represents the Weyl group element w_α.) The calculations on pp. 201-202 of Jantzen (1987) then show how to extend F_μ holomorphically to the set of elements of the form

$$A_\alpha n a \tag{7.42}$$

with $a \in B^+$ and $n \in N^-$. One then has F_μ defined holomorphically on the set E consisting of $N^- B^+$ together with the elements of the form (7.42). One then argues that the complement of E in G has complex codimension 2, and so a standard result in complex analysis shows that F_μ extends holomorphically to all of G.

Putting Propositions 7.36, 7.37, and 7.38 together gives Theorem 7.35.

7.5.5 The $\mathsf{SL}(2;\mathbb{C})$ case

Let us see how the Borel–Weil construction works out in the case $G = \mathsf{SL}(2;\mathbb{C})$. In particular, we will show very explicitly in this case that the space $\mathcal{F}_\mu$ defined in Definition 7.33 is nonzero whenever μ is dominant integral. If X, Y, and H denote the usual basis elements for $\mathsf{sl}(2;\mathbb{C})$, then we may take $\mathfrak{h}$ to be the span of the element H, $\mathfrak{b}^+$ to be the span of the elements X and H, and $\mathfrak{n}^-$ to be the span of the element Y. In that case, the corresponding connected subgroups of $\mathsf{SL}(2;\mathbb{C})$ are

$$B^+ = \left\{ \begin{pmatrix} \alpha & \beta \\ 0 & \frac{1}{\alpha} \end{pmatrix} \middle| \, \alpha \in \mathbb{C}^*, \, \beta \in \mathbb{C} \right\},$$

$$N^- = \left\{ \begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix} \middle| \, \delta \in \mathbb{C} \right\}.$$

An integral element in this setting is simply an integer and a dominant integral element is a non-negative integer. For any integer m, the homomorphism $X_m : B^+ \to \mathbb{C}^*$ is given by

$$X_m \begin{pmatrix} \alpha & \beta \\ 0 & \frac{1}{\alpha} \end{pmatrix} = \alpha^m. \tag{7.43}$$

Let us now see which elements of $\mathsf{SL}(2;\mathbb{C})$ are contained in $N^- B^+$. We compute

$$\begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ 0 & \frac{1}{\alpha} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \alpha\delta & \beta\delta + \frac{1}{\alpha} \end{pmatrix}. \tag{7.44}$$

Consider a matrix

$$g = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \tag{7.45}$$

with determinant one. If $x \neq 0$, then we can express g uniquely in the form of (7.44) if we take $\alpha = x$, $\beta = y$, and $\delta = z/x$. (It is easy to check that with

this choice of α, β, and δ, the $(1,1)$, $(1,2)$, and $(2,1)$ entries in (7.44) agree with the corresponding entries of g. The condition that g have determinant one then guarantees that the $(2,2)$ entry in (7.44) also agrees with the $(2,2)$ entry in g.) If $x = 0$, then g cannot be decomposed into the form (7.44), since the $(1,1)$ entry on the right-hand side in (7.44) cannot be zero. So, the set $N^- B^+$ inside $\mathsf{SL}(2;\mathbb{C})$ is precisely the set of matrices in $\mathsf{SL}(2;\mathbb{C})$ whose $(1,1)$ entry is nonzero.

This calculation (specifically that $\alpha = x$) together with (7.43) shows us that the function F_m in Lemma 7.38 satisfies

$$F_m\begin{pmatrix} x & y \\ y & w \end{pmatrix} = x^m$$

on the set of matrices in $\mathsf{SL}(2;\mathbb{C})$ with $x \neq 0$. When, then, does the function F_m extend to a holomorphic function on all of $\mathsf{SL}(2;\mathbb{C})$? Clearly, it extends precisely when $m \geq 0$. Is the resulting function F_m ($m \geq 0$) an element of the space $\mathcal{F}_\mu$ defined in Definition 7.33? Let us compute and see. The elements of B^- are those of the form

$$b = \begin{pmatrix} \alpha & 0 \\ \beta & \frac{1}{\alpha} \end{pmatrix}. \tag{7.46}$$

Now, if b is as in (7.46) and g as in (7.45), then the $(1,1)$ entry in bg is αx. So, $F_m(g) = x^m$ and $F_m(bg) = \alpha^m x^m$. Then, we compute that

$$\overline{X_m(b^*)} = \overline{X_m\begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ 0 & \frac{1}{\bar{\alpha}} \end{pmatrix}} = \alpha^m.$$

Thus, indeed, $F_m(bg) = \overline{X_m(b^*)} F(g)$ and F_m is a (nonzero!) element of $\mathcal{F}_\mu$.

It is straightforward to extend this calculation to the case of $\mathsf{SL}(n;\mathbb{C})$. See Exercise 6.

7.6 Further Results

Although the main result about the representations of a complex semisimple Lie algebra is the theorem of the highest weight, there are many other useful results about the representations. This section presents some of these results, mostly without proofs. We continue to assume the notation established at the beginning of this chapter.

7.6.1 Duality

Recall from Section 4.7 the notion of the dual representation π^* associated to a finite-dimensional representation π of a group or Lie algebra. We know in general that π^* is irreducible if and only if π is irreducible and that $(\pi^*)^*$ is equivalent (as a representation) to π. In the case of representations of semisimple Lie algebras, we have the following result. (Compare Exercise 2 in Chapter 5.)

Proposition 7.39. *If π is an irreducible finite-dimensional representation of $\mathfrak{g}$, then the weights of π^* are the negatives of the weights of π. Specifically, if μ is a weight of π then $-\mu$ is a weight of π^* with the same multiplicity as μ.*

Proof. Let V be the space on which π acts, let μ be a weight of π, and let V_μ be the corresponding weight space. We know that V is the direct sum of V_μ and the weight spaces for the other weights σ of π. Now let ϕ be a linear functional on V_μ, and extend ϕ to a linear functional on all of V by setting ϕ to zero on all weight spaces V_σ, $\sigma \neq \mu$. Let us compute how the operators $\pi^*(H)$, $H \in \mathfrak{h}$, act on ϕ. If $v \in V_\mu$ then we have, by the definition of the dual representation,

$$[\pi^*(H)\phi](v) = [-\pi(H)^{tr}\phi](v) = -\phi(\pi(H)v) = -\langle \mu, H \rangle \phi(v).$$

If $v \in V_\sigma$, with $\sigma \neq \mu$, then both $\phi(v)$ and $\phi(\pi(H)v)$ are equal to zero, and so we still have

$$[\pi^*(H)\phi](v) = -\langle \mu, H \rangle \phi(v) \tag{7.47}$$

(both sides equal to zero). Thus, actually, (7.47) holds for all $v \in V$ and, therefore, $\pi^*(H)\phi = -\langle \mu, H \rangle \phi$. This shows that $-\mu$ is a weight of π^*, and it is easily seen that the multiplicity of $-\mu$ in π^* is the same as the multiplicity of μ in π. □

Now, we have classified representations of complex semisimple Lie algebras by their highest weights. So it is reasonable to ask how the highest weight of π^* is related to the highest weight of π. The answer is provided in the following result.

Theorem 7.40. *There exists a unique element w_0 of W such that for each dominant integral weight μ, $w_0 \cdot (-\mu)$ is, again, dominant integral. If π is an irreducible representation with highest weight μ, then π^* is an irreducible representation with highest weight $w_0 \cdot (-\mu)$.*

The proof of this result uses standard properties of the Weyl group (Section 8.7) and is omitted.

If it happens that $-I$ is an element of the Weyl group (which is the case for some semisimple Lie algebras but not for others), then we have $w_0 = -I$. This holds, for example, for $\mathfrak{g} = \mathsf{so}(5;\mathbb{C})$, whose root system is "$B_2$." See Section 8.5. In such cases, $w_0 \cdot (-\mu) = \mu$ and, so, the highest weight of π^* is, again, μ. Thus, in Lie algebras where $-I$ is an element of the Weyl group, every irreducible representation is equivalent to its dual.

For $\mathsf{sl}(3;\mathbb{C})$, $-I$ is not an element of the Weyl group. If we take the usual base $\Delta = \{\alpha_1, \alpha_2\}$ for the root system of $\mathsf{sl}(3;\mathbb{C})$, then w_0 will be the reflection about the line perpendicular to the root $\alpha_3 = \alpha_1 + \alpha_2$. (Compare Figures 5.2 and 5.3.) If μ_1 and μ_2 are the fundamental weights for $\mathsf{sl}(3;\mathbb{C})$ (circled in Figure 5.2), then every dominant integral element is of the form $\mu = m_1\mu_1 + m_2\mu_2$. Then, $w_0 \cdot (-\mu) = m_2\mu_1 + m_1\mu_2$. (Compare Exercise 3 in Chapter 5.) In the case of $\mathsf{sl}(3;\mathbb{C})$, a representation with highest weight $\mu = m_1\mu_1 + m_2\mu_2$ is equivalent to its dual if and only if $m_1 = m_2$.

7.6.2 The weights and their multiplicities

It is important to know not only the highest weight of a representation, but all of the weights, along with the multiplicities of the weights. We begin by looking at which weights occur and then turn to the multiplicities. For pictures of the weights of various representations, see Section 5.7, 8.5, and 8.6. Recall from Section 5.7 that the *convex hull* of a finite collection of points in a vector space is the smallest convex set containing those points.

Theorem 7.41. *Suppose that V is a representation with highest weight μ_0. Then, an element $\mu \in \mathfrak{h}$ is a weight of V if and only if the following two conditions are satisfied:*

1. *μ is contained in the convex hull of the orbit of μ_0 under the Weyl group.*
2. *$\mu_0 - \mu$ can be expressed as a linear combination of the positive simple roots with integer coefficients.*

Note that Condition 2 implies that μ is an integral element, since μ_0 and all of the roots are integral. On the other hand, there will typically be integral elements μ contained in the convex hull of the W-orbit of μ_0 that are *not* weights of V. After all, if μ is an integral element, then $\mu_0 - \mu$ is also an integral element, but this does not necessarily mean that $\mu_0 - \mu$ can be expressed as an integer linear combination of roots. See Section 8.10 for more on this issue.

As discussed after the statement of Theorem 7.15, every integral element occurs as a weight of some irreducible representation of $\mathfrak{g}$. In fact, given an integral element μ, it follows from Theorem 7.41 and standard results about the Weyl group (Section 8.7) that there exist infinitely many inequivalent irreducible representations of $\mathfrak{g}$ for which μ is a weight. See Exercises 8 and 9 in Chapter 8.

The proof of Theorem 7.41 is the same as the proof in the $\mathfrak{sl}(3;\mathbb{C})$ case, which is sketched in Section 5.7.

We now turn to the matter of the multiplicities. In the $\mathfrak{sl}(3;\mathbb{C})$ case, there is a simple pattern to the multiplicities, described in Section 5.7. In the general case, things are more complicated and there are two standard results about the multiplicities: Freudenthal's formula and Kostant's formula, both of which allow one, in principle, to compute all of the multiplicities in any representations. Although carrying out these computations in a particular case can be arduous, there exist computer programs that can do them. I present here (without proof) Kostant's formula, which gives the multiplicity of each weight directly. Freudenthal's formula gives a recursive algorithm for computing the multiplicities that may be more computationally efficient than Kostant's formula in high-rank examples. Freudenthal's formula is in Section 22 of Humphreys (1972) and Kostant's formula is in Section 24.

Suppose now that λ is an element of the root lattice (i.e., that λ can be expressed as a linear combination of roots with integer coefficients). Now, let $p(\lambda)$ denote the number of ways that λ can be expressed as a linear combination of *positive* roots with *non-negative* integer coefficients. If λ is higher than

zero, then there will be at least one such way; if λ is not higher than zero, then there will be no such way and $p(\lambda) = 0$. Consider, for example, the case of $\mathsf{sl}(3;\mathbb{C})$, in which we have three positive roots: α_1, α_2, and $\alpha_3 = \alpha_1 + \alpha_2$. If $\lambda = 2\alpha_1 + 3\alpha_2$, then we have

$$\lambda = 2\alpha_1 + 3\alpha_2 = \alpha_1 + 2\alpha_2 + \alpha_3 = \alpha_2 + 2\alpha_3$$

and, so, $p(\lambda) = 3$. More generally, every element of the root lattice for $\mathsf{sl}(3;\mathbb{C})$ can be expressed as $\lambda = k\alpha_1 + l\alpha_2$ for a unique pair of integers k and l. If either k or l is negative, then $p(\lambda) = 0$. If both k and l are non-negative, then $p(\lambda) = 1 + \min(k, l)$.

Theorem 7.42 (Kostant). *Suppose that V is a finite-dimensional irreducible representation of a complex semisimple Lie algebra $\mathfrak{g}$ with highest weight μ_0. If μ is a weight of V, then the multiplicity $m_{\mu_0}(\mu)$ is given by*

$$m_{\mu_0}(\mu) = \sum_{w \in W} \det(w)\, p(w \cdot (\mu_0 + \delta) - (\mu + \delta)),$$

where δ is half the sum of the positive roots.

Let us consider, first, the $w = I$ term in the sum, namely

$$p((\mu_0 + \delta) - (\mu + \delta)) = p(\mu_0 - \mu).$$

Since μ_0 is higher than (or equal to) μ, $\mu_0 - \mu$ is higher than (or equal to) zero, and so this term in the sum is nonzero. If the highest weight μ_0 is sufficiently far from the walls of the fundamental Weyl chamber and μ is sufficiently close to μ_0, then it will happen that for all $w \neq I$, $w \cdot (\mu_0 + \delta)$ is not higher than $\mu + \delta$. So, $p(w \cdot (\mu_0 + \delta) - (\mu + \delta))$ will be zero for all $w \neq I$, and in such cases, $m_{\mu_0}(\mu) = p(\mu_0 - \mu)$. (In the Verma module V_{μ_0}, described in Section 7.3, all weights μ have multiplicity $p(\mu_0 - \mu)$.)

As an example, consider the representation of $\mathsf{sl}(3;\mathbb{C})$ with highest weight $(0, 3)$ and consider the weight $\mu = 0$ occurring in this representation. Figure 7.1 illustrates the computations needed to apply Kostant's formula in this case. In the figure, black dots indicate the weights of the representations and the vertices of the outer hexagon indicate points of the form $w \cdot (\mu_0 + \delta)$, where μ_0 is the highest weight, $(0, 3)$. Of the six elements of the form $w \cdot (\mu_0 + \delta)$, only two are higher than $\mu + \delta$, namely $\mu_0 + \delta$ and $w_{\alpha_1} \cdot (\mu_0 + \delta)$. Since $\det(w_{\alpha_1}) = -1$, we obtain

$$\begin{aligned} m_{\mu_0}(\mu) &= p(\alpha_1 + 2\alpha_2) - p(2\alpha_2) \\ &= 2 - 1 = 1. \end{aligned}$$

In this representation, all weights have multiplicity one.

In the case of $\mathsf{sl}(3;\mathbb{C})$, it is possible to use Kostant's or Freudenthal's formula to show that the multiplicities follow the simple pattern described in Section 5.7. For other Lie algebras, the pattern of multiplicities can be more intricate.

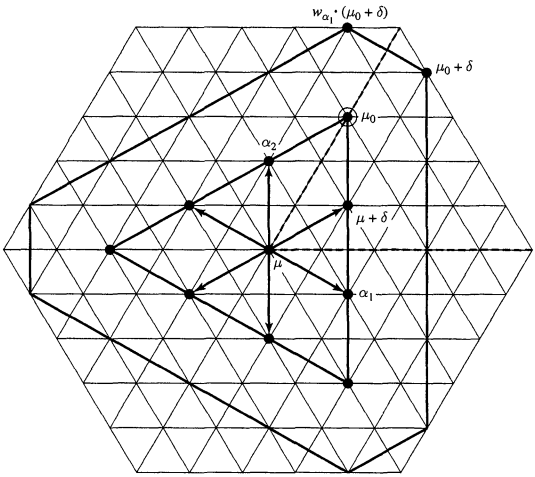


Fig. 7.1. Multiplicity calculation for highest weight (0,3)

7.6.3 The Weyl character formula and the Weyl dimension formula

Let K be a simply-connected compact Lie group, let $\mathfrak{k}$ be its Lie algebra, and let $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$, so that $\mathfrak{g}$ is a complex semisimple Lie algebra. Then, there is a one-to-one correspondence between the continuous representations of K and the complex-linear representations of $\mathfrak{g}$. Let $\mathfrak{t}$ be a maximal commutative subalgebra of $\mathfrak{k}$ and let T be the connected subgroup of K whose Lie algebra is $\mathfrak{t}$. Recall from Section 7.4 that the character of a representation Π of K is the function χ_Π on K defined by

$$\chi_\Pi(x) = \text{trace}(\Pi(x)).$$

Recall also the Weyl character formula (Theorem 7.28) which gives an expression for the restriction to T of the character of an irreducible representation with highest weight μ. (The statement of that theorem is in terms of the real weights, which are $1/i$ times the ordinary weights.)

Note that the value of the character at the identity is equal to the dimension of Π:

$$\chi_\Pi(I) = \text{trace}(I) = \dim \Pi.$$

This means that the dimension of the representation can be obtained by evaluating the character at the identity. Unfortunately, this procedure is not as simple as it sounds, since the Weyl character formula can be taken literally only at points where the denominator is nonzero. At the identity, the denominator of the Weyl character formula *is* zero and the whole formula is of the "zero over zero" form. Nevertheless, it is possible to use a version of

L'Hôpital's rule to evaluate the character at the origin so as to obtain the following result.

Theorem 7.43. *Suppose that π is an irreducible representation of $\mathfrak{g}$ with highest weight μ. Then, the dimension of π is given by*

$$\dim \pi = \frac{\prod_{\alpha \in R^+} \langle \alpha, \mu + \delta \rangle}{\prod_{\alpha \in R^+} \langle \alpha, \delta \rangle},$$

where R^+ denotes the set of positive roots and δ is half the sum of the positive roots.

Note that the denominator of this formula is a constant, independent of μ. So, the dimension is a polynomial function of the highest weight, and the degree of this polynomial is equal to the number of positive roots. The dimension formulas for the cases of $\mathsf{sl}(2;\mathbb{C})$ and $\mathsf{sl}(3;\mathbb{C})$, which we have discussed previously, are consistent with this. For $\mathsf{sl}(2;\mathbb{C})$, the dimension is $m+1$, and for $\mathsf{sl}(3;\mathbb{C})$, it is $\frac{1}{2}(m_1+1)(m_2+1)(m_1+m_2+2)$, reflecting that for $\mathsf{sl}(2;\mathbb{C})$ there is one positive root and for $\mathsf{sl}(3;\mathbb{C})$ there are three.

Let us now verify that Theorem 7.43 agrees with the results we have stated earlier for the $\mathsf{sl}(2;\mathbb{C})$ and $\mathsf{sl}(3;\mathbb{C})$ cases. For $\mathsf{sl}(2;\mathbb{C})$, we think of the roots and weights as being simply numbers (the eigenvalue of H), and the inner product as simply the product. The one positive root α is then the number 2 (reflecting that $\mathrm{ad}_H(X) = 2X$), and half the sum of the positive roots is the number 1. So, if m is the highest eigenvalue of H occurring in an irreducible representation, the dimension predicted by Theorem 7.43 is $(m+1)/1$, in agreement with Chapter 4.

For the case of $\mathsf{sl}(3;\mathbb{C})$, we note that

$$m_1 = \mu(H_1) = 2\frac{\langle \alpha_1, \mu \rangle}{\langle \alpha_1, \alpha_1 \rangle},$$
$$m_2 = \mu(H_2) = 2\frac{\langle \alpha_2, \mu \rangle}{\langle \alpha_2, \alpha_2 \rangle}.$$

Let us normalize all the roots so that $\langle \alpha, \alpha \rangle = 2$ (as in Section 5.6). Using that normalization of the inner product, we have $m_1 = \langle \alpha_1, \mu \rangle$ and $m_2 = \langle \alpha_2, \mu \rangle$. Letting $\alpha_3 = \alpha_1 + \alpha_2$, we have

$$\delta = \frac{1}{2}(\alpha_1 + \alpha_2 + \alpha_3) = \alpha_3.$$

We then note that $\langle \alpha_1, \delta \rangle = 1$, $\langle \alpha_2, \delta \rangle = 1$, and $\langle \alpha_3, \delta \rangle = 2$. So, the numerator in the dimension formula is

$$\begin{aligned}&(\langle \alpha_1, \mu \rangle + \langle \alpha_1, \delta \rangle)\,(\langle \alpha_2, \mu \rangle + \langle \alpha_2, \delta \rangle)\,(\langle \alpha_3, \mu \rangle + \langle \alpha_3, \delta \rangle) \\ &= (m_1+1)(m_2+1)(m_1+m_2+2)\end{aligned}$$

and the denominator is (1)(1)(2). Thus, the dimension formula in this case becomes

$$\dim \pi = \frac{1}{2}(m_1 + 1)(m_2 + 1)(m_1 + m_2 + 2),$$

as stated in Chapter 5.

7.6.4 The analytical proof of the Weyl character formula

Suppose that K is a simply-connected compact Lie group with Lie algebra $\mathfrak{k}$, so that $\mathfrak{g} := \mathfrak{k}_{\mathbb{C}}$ is a complex semisimple Lie algebra. We fix a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$ and we let T be the connected Lie subgroup of K whose Lie algebra is $\mathfrak{t}$. It can be shown that T is a torus; that is, T is isomorphic to $S^1 \times S^1 \times \cdots \times S^1$. Furthermore, it can be shown that every element of K is conjugate to an element of T; that is, given $A \in K$, there exists $B \in K$ such that $BAB^{-1} \in T$. (See Chapter IV of Bröcker and tom Dieck (1985).)

As in Section 7.4, we work with the real roots, which are the elements α of $\mathfrak{t}$ such that there exists a nonzero element X of $\mathfrak{g}$ with

$$[H, X] = i\langle \alpha, H \rangle X$$

for all H in $\mathfrak{t}$ (and, therefore, for all H in $\mathfrak{h}$). We consider also the integral real elements, which are those elements μ of $\mathfrak{t}$ such that $2\langle \alpha, \mu \rangle / \langle \alpha, \alpha \rangle$ is an integer for each real root α. As discussed in Section 7.4, for each integral real element μ, there is a function f_μ on T such that

$$f_\mu(e^H) = e^{i\langle \mu, H \rangle} \tag{7.48}$$

for all H in $\mathfrak{t}$. Functions of this form are called **torus characters**; they are in fact the characters of the irreducible representations of T, which are necessarily one-dimensional since T is commutative.

Proposition 7.44. *The set of torus characters form an orthonormal set in* $L^2(T, dt)$, *where* dt *is the normalized Haar measure on* T.

Actually, the torus characters form an orthonormal *basis*, but the completeness is not relevant here. This result can be thought of as the Peter–Weyl theorem for T (since on the commutative group T every function is a class function), but it can also be proved directly. See Exercise 2.

We now let δ denote half the sum of the positive real roots:

$$\delta = \frac{1}{2} \sum_{\alpha \in R^+} \alpha.$$

Note that the sum is over all positive roots, not just the positive simple roots. It can be shown that δ is an integral element (Section 8.7).

Definition 7.45. *The* ***Weyl denominator*** *is the function* $\sigma : T \to \mathbb{C}$ *given by*

$$\sigma(e^H) = \sum_{w \in W} \det(w)\, e^{i\langle w \cdot \delta, H\rangle}.$$

This is a well-defined function on T because δ, and therefore also $w \cdot \delta$ for all $w \in W$, is an integral element. This is (as the name suggests) the function that occurs in the denominator of the Weyl character formula.

The next key ingredient is the Weyl integral formula. I state there just the special case of the formula that applies to class functions; there is a more general version of the formula that applies to arbitrary functions on K.

Theorem 7.46. *Let* f *be a continuous class function on* K*, let* dA *denote the normalized Haar measure on* K*, and let* dt *denote the normalized Haar measure on* T*. Then,*

$$\int_K f(A)\, dA = \frac{1}{|W|} \int_T f(t)\, |\sigma(t)|^2\, dt.$$

Here, $|W|$ denotes the **order of the Weyl group** (i.e., the number of elements in W).

We are not going to prove this formula here but will only sketch the approach one uses. We consider the map $\Phi : T \times K/T \to K$ given by

$$\Phi(t, A) = AtA^{-1}, \quad t \in T,\ A \in K.$$

We have written Φ as a map from $T \times K$ into K, but if we replace A by At' for some t' in T, then (since T is commutative) the value of Φ does not change. Thus, we may think of Φ as mapping $T \times K/T$ into K; in fact, it maps *onto* K, because every point of K is conjugate to a point in T. We want to use the change-of-variables formula to change the integral over K into an integral over $T \times K/T$.

In the case of a class function, $f(AtA^{-1})$ is independent of A, so the integration over K/T drops out and we are left with just an integral over T. The factor of $|\sigma(t)|^2$ is essentially the Jacobian of the map Φ, which enters as a consequence of the change-of-variables theorem. The factor of $|W|$ arises because an element B of K can be conjugate to several different elements of T. In fact, if A is conjugate to a point t in T, then A is also conjugate to any point in T of the form $w \cdot t$, for w in the Weyl group. (This is true because if B is in $N(\mathfrak{t})$, then Ad_B preserves $\mathfrak{t}$ and thus also T.) "Generic" elements of K are conjugate to exactly $|W|$ elements of T and the map Φ is "generically" $|W|$-to-one, and this is why we need to divide by $|W|$ in the Weyl integral formula. To say the same thing another way, each "generic" conjugacy class intersects T in exactly $|W|$ points and, so, when integrating over T, we are "overcounting" the conjugacy classes by a factor of $|W|$. We then need to divide by $|W|$ to compensate for this.

There are two important features of the Weyl integral formula that we will use in the proof of the Weyl character formula. The first is the factor of $|W|$ which we discussed in the previous paragraph. This factor is extremely important. It is instructive to check the normalization of the Weyl integral formula by checking the formula in the case $f \equiv 1$. Then, the left-hand side of the integral formula is equal to 1, and the right-hand side is equal to $|W|/|W|$ since the terms in the Weyl denominator are orthonormal elements of $L^2(T, dt)$. (The points $w \cdot \delta$, $w \in W$, are *distinct* integral elements.) The second important feature is that the Jacobian factor $|\sigma(t)|^2$ is the absolute value squared of a very nice function σ. Of course, a Jacobian is always non-negative and, thus, has a non-negative square root. This non-negative square root, however, may not be a nice function—for example, it may not be smooth. (For example, the non-negative square root of the function x^2 is $|x|$, which is not differentiable at the origin.) Here, the function σ (which is *not* non-negative) is a very nice function, not only smooth but also having a very simple expansion in terms of torus characters.

We now turn to the proof of the Weyl character formula. The main ingredients are (1) the factor of $|W|$ in the Weyl integral formula, (2) the precise form of the function σ, and (3) the Peter–Weyl theorem. We do not need the full power of the Peter–Weyl theorem here. We need only that the characters of the irreducible representations of K have L^2 norm 1 with respect to the normalized Haar measure on K. This follows from fairly elementary "orthogonality relations." (See Section II.4 of Bröcker and tom Dieck (1985).)

We now start thinking about characters of representations of K.

Proposition 7.47. *If Π is a representation of K, then the restriction of the character χ_Π to T satisfies*

$$\chi_\Pi(e^H) = \sum_\mu m_\mu e^{i\langle \mu, H\rangle},$$

where the sum is over all the real weights of Π and where m_μ is the multiplicity of the weight μ.

Proof. The space V on which Π acts is the direct sum of the weight spaces V_μ associated to the real weights μ. On each V_μ, we have (by definition) that $\pi(H) = i\langle \mu, H\rangle I$. Thus, $\Pi(\exp H) = (\exp i\langle \mu, H\rangle) I$ on V_μ. Taking the trace of this equality gives the proposition. □

This is a "formula" of sorts for the character, but rather cumbersome because the number of weights involved gets larger and larger as the highest weight gets larger, and also because computing the multiplicities can be complicated. For example, trying to compute the dimension of Π from this formula is not so easy—one would have to sum up all the multiplicities of all the weights involved. By contrast, the Weyl character formula yields (with a bit of effort) the simple dimension formula given in Theorem 7.43.

We now turn to the proof of the Weyl character formula, as stated in Section 7.4.

Proof. If the representations Π has highest weight μ, then the character formula is equivalent to saying that

$$\left(\sum_{w\in W} \det(w)\, e^{i\langle w\cdot\delta, H\rangle}\right)\chi_\Pi\left(e^H\right) = \sum_{w\in W} \det(w)\, e^{i\langle w\cdot(\mu+\delta), H\rangle}. \tag{7.49}$$

Now, the Weyl denominator (the sum on the left in (7.49)) is a sum of torus characters with coefficients equal to ± 1 and with $|W|$ terms. The character χ_Π itself (restricted to T) is sum of torus characters with non-negative integer coefficients and with the number of terms getting larger and larger as the highest weight gets larger and larger. When we multiply the Weyl denominator and the character, we will get an apparently huge sum of torus characters with integer coefficients. Note that the product of two torus characters (as in (7.48)) with weights μ_1 and μ_2 is another torus character, with weight $\mu_1 + \mu_2$.

The Weyl character formula says when we multiply the character and the Weyl denominator, all but $|W|$ terms in this huge sum must cancel out. (This can be seen very explicitly in the $\mathsf{SU}(2)$ case. See (7.25) in Section 7.4.) Note that the possibility of cancellation arises because of the minus signs in the formula for the Weyl denominator and because the same torus character can arise in several different ways in the sum.

How do we know this cancellation actually occurs? Well, the Peter–Weyl theorem tells us that the character has L^2 norm 1 over K with respect to normalized Haar measure on K. Then, the Weyl integral formula (Theorem 7.46) implies that the product of the Weyl denominator σ and χ_Π is a function on T whose L^2 norm squared is equal to $|W|$. However, as we have just discussed, $\sigma\chi_\Pi$ is a large sum of torus characters with integer coefficients, and the torus characters are orthonormal (Proposition 7.44). So, the L^2 norm squared of $\sigma\chi_\Pi$ is the sum of the squares of the coefficients, and this must equal $|W|$. This means that there can be at most $|W|$ terms in the expression for $\sigma\chi_\Pi$ in terms of torus characters.

On the other hand, there is at least one torus character in the product $\sigma\chi_\Pi$ that cannot cancel out, namely $e^{i\langle\mu+\delta, H\rangle}$. Since this term comes from the highest weight in σ and the highest weight in χ_Π, it occurs only once and does not cancel out. Meanwhile, the set of weights in σ and in χ_Π are invariant under the Weyl group, and so the set of weights occurring in $\sigma\chi_\Pi$ is also invariant under the Weyl group. So, if $e^{i\langle\mu+\delta, H\rangle}$ does not cancel out, then neither does $e^{i\langle w\cdot(\mu+\delta), H\rangle}$, $w \in W$. This means that we have $|W|$ terms that cannot cancel out in $\sigma\chi_\Pi$; all other terms *must* cancel out or the L^2 norm of χ_Π would be larger than 1.

It remains only to check the coefficients of the terms of the form $e^{i\langle w\cdot(\mu+\delta), H\rangle}$ These terms will arise as the product of $\det(w)e^{i\langle w\cdot\delta, H\rangle}$ in σ and $m(w\cdot\mu)e^{i\langle\mu, H\rangle}$ in χ_Π. However, we know that μ has multiplicity one and that the multiplic-

ities are invariant under the action of the Weyl group. So, $m(w \cdot \mu) = 1$ and the coefficient of $e^{i\langle w\cdot(\mu+\delta),H\rangle}$ will be $\det(w)$. □

7.7 Exercises

1. For any complex number μ, consider the Verma module V_μ for $\mathsf{sl}(2;\mathbb{C})$ described in Section 7.3.4. Show that if μ is not a non-negative integer, then V_μ is irreducible.
2. Show that every continuous homomorphism of S^1 into $\mathbb{C}^*$ is of the form $e^{i\theta} \to e^{in\theta}$ for some $n \in \mathbb{Z}$. Show that the functions $\{e^{in\theta}\}_{n\in\mathbb{Z}}$ form an orthonormal set inside $L^2(S^1, d\theta/2\pi)$.
   Note: The functions $e^{in\theta}$ are the "torus characters" for the one-dimensional torus S^1. Compare Proposition 7.44.
3. Let T be the subgroup of $\mathsf{SU}(n)$ consisting of diagonal elements. (Then, T is a "maximal torus" in $\mathsf{SU}(n)$.) Show that every element A of $\mathsf{SU}(n)$ is conjugate to an element of T. Let W denote the Weyl group for $\mathsf{SU}(n)$, namely the permutation group acting on T by permuting the diagonal elements. Under what conditions is A conjugate to exactly $|W|$ elements of T?
4. If V is an irreducible representation of $\mathfrak{g}$ and u is a nonzero element of V, show that the map $u \to F_{u,v}$ in (7.34) is injective.
5. Verify directly the Weyl character formula for the representation of $\mathsf{sl}(3;\mathbb{C})$ with highest weight $(1,2)$, using the weights and multiplicities of this representation given in Section 5.7. To do this, consider the Weyl denominator σ given in Definition 7.45 and the formula for χ_Π given in Proposition 7.47. Now, compute $\sigma\chi_\Pi$ using that $e^{i\langle\mu,H\rangle}e^{i\langle\lambda,H\rangle} = e^{i\langle\mu+\lambda,H\rangle}$.
6. This exercise concerns the Borel–Weil construction in the case $G = \mathsf{SL}(n;\mathbb{C})$. Let B^+ denote the subgroup of $\mathsf{SL}(n;\mathbb{C})$ consisting of matrices of the form
$$a = \begin{pmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} \tag{7.50}$$
   and let N^- denote the subgroup of $\mathsf{SL}(n;\mathbb{C})$ consisting of lower triangular matrices with ones on the diagonal. Let B^- denote the matrices of the same form as in (7.50) except with the nonzero elements below the diagonal. Consider the homomorphism $X : B^+ \to \mathbb{C}^*$ given by
$$X(a) = \alpha_1^{m_1}(\alpha_1\alpha_2)^{m_2}\cdots(\alpha_1\alpha_2\cdots\alpha_{n-1})^{m_{n-1}}, \tag{7.51}$$
   where $m_1, \ldots, m_{n-1}$ are non-negative integers. For any $k = 1, 2, \ldots, n-1$ and any $g \in \mathsf{SL}(n;\mathbb{C})$, let $\det_k(g)$ denote the determinant of the $k \times k$ block in the upper left corner of g.
   Now, consider the function $F : \mathsf{SL}(n;\mathbb{C}) \to \mathbb{C}^*$ given by

$$F(g) = [\det_1(g)]^{m_1}[\det_2(g)]^{m_2} \cdots [\det_{n-1}(g)]^{m_{n-1}}.$$

Show that F is a polynomial in the entries of g and satisfies

$$F(bg) = \overline{X(b^*)}F(g)$$

for all $b \in B^-$ and all $g \in \mathsf{SL}(n;\mathbb{C})$. (Compare Definition 7.33.)
Note: It can be shown that the homomorphisms in (7.51) are precisely the X_μ's corresponding to dominant integral elements μ for $\mathsf{sl}(n;\mathbb{C})$.

7. Verify the Weyl dimension formula for the representations of $\mathsf{sl}(3;\mathbb{C})$ with highest weights $(1,2)$, $(2,2)$, and $(0,4)$, using the weights and multiplicities given in Section 5.7.
8. Consider the Cartan subalgebra for $\mathsf{sl}(n;\mathbb{C})$ given in Section 6.9 and consider the system of positive roots described in that section. Show that half the sum of the positive roots is an integral element.

8

More on Roots and Weights

8.1 Abstract Root Systems

In this section (and the next several sections), we consider root systems apart from their origins in semisimple Lie algebras. There are many results about root systems that are relevant to the understanding of semisimple Lie algebras but whose proofs involve only the root systems and not the Lie algebras from which they came. Therefore, it is convenient to separate the theory of root systems from Lie algebras.

Definition 8.1. *A **root system** is a finite-dimensional real vector space E with an inner product $\langle \cdot, \cdot \rangle$, together with a finite collection R of nonzero vectors in E satisfying the following properties:*

1. *The vectors in R span E.*
2. *If α is in R, then so is $-\alpha$.*
3. *If α is in R, then the only multiples of α in R are α and $-\alpha$.*
4. *If α and β are in R, then so is $w_\alpha \cdot \beta$, where w_α is the linear transformation of E defined by*
$$w_\alpha \cdot \beta = \beta - 2\frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}\alpha, \quad \beta \in E.$$
5. *For all α and β in R, the quantity*
$$2\frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}$$
is an integer.

*The dimension of E is called the **rank** of the root system and the elements of R are called **roots**.*

Property 2 is actually redundant, since $w_\alpha(\alpha) = -\alpha$. In the theory of symmetric spaces, there arise systems satisfying Properties 1, 2, 4, and 5 but

not Property 3. These are called "unreduced root systems." We will consider only "reduced root systems"—those satisfying Property 3 as well as the other four properties.

The map w_α is the reflection about the hyperplane perpendicular to α; that is, $w_\alpha \cdot \alpha = -\alpha$ and $w_\alpha \cdot \beta = \beta$ for all β in E that are perpendicular to α, as is easily verified from the formula for w_α. From this description, it should be evident that w_α is an orthogonal transformation of E with determinant -1.

We can interpret Property 5 geometrically in one of two ways. In light of the formula for w_α, Property 5 is equivalent to saying that $w_\alpha \cdot \beta$ should differ from β by an *integer* multiple of α. Alternatively, if we recall that the orthogonal projection of β onto α is given by $(\langle \beta, \alpha \rangle / \langle \alpha, \alpha \rangle)\alpha$, we note that the quantity in Property 5 is twice the coefficient of α in this projection. Thus, Property 5 is equivalent to saying that *the projection of β onto α is an integer or half-integer multiple of α.*

If the rank of the root system is one, then there is only one possibility: R must consist of a pair $\{\alpha, -\alpha\}$, where α is a nonzero element of E. This root system is customarily called "A_1" and is shown in Figure 8.1. In rank two, there are four possibilities, pictured in Figure 8.2 with their conventional names. In the case of $A_1 \times A_1$, the lengths of the horizontal roots are unrelated to the lengths of the vertical roots. In A_2, all roots have the same length; in B_2, the length of the longer roots is $\sqrt{2}$ times the length of the shorter roots; in G_2, the length of the longer roots is $\sqrt{3}$ times the length of the shorter roots. The angle between successive roots is 90° for $A_1 \times A_1$, 60° for A_2, 45° for B_2, and 30° for G_2. The reader is invited to verify that each of these systems is actually a root system. That every root system in rank two is actually one of the ones pictured here is proved in Section 8.5. Examples of rank-three root systems are given in Section 8.6.

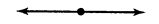


Fig. 8.1. The root system A_1

We have shown that one can associate a root system to every complex semisimple Lie algebra. It turns out that *every* root system arises in this way, although this is very far from obvious—see Section 8.9.

Definition 8.2. *If (E, R) is a root system, then the* ***Weyl group*** *W of R is the subgroup of the orthogonal group of E generated by the reflections w_α, $\alpha \in R$.*

By assumption, each w_α maps R into itself, indeed *onto* itself, since each $\alpha \in R$ satisfies $\alpha = w_\alpha \cdot (w_\alpha \cdot \alpha)$. It follows that every element of W maps R onto itself. Since the roots span E, a linear transformation of E is determined by its action on R. This shows (compare Proposition 6.29) that the Weyl

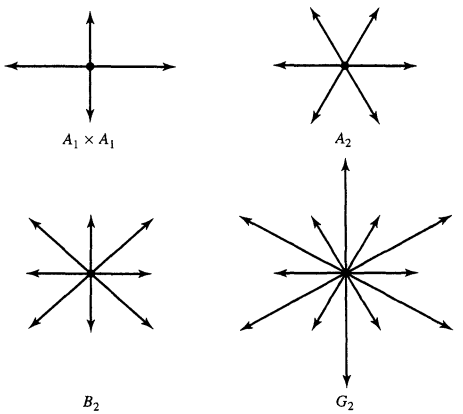


Fig. 8.2. The rank-two root systems

group is a finite subgroup of $\mathsf{O}(E)$ and may be regarded as a subgroup of the permutation group on R. We follow the notation of Chapter 6 and write the action of a Weyl group element w on an element H of E as $w \cdot H$.

Proposition 8.3. *Suppose (E, R) and (F, S) are root systems. Consider the vector space $E \oplus F$, with the natural inner product determined by the inner products on E and F. Then, $R \cup S$ is a root system in $E \oplus F$, called the* ***direct sum*** *of R and S.*

Here, we are identifying E with the subspace of $E \oplus F$ consisting of all vectors of the form $(e, 0)$ with e in E, and similarly for F. So, more precisely, $R \cup S$ means the elements of the form $(\alpha, 0)$ with α in R together with elements of the form $(0, \beta)$ with β in S. (Elements of the form (α, β) with $\alpha \in R$ and $\beta \in S$ are *not* in $R \cup S$.)

Proof. If R spans E and S spans F, then $R \cup S$ spans $E \oplus F$, so Condition 1 is satisfied. Conditions 2 and 3 are easily verified. For Conditions 4 and 5, consider, say, $\alpha \in R$ and some root β in $R \cup S$. If β is in R, then Conditions 4 and 5 hold because R is a root system. If β is in S, then $\langle \beta, \alpha \rangle = 0$ (since E is orthogonal to F inside $E \oplus F$) and so $w_\alpha \cdot \beta = \beta \in R \cup S$ and the quantity in Condition 5 is zero. The same argument applies if α is in S. □

Definition 8.4. *A root system (E, R) is called* ***reducible*** *if there exists an orthogonal decomposition $E = E_1 \oplus E_2$ with $\dim E_1 > 0$ and $\dim E_2 > 0$ such that every element of R is either in E_1 or in E_2. If no such decomposition exists, (E, R) is called* ***irreducible.***

If (E, R) is reducible, then it is not hard to see that the part of R in E_1 is a root system in E_1 and the part of R in E_2 is a root system in E_2. So, a

root system is reducible precisely if it can be realized as a direct sum of two other root systems. In the Lie algebra setting, the root system associated to a complex semisimple Lie algebra $\mathfrak{g}$ is irreducible precisely if $\mathfrak{g}$ is simple.

Definition 8.5. *Two root systems (E, R) and (F, S) are said to be* ***equivalent*** *if there exists an invertible linear transformation $A : E \to F$ such that A maps R onto S and such that for all $\alpha \in R$ and $\beta \in E$, we have*

$$A(w_\alpha \cdot \beta) = w_{A\alpha} \cdot A\beta.$$

A map A with this property is called an ***equivalence.***

Note that the linear map A is *not* required to preserve inner products, but only to preserve the reflections about the roots. It is easily seen that *if* A is an orthogonal transformation of E to F and if A takes R onto S, then A is an equivalence. However, not every equivalence is of this form. For example, we may take $F = E$ and $S = \lambda R$, where λ is a nonzero real number. (In this case (E, S) is again a root system, as is easily verified.) Then, $A = \lambda I$ is an equivalence of (E, R) with (E, S). To see this, note that $w_{\lambda\alpha}$ is the same as w_α since the hyperplane perpendicular to $\lambda\alpha$ is the same as the hyperplane perpendicular to α.

We see, then, that if one multiplies all of the roots in a root system by a nonzero constant, one gets another root system that is "the same as" (i.e., equivalent to) the original root system. Note that the quantity $2\langle\alpha, \beta\rangle / \langle\alpha, \alpha\rangle$ is unchanged if both α and β are multiplied by the same constant. So, the actual lengths of the roots in a root system are not important. Nevertheless, the *ratios* of lengths of different roots in the root systems are important (and unchanged if all the roots are multiplied by the same constant).

Proposition 8.6. *Suppose α and β are roots, α is not a multiple of β, and $\langle\alpha, \alpha\rangle \geq \langle\beta, \beta\rangle$. Then, one of the following holds:*

1. $\langle\alpha, \beta\rangle = 0$
2. $\langle\alpha, \alpha\rangle = \langle\beta, \beta\rangle$ *and the angle between α and β is 60° or 120°*
3. $\langle\alpha, \alpha\rangle = 2\langle\beta, \beta\rangle$ *and the angle between α and β is 45° or 135°*
4. $\langle\alpha, \alpha\rangle = 3\langle\beta, \beta\rangle$ *and the angle between α and β is 30° or 150°*

So, if two roots are not multiples of one another and are not perpendicular to one another, then the ratio of the length of the longer to the length of the shorter must be 1, $\sqrt{2}$, or $\sqrt{3}$, and for each case, there is only one possible acute angle and one possible obtuse angle. If the roots are perpendicular, then there is no constraint on the ratio of their lengths. The rank-two examples in Figure 8.2 demonstrate that each of the angles and length ratios permitted by Proposition 8.6 actually occurs. Figure 8.3 shows the allowed angles and length ratios, for the case of an acute angle.

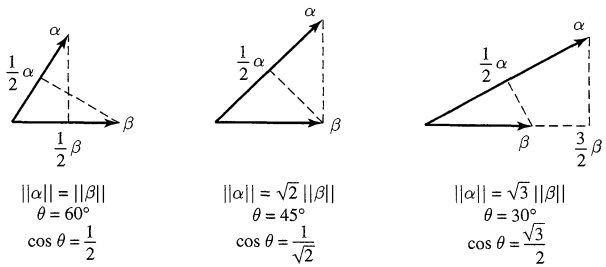


Fig. 8.3. Basic lengths and angles

Proof. Suppose that α and β are roots and let $m_1 = 2\langle\alpha,\beta\rangle/\langle\alpha,\alpha\rangle$ and $m_2 = 2\langle\beta,\alpha\rangle/\langle\beta,\beta\rangle$. Assume $\langle\alpha,\alpha\rangle \geq \langle\beta,\beta\rangle$. (If not, reverse the labeling of the two.) By Property 5, m_1 and m_2 must be integers. Note that

$$m_1 m_2 = 4\frac{\langle\alpha,\beta\rangle^2}{\langle\alpha,\alpha\rangle\langle\beta,\beta\rangle} = 4\cos^2\theta, \tag{8.1}$$

where θ is the angle between α and β, and that

$$\frac{m_2}{m_1} = \frac{\langle\alpha,\alpha\rangle}{\langle\beta,\beta\rangle} \tag{8.2}$$

whenever $\langle\alpha,\beta\rangle \neq 0$. From (8.1), we conclude that $0 \leq m_1 m_2 \leq 4$. If $m_1 m_2 = 0$, then $\cos\theta = 0$, so α and β are perpendicular. If $m_1 m_2 = 4$, then $\cos^2\theta = 1$ so θ is zero or 180° (i.e., α and β are multiples of one another (and so $\alpha = \pm\beta$)).

The remaining possible values for $m_1 m_2$ are 1, 2, and 3, which we consider in turn. If $m_1 m_2 = 1$, then $\cos^2\theta = 1/4$, so θ is 60° or 120°. Since m_1 and m_2 are both integers, we must have $m_1 = 1$ and $m_2 = 1$ or $m_1 = -1$ and $m_2 = -1$, and in either case, (8.2) tells us that α and β have the same length, and we get Case 2 of the proposition.

If $m_1 m_2 = 2$, then $\cos^2\theta = 1/2$ so θ is 45° or 135°. Since we assume $\langle\alpha,\alpha\rangle \geq \langle\beta,\beta\rangle$, (8.2) tells us that $|m_2| \geq |m_1|$, so we have either $m_2 = 2$, $m_1 = 1$ or $m_2 = -2$, $m_1 = -1$, and we are in Case 3 of the proposition.

Finally, if $m_1 m_2 = 3$, then $\cos^2\theta = 3/4$ and so θ is 30° or 150°. Since we assume $\langle\alpha,\alpha\rangle \geq \langle\beta,\beta\rangle$, (8.2) tells us that $|m_2| \geq |m_1|$, so we have either $m_2 = 3$, $m_1 = 1$ or $m_2 = -3$, $m_1 = -1$, and we are in Case 4 of the proposition.

In the last three cases, having m_1 and m_2 both positive corresponds to an acute angle ($\langle\alpha,\beta\rangle > 0$) and having m_1 and m_2 both negative corresponds to an obtuse angle ($\langle\alpha,\beta\rangle < 0$). □

Corollary 8.7. *Suppose that α and β are roots. If the angle between α and β is strictly obtuse (i.e., strictly between 90° and 180°), then $\alpha+\beta$ is a root. If the angle between α and β is strictly acute (i.e., strictly between 0 and 90°), then $\alpha-\beta$ and $\beta-\alpha$ are roots.*

Proof. The proof is by examining each of the three obtuse angles and each of the three acute angles allowed by Proposition 8.6. Consider first the acute case and adjust the labeling so that $\langle \alpha, \alpha \rangle \geq \langle \beta, \beta \rangle$. An examination of Cases 2, 3, and 4 in the proposition (see Figure 8.3) shows that in each of these cases, the projection of β onto α is equal to $\frac{1}{2}\alpha$. (This corresponds to $m_1 = 1$ in the notation of the proof of the proposition.) In that case, $w_\alpha \cdot \beta = \beta - \alpha$. Thus, $\beta - \alpha$ and $\alpha - \beta = -(\beta - \alpha)$ are roots. In the obtuse case (with $\langle \alpha, \alpha \rangle \geq \langle \beta, \beta \rangle$), we have the projection of β onto α equal to $-\frac{1}{2}\alpha$, and, so, $w_\alpha \cdot \beta = \alpha + \beta$. □

8.2 Duality

Definition 8.8. *If (E, R) is a root system, then for each root $\alpha \in R$, define the* ***co-root*** *H_α by*

$$H_\alpha = 2\frac{\alpha}{\langle \alpha, \alpha \rangle}.$$

The set of all co-roots is denoted $R^\vee$ and is called the ***dual root system*** *to R.*

Proposition 6.26 shows that this definition is consistent with the use of the term "co-root" in Chapter 6. Property 5 in the definition of a root system may be restated as the condition that $\langle \beta, H_\alpha \rangle$ be an integer for all roots α and β. We compute that

$$\langle H_\alpha, H_\alpha \rangle = 4\frac{\langle \alpha, \alpha \rangle}{\langle \alpha, \alpha \rangle^2} = \frac{4}{\langle \alpha, \alpha \rangle}$$

and, therefore,

$$2\frac{H_\alpha}{\langle H_\alpha, H_\alpha \rangle} = 2\left(\frac{2\alpha}{\langle \alpha, \alpha \rangle}\right)\frac{\langle \alpha, \alpha \rangle}{4} = \alpha, \tag{8.3}$$

and so the formula for α in terms of H_α is exactly the same as the formula for H_α in terms of α. Furthermore, if we take the inner product of (8.3) with H_β, we see that

$$2\frac{\langle H_\alpha, H_\beta \rangle}{\langle H_\alpha, H_\alpha \rangle} = \langle \alpha, H_\beta \rangle = 2\frac{\langle \alpha, \beta \rangle}{\langle \beta, \beta \rangle}. \tag{8.4}$$

This means that the left-hand side of (8.4) is an integer.

Furthermore, since H_α is a multiple of α, the hyperplane perpendicular to α is the same as the hyperplane perpendicular to H_α. This means that the reflection associated to H_α is the same as the reflection associated to α (i.e., $w_{H_\alpha} = w_\alpha$). However, since w_α is an orthogonal transformation, we have

$$w_\alpha \cdot H_\beta = 2\frac{w_\alpha \cdot \beta}{\langle \beta, \beta \rangle} = 2\frac{w_\alpha \cdot \beta}{\langle w_\alpha \cdot \beta, w_\alpha \cdot \beta \rangle} = H_{w_\alpha \cdot \beta}.$$

This shows that the set of co-roots is invariant under each reflection w_α $(= w_{H_\alpha})$. This, together with (8.4), shows that $R^\vee$ is again a root system.

(The remaining properties of root systems for $R^\vee$ follow immediately from the corresponding properties of R.) So, we have established the following result.

Proposition 8.9. *If R is a root system, then $R^\vee$ is also a root system and the Weyl group for $R^\vee$ is the same as the Weyl group for R. Furthermore, $(R^\vee)^\vee = R$.*

Note from (8.4) that the integer associated to the pair (H_α, H_β) in $R^\vee$ is the same as the integer associated to the pair (α, β) (not (β, α)) in R.

It follows from (8.3) that $(R^\vee)^\vee = R$; that is, if $R^\vee$ is dual to R, then R is also dual to $R^\vee$. If all the roots in R have the same length L, as in the case of the root system associated to $\mathfrak{sl}(n;\mathbb{C})$, then $R^\vee$ is obtained from R by multiplying each of the elements of R by a constant (namely $2/L^2$). In that case, the root system $R^\vee$ will be equivalent to R. Even if not all of the elements of R have the same length, it could still happen that $R^\vee$ is equivalent to R. This is the case, for example, for the rank-two root systems B_2 and G_2 (Figure 8.2). In general, however, $R^\vee$ need not be equivalent to R. For example, the rank-three root systems B_3 and C_3 (Section 8.6) are dual to each other but not equivalent to each other.

8.3 Bases and Weyl Chambers

Definition 8.10. *A subset Δ of R is called a **base** for R if the following conditions hold:*

1. *Δ is a basis for E as a vector space.*
2. *Each root $\alpha \in R$ can be expressed as a linear combination of elements of Δ with integer coefficients and in such a way that the coefficients are either all non-negative or all nonpositive.*

*The roots for which the coefficients are non-negative are called **positive roots** and the others are called **negative roots** (relative to the base Δ). The set of positive roots relative to a fixed base Δ is denoted R^+. The elements of Δ are called the **positive simple roots**.*

Note that since Δ is a basis for E, each α can be expressed *uniquely* as a linear combination the elements of Δ. We require that Δ be such that the coefficients in the expansion of each $\alpha \in R$ be integers and such that all the nonzero coefficients have the same sign.

Proposition 8.11. *If α and β are distinct elements of a base Δ for R then, $\langle \alpha, \beta \rangle \leq 0$.*

Geometrically, this means that either α and β are perpendicular or the angle between them is obtuse.

Proof. Since $\alpha \neq \beta$, if we had $\langle \alpha, \beta \rangle > 0$, then the angle between α and β would be strictly between 0 and 90°. Then, by Corollary 8.7, $\alpha - \beta$ would be an element of R. Since the elements of Δ form a basis for E as a vector space, each element of R has a *unique* expansion in terms of elements of Δ, and the coefficients of that expansion are supposed to be either all non-negative or all nonpositive. However, the expansion of $\alpha - \beta$ has one positive and one negative coefficient. So, $\alpha - \beta$ must actually not be a root, which means $\langle \alpha, \beta \rangle \leq 0$. □

It is far from obvious that a base exists. The reader is invited to look at the rank-two root systems in Figure 8.2 and find a base for each one. We now prove that a base always exists and give a constructive method for finding one.

Proposition 8.12. *If E is a finite-dimensional real vector space and R is any finite subset of E not containing* 0, *then there exists a hyperplane V through the origin in E that does not contain any element of R.*

Proof. To prove this, we try to find a vector H in E such that $\langle H, \alpha \rangle$ is nonzero for each α in R (in which case H itself must certainly be nonzero). If we can find such an H, then we may take V to be the orthogonal complement of H, and V will be a hyperplane through the origin and (by the way H is constructed) V will not contain any element of R. How do we find H? Well, H cannot be contained in any of the hyperplanes $V_\alpha = \{H \in E | \langle \alpha, H \rangle = 0\}$. So, if we can prove that the finite collection of hyperplanes V_α cannot fill up all of E, then there will be points not in any V_α and so we can find H and, thus, V.

It remains to prove that the union of the finite collection of hyperplanes $\{V_\alpha\}_{\alpha \in R}$ cannot be all of E. This can be done, for example, using measure theory: Each V_α is a set of Lebesgue measure zero in E, and so the union of the V_α's is again a set of measure zero, and so the union cannot be all of E. □

Definition 8.13. *Let (E, R) be a root system. Let V be a hyperplane through the origin in E such that V does not contain any root. Choose one "side" of V and let R^+ denote the set of roots on this side of V. An element α of R^+ is called* ***decomposable*** *if there exist β and γ in R^+ such that $\alpha = \beta + \gamma$; if no such elements exist, α is called* ***indecomposable****.*

The "sides" of V can be defined as the connected components of the set $E - V$. Alternatively, choose a nonzero element μ of the (one-dimensional) orthogonal complement of V. Then, V is precisely the set of H in E for which $\langle \mu, H \rangle = 0$. The two "sides" of V are then the sets $\{H \in E | \langle \mu, H \rangle > 0\}$ and $\{H \in E | \langle \mu, H \rangle < 0\}$. Choosing a different nonzero μ in $V^\perp$ (which must be a constant multiple of the original choice) either leaves these two sets unchanged or interchanges them.

The notion of indecomposable is understood to be relative to the choice of V and the choice of a side of V. This construction of a base for R motivates

the use of the phrase "positive simple root" for the elements of the base Δ. In this construction, we first find the set of positive roots; then, the elements that are positive and "simple" (i.e., indecomposable) form the base.

Theorem 8.14. *Suppose (R, E) is a root system, V is a hyperplane through the origin in E not containing any element of R, and R^+ is the set of roots lying on one fixed side of V. Then, the set of indecomposable elements of R^+ is a base for R.*

The reader is invited to carry out this procedure for each of the rank-two root systems listed in Section 8.5.

Proof. Choose a nonzero vector γ in the orthogonal complement of V (which is one dimensional) in such a way that γ lies on the positive side of V (i.e., the side containing R^+). Then, V is simply the set of $H \in E$ such that $\langle \gamma, H \rangle = 0$ and the positive side of V is the set of $H \in E$ such that $\langle \gamma, H \rangle > 0$. Let Δ denote the set of indecomposable elements in R^+.

Step 1. Every $\alpha \in R^+$ can be expressed as a linear combination of elements of Δ with non-negative integer coefficients. Suppose not. Then, among all of the elements of R^+ that cannot be expressed in this way, choose α so that $\langle \gamma, \alpha \rangle$ is a small as possible. Certainly α cannot be an element of Δ, so α must be decomposable, $\alpha = \beta_1 + \beta_2$, with $\beta_1, \beta_2 \in R^+$. Now, β_1 and β_2 cannot both be expressible as linear combinations of elements of Δ with non-negative integer coefficients, or else α would be expressible in this way. However, $\langle \gamma, \alpha \rangle = \langle \gamma, \beta_1 \rangle + \langle \gamma, \beta_2 \rangle$, and since the numbers $\langle \gamma, \beta_1 \rangle$ and $\langle \gamma, \beta_2 \rangle$ are both positive, they must be smaller than $\langle \gamma, \alpha \rangle$, contradicting the minimality of α.

Step 2. If α and β are distinct elements of Δ, then $\langle \alpha, \beta \rangle \leq 0$. If we had $\langle \alpha, \beta \rangle > 0$, then by Corollary 8.7, $\alpha - \beta$ and $\beta - \alpha$ would both be roots, one of which would have to be positive. If $\alpha - \beta$ were positive, then we would have $\alpha = (\alpha - \beta) + \beta$ and α would be decomposable. If $\beta - \alpha$ were positive, then we would have $\beta = (\beta - \alpha) + \alpha$, and α would be decomposable. Since α and β are assumed indecomposable, we must have $\langle \alpha, \beta \rangle \leq 0$.

Step 3. The elements of Δ are linearly independent. Suppose we have

$$\sum_{\alpha \in \Delta} c_\alpha \alpha = 0 \tag{8.5}$$

for some collection of constants c_α. Then, we may separate the sum into those terms where $c_\alpha \geq 0$ and those where $c_\alpha = -d_\alpha < 0$ and we obtain

$$\sum c_\alpha \alpha = \sum d_\beta \beta \tag{8.6}$$

where the sums range over disjoint subsets of Δ and where $c_\alpha \geq 0$ and $d_\alpha \geq 0$. Let $u = \sum c_\alpha \alpha$. From (8.6), we have

$$\begin{aligned}\langle u, u\rangle &= \left\langle \sum c_\alpha \alpha, \sum d_\beta \beta \right\rangle \\ &= \sum \sum c_\alpha d_\beta \langle \alpha, \beta\rangle.\end{aligned}$$

However, c_α and d_α are non-negative and (by Step 2) $\langle \alpha, \beta\rangle \leq 0$. So, $\langle u, u\rangle \leq 0$, which can only happen if u is zero.

Now, if $u = 0$, then $\langle \gamma, u\rangle = \sum c_\alpha \langle \gamma, \alpha\rangle = 0$, which implies that all the c_α's are zero since $c_\alpha \geq 0$ and $\langle \gamma, \alpha\rangle > 0$. The same sort of reasoning shows that all the d_α's are zero, so that all of the coefficients in the original expansion (8.5) are zero. This shows that Δ is linearly independent.

Step 4. Δ *is a base.* We have shown that Δ is linearly independent and that all of the elements of R^+ can be expressed as linear combinations of elements of Δ with non-negative integer coefficients. The remaining elements of R, namely the elements of R^-, are simply the negatives of the elements of R^+, and so they can be expressed as linear combinations of elements of Δ with nonpositive integer coefficients. Since the elements of R span E, Δ must also span E and it is a base. □

Theorem 8.15. *Given any base* Δ *for* R*, there exists a hyperplane* V *such that* Δ *arises as in Theorem 8.14.*

Proof. If $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ is a base for R, then Δ is a basis for E in the vector space sense. Then, by elementary linear algebra, for any sequence of numbers $c_1, \ldots, c_r$ there exists a unique $\gamma \in E$ with $\langle \gamma, \alpha_k\rangle = c_k$, $k = 1, \ldots, r$. In particular, we can choose γ so that $\langle \gamma, \alpha_k\rangle > 0$ for $k = 1, \ldots, r$. Then, if R^+ denotes the positive roots with respect to Δ, we will have $\langle \gamma, \alpha\rangle > 0$ for all $\alpha \in R^+$, since α is a linear combination of elements of Δ with non-negative coefficients. So, all of the elements of R^+ lie on the same side of the hyperplane $V = \{H \in E \,|\, \langle \gamma, H\rangle = 0\}$. It is not hard to see that the elements of the original base Δ are indecomposable and so are contained in the base constructed in Theorem 8.14. However, the number of elements in a base must equal $\dim E$, so, actually, Δ *is* the base constructed in Theorem 8.14. □

Proposition 8.16. *If* Δ *is a base for* R*, then the set of all co-roots* H_α, $\alpha \in \Delta$, *is a base for the dual root system* $R^\vee$.

Proof. Choose a hyperplane V such that the base Δ for R arises as in Theorem 8.14, and call the side of V on which Δ lies the positive side. Let R^+ denote the set of positive roots in R relative to the base Δ. Then the co-roots H_α, $\alpha \in R^+$, also lie on the positive side of V, and all the remaining co-roots lie on the negative side of V. Thus, applying Theorem 8.14 to $R^\vee$, there exist $\beta_1, \ldots, \beta_r$ in R^+ such that $H_{\beta_1}, \ldots H_{\beta_r}$ form a base $\Delta^\vee$ for $R^\vee$. We want to show that $\{\beta_1, \ldots, \beta_r\} = \{\alpha_1, \ldots, \alpha_r\}$.

Now, since the β's are in R^+, they can be expanded as linear combinations of the α's with non-negative (integer) coefficients. On the other hand, for each $\alpha_k \in \Delta$, H_{α_k} lies on the positive side of V, and so H_{α_k} is a positive root relative

to the base $\Delta^\vee$ for $R^\vee$. This means that H_{α_k} can be expanded in terms of $H_{\beta_1}, \ldots, H_{\beta_r}$ with non-negative (integer) coefficients. Since each co-root is simply a positive multiple of the corresponding root, it follows that α_k can be written as a linear combination of $\beta_1, \ldots, \beta_r$ with non-negative coefficients.

We conclude, then, that each β_k can be expanded in terms of the α's with non-negative coefficients and each α_k can be expanded in terms of the β's with non-negative coefficients. Now, suppose there is some element of $\{\beta_1, \ldots, \beta_r\}$, which we may call β_1, that is not an element of $\Delta = \{\alpha_1, \ldots, \alpha_r\}$. Then β_1 cannot be a multiple of any α_k, since if β_1 were equal to some α_k then β_1 would be in Δ and if β_1 were equal to $-\alpha_k$ then β_1 would be on the negative side of V. Thus, the expansion of β_1 in terms of the α's must have at least two nonzero coefficients. Then choose some α_k so that the expansion of α_k in terms of the β's has a nonzero coefficient for β_1. (There must exist such an α_k, or else all the α's would be contained in the span of $\beta_2, \ldots, \beta_r$ and the α's would not span E.) Now, expand α_k in terms of the β's and then expand each of the β's in terms of the α's. Since all the coefficients in both expansions are non-negative and since the coefficient of β_1 in the expansion of α_k is nonzero, we will get an expansion of α_k in terms of the α_l's in which at least two of the coefficients are nonzero. This contradicts the linear independence of the α's, and so we must, after all, have $\{\beta_1, \ldots, \beta_r\} = \{\alpha_1, \ldots, \alpha_r\}$. □

Definition 8.17. *If $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ is a base for R, then the **open fundamental Weyl chamber** in E (relative to Δ) is the set of all H in E such that $\langle \alpha_k, H \rangle > 0$ for all $k = 1, \ldots, r$. The **closed fundamental Weyl chamber** in E (relative to Δ) is the set of all α in E such that $\langle \alpha_k, \alpha \rangle \geq 0$ for all $k = 1, \ldots, r$.*

Since $\alpha_1, \ldots, \alpha_r$ form a basis for E in the vector space sense, elementary linear algebra shows that for any collection $a_1, \ldots, a_r$ of real numbers, there exists a unique vector $H \in E$ with $\langle \alpha_i, H \rangle = a_i$, $i = 1, \ldots, r$. This shows that the fundamental Weyl chamber (open or closed) is nonempty. It is easily seen that the open fundamental Weyl chamber is an open set in E and that its closure is the closed fundamental Weyl chamber.

Definition 8.18. *For each $\alpha \in R$, let V_α denote the hyperplane perpendicular to α. Then, an **open Weyl chamber** in E (relative to R) is a connected component of the set*

$$E - \bigcup_{\alpha \in R} V_\alpha.$$

The following result shows that the use of the term "Weyl chamber" in Definition 8.18 is consistent with its use in Definition 8.17. (The remaining results of this section are offered without proof.)

Proposition 8.19. *For each open Weyl chamber C, there exists a unique base Δ_C for R such that C is the open fundamental Weyl chamber associated to Δ_C. The positive roots with respect to Δ_C are precisely those elements α of R such that α has positive inner product with each element of C.*

So, there is a one-to-one correspondence between Weyl chambers and bases. Given a base Δ, we construct a Weyl chamber by looking at vectors having a positive inner product with each element of Δ. Given a Weyl chamber C, we define R^+ to be the set of roots α that have positive inner product with each element of C and we then take Δ to be the set of indecomposable elements in R^+.

Theorem 8.20. *The Weyl group acts simply and transitively on the set of bases and also on the set of open Weyl chambers.*

This means, in particular, that the base is unique up to the action of the Weyl group.

8.4 Integral and Dominant Integral Elements

Definition 8.21. *An element μ of E is called an **integral element** if for all α in R, the quantity*

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle}$$

is an integer.

Definition 8.22. *If Δ is a base for R, then an integral element μ is called **dominant integral** if*

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \geq 0$$

*for all $\alpha \in \Delta$. A dominant integral element is called **strictly dominant** if*

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle} > 0$$

for all $\alpha \in \Delta$.

This means that an integral element is dominant if and only it is contained in closed fundamental Weyl chamber and strictly dominant if and only if it is contained in the open fundamental Weyl chamber.

Proposition 8.23. *If $\mu \in E$ has the property that*

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle}$$

is an integer for all $\alpha \in \Delta$, then the same holds for all $\alpha \in R$ and, thus, μ is an integral element.

Proof. This follows from Proposition 8.16. □

Suppose μ is an element of E for which $2\langle \mu, \alpha\rangle / \langle \alpha, \alpha\rangle$ is a non-negative integer for all α in Δ. Then, the proposition tells us that μ is an integral element and it then follows that μ is a dominant integral element.

Definition 8.24. *Let $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ be a base. Then, the **fundamental weights** (relative to Δ) are the elements $\mu_1, \ldots, \mu_r$ with the property that*

$$2\frac{\langle \mu_k, \alpha_l\rangle}{\langle \alpha_l, \alpha_l\rangle} = \delta_{kl}, \quad k, l = 1, \ldots, r. \tag{8.7}$$

Let us see that there really are elements with this property. Fix k with $1 \leq k \leq r$ and let V_k be the span of $\alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_r$. Then, V_k is a $(r-1)$-dimensional subspace of E and the orthogonal complement $V_k^{\perp}$ of V_k is one-dimensional. If μ is a nonzero element of $V_k^{\perp}$, then μ cannot be orthogonal to α_k, or else μ would be orthogonal to all of the α's and, hence, to every element of E, including μ itself. Now, set

$$\mu_k = \frac{1}{2}\mu\frac{\langle \alpha_k, \alpha_k\rangle}{\langle \mu, \alpha_k\rangle}$$

and μ_k will be the k^{th} fundamental weight. (The same sort of argument shows that the fundamental weights are unique.)

Geometrically, the k^{th} fundamental weight is the unique element of E that is perpendicular to each α_l, $l \neq k$, and whose orthogonal projection onto α_k is one-half of α_k. The reader should look back at the picture of the dominant integral elements for $\mathfrak{sl}(3;\mathbb{C})$ in Figure 5.2, identify the two fundamental weights, and verify that the fundamental weights have this property.

Once we have found the fundamental weights, then the set of dominant integral elements is precisely the set of linear combinations of the fundamental weights with non-negative integer coefficients, and the set of all integral elements is the set of linear combinations of fundamental weights with arbitrary integer coefficients.

Note that every root is an integral element and, therefore, any linear combination of roots with integer coefficients is also an integral element. However, it is not necessarily the case that every integral element is an integer linear combination of roots—see Section 8.10.

Definition 8.25. *If $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ is a base, then an element μ_0 is said to be **higher** than μ if $\mu_0 - \mu$ can be expressed as*

$$\mu_0 - \mu = a_1\alpha_1 + \cdots + a_r\alpha_r,$$

*with each $a_i \geq 0$. We equivalently say that μ is **lower** than μ_0 and we write this relation as $\mu_0 \succeq \mu$ or $\mu \preceq \mu_0$.*

It is understood that the notion of higher and lower is relative to the chosen base Δ.

8.5 Examples in Rank Two

8.5.1 The root systems

Figure 8.2 shows four examples of rank-two root systems. We now prove that every rank-two root system is equivalent to one of these four. So, let $R \subset \mathbb{R}^2$ be a root system. Let θ be the smallest angle occurring between any two vectors in R. Since the elements of R span $\mathbb{R}^2$, we can find two linearly independent vectors α and β in R. If the angle between α and β is greater than 90°, then the angle between α and $-\beta$ is less than 90°; thus, the minimum angle θ is at most 90°. Then, according to Proposition 8.6, θ must be one of the following: 90°, 60°, 45°, 30°.

Let α and β be two elements of R such that the angle between them is the minimum angle θ. I claim that for each integer n there must be a unique element of R whose angle with α is $n\theta$. To see this, apply to α the reflection w_β about the line perpendicular to β. Then, the vector $-w_\beta(\alpha)$ will be a vector that is at angle θ to β but on the opposite side of β from α, as shown in Figure 8.4. Thus, $-w_\beta(\alpha)$ is at angle 2θ to α. If we then apply to β the reflection associated to the vector $w_\beta(\alpha)$, the negative of this vector will be a vector at angle 3θ to α. Continuing in the same way, we can obtain vectors at angle $n\theta$ to α for all n. These vectors are unique since a nontrivial positive multiple of a root is not allowed to be a root.

Note that all of the possible values of θ evenly divide 360°, so, eventually, we will come around to α again. Note also that the vector at angle 2θ to α must have the same length as α, since it is obtained from α by applying the (length-preserving) reflection w_β. The same sort of reasoning shows that any two vectors in R whose angle is an even multiple of θ must have the same length. For pairs of vectors whose angle is an odd multiple of θ, the ratio of the lengths must be consistent with Proposition 8.6.

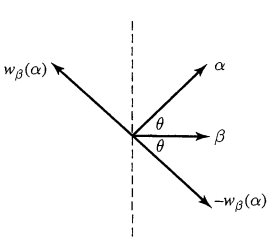


Fig. 8.4. α and $-w_\beta(\alpha)$

We now consider, in turn, each of the possible values for the minimum angle θ.

Case 1: $\theta = 90°$. In this case, the ratio of the lengths of perpendicular vectors is undetermined, so we have a pair of vectors α and $-\alpha$ and a perpendicular pair of vectors β and $-\beta$, with no restrictions on the lengths of α and β.

Case 2: $\theta = 60°$. In this case, Proposition 8.6 tells us that all vectors in R must have the same length, since the angle between nonparallel elements must be either 60° or 120°. Thus, we have six vectors of the same length, each one at an angle of 60° to the adjacent ones.

Case 3: $\theta = 45°$. In this case, Proposition 8.6 tells us that the vectors at 45° to each other must have a length ratio of $\sqrt{2}$ (or $1/\sqrt{2}$). Meanwhile, vectors at an angle of 90° to each other must have the same length since one can be obtained from the other by the reflection associated to the vector at 45° to both of them. So, we have eight vectors at 45° angles, with lengths alternating between a shorter length L and a longer length $\sqrt{2}L$.

Case 4: $\theta = 30°$. This case is similar to the previous one except that now the ratio of lengths of consecutive vectors must be $\sqrt{3}$ rather than $\sqrt{2}$. So, we have 12 vectors at 30° angles, with lengths alternating between some shorter length L and a longer length $\sqrt{3}L$.

These four cases correspond to the root systems $A_1 \times A_1$, A_2, B_2, and G_2 (in that order), as pictured in Figure 8.2. It is left as an exercise to verify that each of these collections of vectors is actually a root system. This is true essentially because for any pair of vectors in each system, the angles and length ratios are consistent with Proposition 8.6.

We also need to show that any two root systems falling under a particular case are equivalent. This is perhaps least obvious for Case 1. If we have one root system $R_1 = \{\pm\alpha_1, \pm\beta_1\}$ (with α_1 perpendicular to β_1) and another $R_2 = \{\pm\alpha_2, \pm\beta_2\}$ (with α_2 perpendicular to β_2), then there is a unique linear map A from $\mathbb{R}^2$ to $\mathbb{R}^2$ that takes α_1 to α_2 and β_1 to β_2. This map will be the desired equivalence, as is easily verified. Note that A need not be an isometry. The verification for the three remaining cases is left as an exercise—in those cases the equivalence will be a combination of a rotation and a dilation.

8.5.2 Connection with Lie algebras

The root systems $A_1 \times A_1$, A_2, and B_2 arise as root systems of "classical" Lie algebras as follows. The root system $A_1 \times A_1$ is the root system of $\mathsf{so}(4;\mathbb{C})$, which is isomorphic to $\mathsf{sl}(2;\mathbb{C}) \oplus \mathsf{sl}(2;\mathbb{C})$; A_2 is the root system of $\mathsf{sl}(3;\mathbb{C})$; and B_2 is the root system of $\mathsf{so}(5;\mathbb{C})$, which is isomorphic to $\mathsf{sp}(2;\mathbb{C})$. The root system G_2 is the root system of an "exceptional" Lie algebra, also called G_2. The Lie algebra G_2 can be represented as a Lie algebra of matrices (indeed every Lie algebra can!), but not in any particularly simple way.

8.5.3 The Weyl groups

We now compute the Weyl group in each case. Suppose our root system has n elements, in which case $\theta = 360/n$. There will then be $n/2$ reflections, one for

each pair $\pm\alpha$ of roots. If α and β are roots with angle ϕ between them, then the composition of the rotations w_α and w_β will be a rotation by angle $\pm 2\phi$, with the direction of the rotation depending on the order of the composition. To see that this is the case, note that w_α and w_β both have determinant -1 and so $w_\alpha w_\beta$ has determinant 1 and is, therefore, a rotation by some angle. To determine the angle, it suffices to apply $w_\alpha w_\beta$ to any nonzero vector, for example, β. However, $w_\beta(\beta) = -\beta$ and $w_\alpha(-\beta)$ is the vector at angle ϕ to α but on the opposite side of α as β, hence at angle 2ϕ to β. Then, the composition of a reflection w_α and a rotation by angle 2ϕ will be another reflection w_β, where β is a root at angle ϕ to α. (I leave it to the reader to verify this.) Thus, the set of $n/2$ reflections together with rotations by integer multiples of 2θ form a group; this is the Weyl group of the rank-two root system.

Therefore, if there are n elements in the rank-two root system, then the Weyl group consists of the $n/2$ reflections together with $n/2$ rotations, namely rotations by *even* integer multiples of $\theta = 360°/n$. The Weyl group is, therefore, the dihedral group on $n/2$ elements (i.e., the symmetry group of a regular $(n/2)$-gon). Note that in the case of A_2 the Weyl group consists of three reflections together with three rotations (by multiples of 120°). In this case, the Weyl group is not the full symmetry group of the root system: Rotations by 60° map R onto itself but are not elements of the Weyl group.

8.5.4 Duality

Recall that the map $\alpha \to 2\alpha/\langle\alpha,\alpha\rangle$ turns each root system into a new root system that is called the dual of the original root system. This new root system in general may not be equivalent to the original root system. However, in the rank-two case, one can see that the dual root system is always equivalent to the original system. Recall that duality makes long roots short and short roots long. So, in the case of B_2, the dual root system can be converted back to the original one by rotating by 45° and then rescaling all roots by a constant, and similarly for G_2.

8.5.5 Positive roots and dominant integral elements

For each of the four rank-two root systems, we select a base $\{\alpha_1, \alpha_2\}$ and then indicate the associated fundamental weights μ_1 and μ_2. Here, μ_1 is orthogonal to α_2 and its orthogonal projection onto α_1 is one-half of α_1, whereas μ_2 is orthogonal to α_1 and its orthogonal projection onto α_2 is one-half of α_2. The case of A_2 was shown in Figure 5.2. We now consider the remaining three cases. In each of Figures 8.5, 8.6, and 8.7, the fundamental weights for the relevant root system are circled and the other dominant integral elements are indicated by black dots. The dashed lines indicate the boundary of the fundamental Weyl chamber. The grid in the background indicates the set of all integral elements. Note that B_2 is presented here rotated by 45° from its

orientation in Figure 8.2. This rotation allows the set of integral elements to be a square lattice with edges oriented horizontally and vertically. Note also that $A_1 \times A_1$ is presented here with all roots having the same length, even though they are not required to do so. The reader should verify that, in each case, the set $\{\alpha_1, \alpha_2\}$ is actually a base. For example, in the case of G_2, the positive roots are α_1, α_2, $\alpha_2 + \alpha_1$, $\alpha_2 + 2\alpha_1$, $\alpha_2 + 3\alpha_1$, and $2\alpha_2 + 3\alpha_1$, and the remaining roots are the negatives of these six.

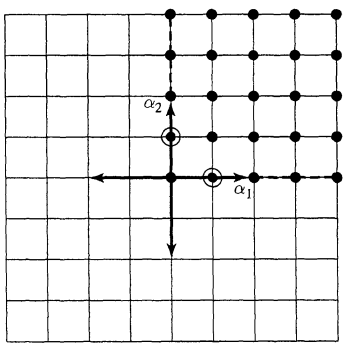


Fig. 8.5. Roots and dominant integral elements for $A_1 \times A_1$

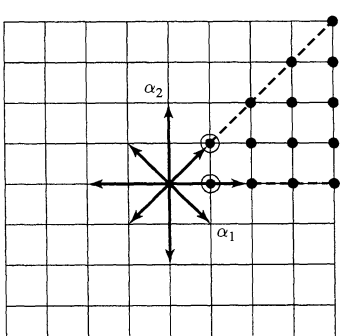


Fig. 8.6. Roots and dominant integral elements for B_2

8.5.6 Weight diagrams

Since we have already seen (in Chapter 5) weight diagrams for the A_2 case (i.e., the Lie algebra $\mathfrak{sl}(3;\mathbb{C})$), we will content ourselves here with one representative

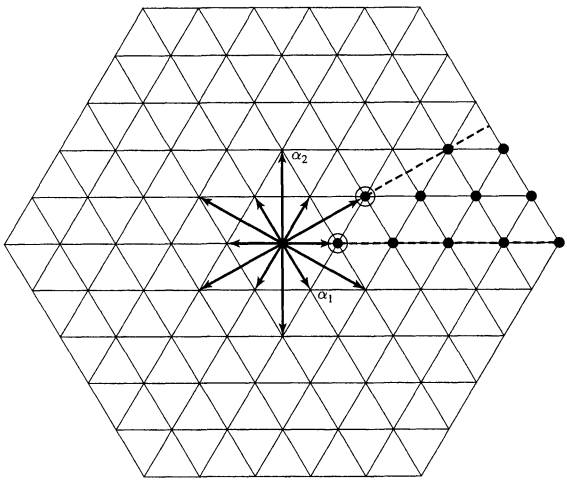


Fig. 8.7. Roots and dominant integral elements for G_2

diagram for each of the remaining three rank-two cases, $A_1 \times A_1$, B_2, and G_2 (corresponding to the Lie algebras $\mathsf{so}(4;\mathbb{C}) \cong \mathsf{sl}(2;\mathbb{C}) \oplus \mathsf{sl}(2;\mathbb{C})$, $\mathsf{so}(5;\mathbb{C}) \cong \mathsf{sp}(2;\mathbb{C})$, and G_2).

We make use of Theorem 7.41, which asserts that if π is representation with highest weight μ_0, then μ is a weight for π if and only if (1) μ is contained in the convex hull of the orbit of μ_0 under the Weyl group and (2) $\mu_0 - \mu$ can be expressed as a linear combination of roots with integer coefficients. Note that in most cases, there will be integral elements contained in the convex hull of the orbit of μ_0 that are not weights of π because they do not satisfy the second property in Theorem 7.41. In the case of $A_1 \times A_1$, the Weyl group is generated by reflections about the x-axis and the y-axis, so the orbit of a typical point is a rectangle. For B_2, every element of the Weyl group is either a rotation by a multiple of 90° or a combination of the reflection about α_1 and such a rotation. So, to obtain the orbit of μ_0, we first look at the two points μ_0 and $w_{\alpha_1} \cdot \mu_0$ and then at the eight points obtained by rotating μ_0 and $w_{\alpha_1} \cdot \mu_0$ by multiples of 90°. (If μ_0 is on the boundary of the fundamental Weyl chamber, then these eight points will not be distinct.) For G_2, the orbit of μ_0 is obtained by looking at rotations by multiples of 60° applied to the two points μ_0 and $w_{\alpha_1} \cdot \mu_0$, yielding 12 (generically distinct) points.

In each figure, a black dot indicates a weight of the given representation and the highest weight is circled. A number next to a dot indicates the multiplicity of the corresponding weight. A dot without a number indicates a weight of multiplicity one. One set of dashed lines indicates the boundary of the fundamental Weyl chamber and another set of dashed lines indicates the boundary of the set of points lower than the highest weight.

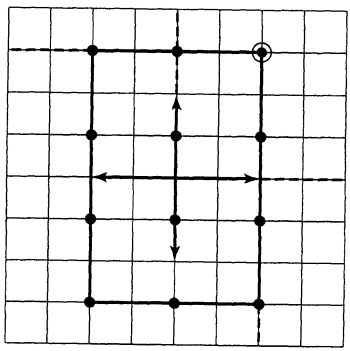


Fig. 8.8. Typical weight diagram for $A_1 \times A_1$

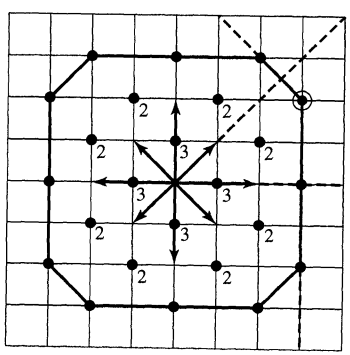


Fig. 8.9. Typical weight diagram for B_2

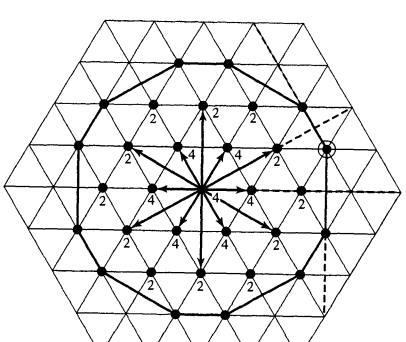


Fig. 8.10. Typical weight diagram for G_2

8.6 Examples in Rank Three

In rank three, we can have a reducible root system, which must be a direct sum of A_1 with one of the rank-two root systems described in the previous section. In this section, we will consider only the *irreducible* root systems of rank three. There are, up to equivalence, three irreducible root systems in rank three, customarily denoted A_3, B_3, and C_3. They all arise as root systems of classical Lie algebras. The root system A_3 comes from the Lie algebra $\mathsf{sl}(4;\mathbb{C})$ (or $\mathsf{so}(6;\mathbb{C})$, which is isomorphic to $\mathsf{sl}(4;\mathbb{C})$). The root system B_3 comes from the Lie algebra $\mathsf{so}(7;\mathbb{C})$ and the root system C_3 comes from the Lie algebra $\mathsf{sp}(3;\mathbb{C})$. The connection with Lie algebras is described in greater detail later in this section and in Section 8.8.

The color plates (p. 162) show models related to rank-three root systems, built using the Zome system, available at www.zometool.com. Many more images can be seen on the author's web site, at www.nd.edu/~bhall. The images in the color plates were modeled using the vZome program, available free from the programmer, Scott Vorthmann, at www.vorthmann.org/zome. The models were then rendered in the POV program by Charles Albrecht. The reader is encouraged to obtain enough Zome to make the root systems for him- or herself. The root systems require the "green lines" which are not part of the simplest Zome kits. The model of the dominant integral elements for C_3 (Plate 6) also makes use of "half-length blue lines" which are available by special order. (However, one can easily "fake it" using only whole blue lines.) In each plate, the roots marked with red balls form a base for the relevant root system.

The first three color plates show the root systems A_3, B_3, and C_3. Let us make a few observations about these root systems. First, B_3 and C_3 are obtained by adding more vectors to A_3. In higher dimensions, B_n and C_n are obtained by adding vectors to D_n, not to A_n. However, we have a low-dimensional coincidence in rank three, namely that A_3 is the same as D_3, reflecting that $\mathsf{sl}(4;\mathbb{C})$ is isomorphic to $\mathsf{so}(6;\mathbb{C})$.

Second, B_3 and C_3 are dual to each other. Specifically, let us normalize the inner product so that the green lines have length $\sqrt{2}$. In that case, the blue lines in B_3 (Plate 2) have length 1. Thus when we replace each root α by $2\alpha/\langle\alpha,\alpha\rangle$, the green lines are unchanged and the blue lines are replaced by blue lines of twice the length. This gives C_3 (Plate 3).

Third, in both B_3 and C_3, the long roots by themselves form a root system, as do the short roots by themselves. For B_3, the long roots form A_3 and the short roots form $A_1 \times A_1 \times A_1$; for C_3, it is the reverse.

I will not make any attempt in this section to prove that these are the only irreducible root systems in rank three. See the next section for a discussion of the classification of root systems (and semisimple Lie algebras).

The calculations in Section 8.8 will show that each of these rank-three root systems arises from one of the Lie algebras $\mathsf{sl}(4;\mathbb{C}) \cong \mathsf{so}(6;\mathbb{C})$, $\mathsf{so}(7;\mathbb{C})$, and $\mathsf{sp}(3;\mathbb{C})$, as follows.

Case 1: $\mathsf{sl}(4;\mathbb{C})$ or $\mathsf{so}(6;\mathbb{C})$. In these two cases (which are isomorphic), one obtains 12 roots, all of the same length and with all angles between nonparallel roots equal to 60°, 90°, or 120°. This is the A_3 root system.

Case 2: $\mathsf{so}(7;\mathbb{C})$. In this case, one gets 12 "long" roots that form, among themselves, a root system isomorphic to A_3. One gets, in addition, six "short" roots that are shorter by a factor of $\sqrt{2}$ than the long roots. The short roots come in pairs of the form $\pm\alpha$, with each pair perpendicular to the other two pairs of short roots, so that the short roots among themselves form a root system isomorphic to $A_1 \times A_1 \times A_1$. The total root system (including both long and short roots) is the B_3 root system.

Case 3: $\mathsf{sp}(3;\mathbb{C})$. This case is the same as Case 2 except that the 12 roots forming A_3 are short and the 6 roots forming $A_1 \times A_1 \times A_1$ are longer by a factor of $\sqrt{2}$ than the 12 short roots. This is the C_3 root system, which is the dual root system to B_3.

We now look at the set of dominant integral elements for each of the three irreducible root systems in rank three. In each case, we identify three positive simple roots (indicated by a red ball), denoted α_1, α_2, and α_3, and we then consider the associated fundamental weights, denoted μ_1, μ_2, and μ_3. This means that μ_1 is orthogonal to α_2 and α_3 and that the orthogonal projection of μ_1 onto α_1 is equal to $\frac{1}{2}\alpha_1$, and similarly for μ_2 and μ_3. In each case, the plate shows elements of the form $n_1\mu_1+n_2\mu_2+n_3\mu_3$, with each n_k ranging over the set $\{0,1,2\}$. In the case of A_3 (Plate 4) we obtain two yellow fundamental weights and one blue one. In the case of B_3 and C_3 (Plates 5 and 6), we obtain one green fundamental weight, one blue one, and one yellow one. The directions for both roots and weights are the same in B_3 as they are for C_3; only the lengths have changed. (Compare Proposition 6.37.)

Plate 7 shows a representative weight diagram for the Lie algebra $\mathsf{sl}(4;\mathbb{C})$, whose root system is A_3. The highest weight of this representation is the sum of the three fundamental weights. (The image includes one "cell" of the set of dominant integral elements for reference.) The weights of this representation split up into four orbits under the Weyl group. The 24 vertices of the outer polyhedron make up a single orbit, and each weight in this orbit has multiplicity 1. The centers of the 8 hexagons break up into two orbits of the Weyl group, with four elements each. Each of these weights has multiplicity 2. Finally, the 6 vertices of the inner octahedron make up a single orbit, and the weights in this orbit have multiplicity 4. Counting the weights with their multiplicities shows that the dimension of the representation is 64.

8.7 Additional Properties

There are many other properties of roots and the Weyl group that are known and worth knowing. We have focused, up to now, on the most essential properties and on examples of root systems. We now list some of the remaining

properties without proof. Proofs may be found in any number of standard references, including Chapter III of Humphreys (1972) and Chapter V of Bröcker and tom Dieck (1985). (See also Chapter V of Serre (1987).) In all of these properties, R is a root system in the Euclidean space E, Δ is a fixed base of R, C is the open fundamental Weyl chamber, and $\bar{C}$ is the closed fundamental Weyl chamber.

1. Every root is an element of some base.
2. If R is irreducible, then the Weyl group acts irreducibly on E.
3. If R is irreducible, then at most two different lengths of roots can arise in R and the Weyl group acts transitively on each length of root.
4. The reflections w_α, $\alpha \in \Delta$, generate the Weyl group for R.
5. If α is an element of Δ and β is a positive root different from α, then $w_\alpha \cdot \beta$ is a positive root; that is, w_α permutes the positive roots different from α.
6. Each orbit of the Weyl group contains exactly one point in $\bar{C}$.
7. If $\mu \in E$ is not contained in any of the hyperplanes perpendicular to the roots, then they Weyl group acts freely on μ.
8. If μ_0 is an element of $\bar{C}$, then $\mu_0 \succeq w \cdot \mu_0$ for all $w \in W$.
9. If $\mu_0 \in \bar{C}$, then for all $\mu \in E$, μ is contained in the convex hull of the W-orbit of μ_0 if and only if $w \cdot \mu \preceq \mu_0$ for all $w \in W$.
10. Let δ denote half the sum of the positive roots. Then, for each positive simple root α,
$$\frac{2\langle \alpha, \delta \rangle}{\langle \alpha, \alpha \rangle} = 1.$$
Thus, δ is a strictly dominant integral element and every strictly dominant integral element μ can be expressed as $\mu = \lambda + \delta$, where λ is a dominant integral element.

Let us consider examples of the ways in which these properties of root systems can be used. In Section 7.3, we needed to show that the weights of a certain quotient representation V_μ / U_μ were invariant under the action of the Weyl group. To do this, it is sufficient to show that the weights are invariant under the action of some generating set for W, and we use Property 4, that the w_α's with α in the base Δ are a generating set. It is not feasible to prove directly that the weights are invariant under all the w_α's, $\alpha \in R$; we make use of special properties of the w_α's, $\alpha \in \Delta$.

As another example, we use Property 9 to show that all the weights of the irreducible finite-dimensional representation with highest weight μ_0 are contained in the convex hull of the W-orbit of μ_0 (Theorem 7.41). We use Properties 7 and 10 in the proof of the Weyl character formula to show that the weights $w \cdot (\mu_0 + \delta)$, $w \in W$, are distinct. We also use the integrality of δ (part of Property 10) to show that the numerator and denominator of the Weyl character formula are well-defined functions on T (in the case that K is simply connected).

See the exercise for additional examples of how the results in this section can be used.

8.8 The Root Systems of the Classical Lie Algebras

In this section, we consider the root systems of the "classical" complex semisimple Lie algebras, namely $\mathsf{sl}(n;\mathbb{C})$, $\mathsf{so}(n;\mathbb{C})$, and $\mathsf{sp}(n;\mathbb{C})$. We have already worked out in detail the case of $\mathsf{sl}(n;\mathbb{C})$ in Section 6.9. The root system for $\mathsf{sl}(n;\mathbb{C})$ is denoted A_{n-1}. The analysis of the orthogonal algebras $\mathsf{so}(n;\mathbb{C})$ builds on the calculations in Exercises 12 and 13 of Chapter 6 and is divided into the cases n even and n odd. The analysis of the symplectic algebras $\mathsf{sp}(n;\mathbb{C})$ builds on Exercise 14 of Chapter 6.

8.8.1 The orthogonal algebras $\mathsf{so}(2n;\mathbb{C})$

The root system for $\mathsf{so}(2n;\mathbb{C})$ is denoted D_n. We consider $\mathsf{so}(2n;\mathbb{C})$, the space of $2n \times 2n$ skew-symmetric complex matrices, with compact real form $\mathsf{so}(2n)$, the space of $2n \times 2n$ skew-symmetric real matrices. We consider in $\mathsf{so}(2n)$ the maximal commutative subalgebra $\mathfrak{t}$ consisting of 2×2 block-diagonal matrices in which the k^{th} diagonal block is of the form

$$\begin{pmatrix} 0 & a_k \\ -a_k & 0 \end{pmatrix} \tag{8.8}$$

for some $a_k \in \mathbb{R}$. We then consider the Cartan subalgebra $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$ of $\mathsf{so}(2n;\mathbb{C})$, which consists of 2×2 block-diagonal matrices in which the k^{th} diagonal block is of the form (8.8) with $a_k \in \mathbb{C}$. (The calculations in the next two paragraphs show that $\mathsf{so}(2n;\mathbb{C})$ decomposes as a direct sum of $\mathfrak{h}$ and root spaces $\mathfrak{g}_\alpha$ corresponding to (nonzero) elements $\alpha \in \mathfrak{h}^*$. It follows from this that $\mathfrak{t}$ is actually a *maximal* commutative subalgebra of $\mathsf{so}(2n)$, which is not obvious from the definition of $\mathfrak{t}$. Similar remarks apply in the next two subsections.)

The root vectors are now 2×2 block matrices having a 2×2 matrix C in the (k,l) block $(k < l)$, the matrix $-C^{tr}$ in the (l,k) block, and zero in all other blocks, where C is one of the four matrices

$$C_1 = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix},$$
$$C_3 = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}, \quad C_4 = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

A little calculation shows that these are, indeed, root vectors and that the corresponding roots are the linear functionals on $\mathfrak{h}$ given by $i(a_k + a_l)$, $-i(a_k + a_l)$, $i(a_k - a_l)$, and $-i(a_k - a_l)$, respectively. (Compare Exercise 12 in Chapter 6.)

We may consider the inner product $\langle X, Y\rangle = \text{trace}(X^*Y)$ on $\mathsf{so}(2n;\mathbb{C})$, which is invariant under the adjoint action of $\mathsf{SO}(2n)$. If we use this inner product to identify $\mathfrak{h}^*$ with $\mathfrak{h}$, then the roots are thought of as elements of $\mathfrak{h}$ instead of $\mathfrak{h}^*$. Let Θ_k denote the 2×2 block-diagonal matrix whose k^{th} diagonal block is

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and whose other diagonal blocks are zero. The roots (as elements of $\mathfrak{h}$) are then the matrices

$$\frac{i}{2}\left(\pm\Theta_k \pm \Theta_l\right),$$

with $1 \leq k < l \leq n$. Each of the roots has length 1 with respect to the given inner product. The inner product of $(i/2)(\pm\Theta_k \pm \Theta_l)$ with $(i/2)(\pm\Theta_{k'} \pm \Theta_{l'})$ is zero if the set $\{k, l\}$ is disjoint from $\{k', l'\}$, and the inner product is $\pm 1/2$ if the intersection of $\{k, l\}$ and $\{k', l'\}$ has one element. The root $(i/2)(\Theta_k - \Theta_l)$ is orthogonal to the root $(i/2)(\Theta_k + \Theta_l)$.

As a base, we may take the $n-1$ roots

$$\frac{i}{2}(\Theta_1 - \Theta_2), \frac{i}{2}(\Theta_2 - \Theta_3), \cdots, \frac{i}{2}\left(\Theta_{n-2} - \Theta_{n-1}\right), \frac{i}{2}\left(\Theta_{n-1} - \Theta_n\right) \tag{8.9}$$

together with the one additional root,

$$\frac{i}{2}\left(\Theta_{n-1} + \Theta_n\right). \tag{8.10}$$

Note that for $1 \leq k < l \leq n$, we have the following formulas:

$$\begin{aligned}
\Theta_k - \Theta_l &= (\Theta_k - \Theta_{k+1}) + (\Theta_{k+1} - \Theta_{k+2}) + \cdots + (\Theta_{l-1} - \Theta_l), \\
\Theta_k + \Theta_n &= (\Theta_k - \Theta_{n-1}) + (\Theta_{n-1} + \Theta_n), \\
\Theta_k + \Theta_l &= (\Theta_k + \Theta_n) + (\Theta_l - \Theta_n).
\end{aligned}$$

This shows that every root of the form $(i/2)\left(\Theta_k - \Theta_l\right)$ or $(i/2)(\Theta_k + \Theta_l)$ $(k < l)$ can be written as a linear combination of the base in (8.9) and (8.10) with non-negative integer coefficients. The roots of this form are then positive and the remaining roots are negative.

Two consecutive roots in the list (8.9) have an angle of 120° and two nonconsecutive roots in the list (8.9) are orthogonal. The angle between the root in (8.10) and the *second-to-last* element in the list (8.9) is 120°; the root in (8.10) is orthogonal to all the other roots in (8.9).

8.8.2 The orthogonal algebras $\mathbf{so(2n+1;\mathbb{C})}$

The root system for $\mathsf{so}(2n+1;\mathbb{C})$ is denoted B_n. We consider $\mathsf{so}(2n+1;\mathbb{C})$, the space of $(2n+1) \times (2n+1)$ skew-symmetric complex matrices, with compact real form $\mathsf{so}(2n+1)$, the space of $(2n+1) \times (2n+1)$ skew-symmetric real

matrices. We consider in $\mathsf{so}(2n+1)$ the maximal commutative subalgebra $\mathfrak{t}$ consisting of block diagonal matrices with n blocks of size 2×2 followed by one block of size 1×1. We take the 2×2 blocks to be of the same form as in $\mathsf{so}(2n)$ and we take the 1×1 block to be zero. The associated Cartan subalgebra $\mathfrak{h}$ of $\mathsf{so}(2n+1;\mathbb{C})$ is then matrices of the same form as in $\mathfrak{t}$ except that the off-diagonal elements of the 2×2 blocks are permitted to be complex.

The Cartan subalgebra in $\mathsf{so}(2n+1;\mathbb{C})$ is identifiable in an obvious way with the Cartan subalgebra in $\mathsf{so}(2n;\mathbb{C})$. In particular, both $\mathsf{so}(2n;\mathbb{C})$ and $\mathsf{so}(2n+1;\mathbb{C})$ have rank n. With this identification of the Cartan subalgebras, every root for $\mathsf{so}(2n;\mathbb{C})$ is also a root for $\mathsf{so}(2n+1;\mathbb{C})$. There are $2n$ additional roots for $\mathsf{so}(2n+1;\mathbb{C})$. The root vectors for these additional roots are as follows. First, the matrices having

$$B_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

in entries $(2k, 2n+1)$ and $(2k+1, 2n+1)$ and having $-B_1^{tr}$ in entries $(2n+1, 2k)$ and $(2n+1, 2k+1)$. Second, the matrices having

$$B_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

in entries $(2k, 2n+1)$ and $(2k+1, 2n+1)$ and $-B_2^{tr}$ in entries $(2n+1, 2k)$ and $(2n+1, 2k+1)$. The corresponding roots, viewed as elements of $\mathfrak{h}^*$, are given by ia_k and $-ia_k$. (Compare Exercise 13 in Chapter 6.)

Let Θ_k have the same meaning as in the previous subsection, except that now Θ_k is a $(2n+1) \times (2n+1)$ matrix. We use the inner product $\langle X, Y \rangle = \text{trace}(X^*Y)$, which is invariant under the adjoint action of $\mathsf{SO}(2n+1)$, to identity $\mathfrak{h}^*$ with $\mathfrak{h}$. In that case, the additional roots for the $\mathsf{so}(2n+1;\mathbb{C})$ case are given by

$$\pm \frac{i}{2} \Theta_k.$$

These additional roots have length $1/\sqrt{2}$ with respect to the given inner product, whereas the roots that are the same as for $\mathsf{so}(2n;\mathbb{C})$ have length 1.

As a base for our root system, we may take the $n-1$ roots

$$\frac{i}{2}(\Theta_1 - \Theta_2), \frac{i}{2}(\Theta_2 - \Theta_3), \cdots, \frac{i}{2}(\Theta_{n-2} - \Theta_{n-1}), \frac{i}{2}(\Theta_{n-1} - \Theta_n) \tag{8.11}$$

(exactly as in the $\mathsf{so}(2n;\mathbb{C})$ case) together with the one additional root,

$$\frac{i}{2} \Theta_n. \tag{8.12}$$

The positive roots are those of the form $(i/2)(\Theta_k - \Theta_l)$ or $(i/2)(\Theta_k + \Theta_l)$ $(k < l)$ and those of the form $(i/2)\Theta_k$ $(1 \leq k \leq n)$. As in the $\mathsf{so}(2n;\mathbb{C})$ case, consecutive roots in the list (8.11) have an angle of $120°$, whereas nonconsecutive roots on the list (8.11) are orthogonal. Meanwhile, the root in (8.12) has an angle of $135°$ with the *last* root in (8.11) and is orthogonal to the remaining roots in (8.11).

8.8.3 The symplectic algebras $\mathsf{sp}(n;\mathbb{C})$

The root system for $\mathsf{sp}(n;\mathbb{C})$ is denoted C_n. We consider $\mathsf{sp}(n;\mathbb{C})$, the space of $2n \times 2n$ complex matrices X satisfying $JX^{tr}J = X$, where J is the $2n \times 2n$ matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Explicitly, the elements of $\mathsf{sp}(n;\mathbb{C})$ are matrices of the form

$$X = \begin{pmatrix} A & B \\ C & -A^{tr} \end{pmatrix},$$

where A is an arbitrary $n \times n$ matrix and B and C are arbitrary symmetric matrices. We consider the compact real form $\mathsf{sp}(n) = \mathsf{sp}(n;\mathbb{C}) \cap \mathsf{u}(2n)$.

We consider the maximal commutative subalgebra $\mathfrak{t}$ of $\mathsf{sp}(n)$ consisting of matrices of the form

$$\begin{pmatrix} ia_1 & & & & & & \\ & \ddots & & & & & \\ & & ia_n & & & & \\ & & & -ia_1 & & & \\ & & & & \ddots & & \\ & & & & & -ia_n \end{pmatrix},$$

with $a_1, \ldots, a_n \in \mathbb{R}$. We then consider the Cartan subalgebra $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$ of $\mathsf{sp}(n;\mathbb{C})$, which consists of matrices of the same form but with $a_1, \ldots, a_n \in \mathbb{C}$.

Let E_{kl} denote the $n \times n$ matrix whose (k,l) entry is one and whose other entries are zero. Then, the $2n \times 2n$ matrices of the block form

$$\begin{pmatrix} 0 & E_{kl} + E_{lk} \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ E_{kl} + E_{lk} & 0 \end{pmatrix} \tag{8.13}$$

$(k \neq l)$ are root vectors for which the corresponding roots are $i(a_k + a_l)$ and $-i(a_k + a_l)$. Next, matrices of the block form

$$\begin{pmatrix} E_{kl} + E_{lk} & 0 \\ 0 & 0 \end{pmatrix} \tag{8.14}$$

$(k \neq l)$ are root vectors for which the corresponding roots are $i(a_k - a_l)$. Finally, matrices of the block form

$$\begin{pmatrix} 0 & E_{kk} \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ E_{kk} & 0 \end{pmatrix} \tag{8.15}$$

are root vectors for which the corresponding roots are $2ia_k$ and $-2ia_k$.

We may use the inner product $\langle X, Y \rangle = \mathrm{trace}(X^*Y)$ on $\mathsf{sp}(n;\mathbb{C})$, which is invariant under the adjoint action of $\mathsf{Sp}(n) \subset \mathsf{U}(2n)$. If we use this inner

product to identify $\mathfrak{h}^*$ with $\mathfrak{h}$, then the roots become the following elements of $\mathfrak{h}$. The roots coming from (8.13) are

$$\frac{i}{2}\begin{pmatrix} E_{kk}+E_{ll} & 0 \\ 0 & -E_{kk}-E_{ll} \end{pmatrix}, \quad \frac{i}{2}\begin{pmatrix} -E_{kk}-E_{ll} & 0 \\ 0 & E_{kk}+E_{ll} \end{pmatrix}, \tag{8.16}$$

the roots coming from (8.14) are

$$\frac{i}{2}\begin{pmatrix} E_{kk}-E_{ll} & 0 \\ 0 & -E_{kk}+E_{ll} \end{pmatrix}, \tag{8.17}$$

and the roots coming from (8.15) are

$$i\begin{pmatrix} E_{kk} & 0 \\ 0 & -E_{kk} \end{pmatrix}, \quad i\begin{pmatrix} -E_{kk} & 0 \\ 0 & E_{kk} \end{pmatrix}. \tag{8.18}$$

The roots in (8.16) and (8.17) have length 1 and the roots in (8.18) have length $\sqrt{2}$.

As a base, we may take the $n-1$ roots

$$\frac{i}{2}\begin{pmatrix} E_{kk}-E_{k+1,k+1} & 0 \\ 0 & -E_{kk}+E_{k+1,k+1} \end{pmatrix} \tag{8.19}$$

together with the one additional root

$$i\begin{pmatrix} E_{nn} & 0 \\ 0 & -E_{nn} \end{pmatrix}. \tag{8.20}$$

The angle between two consecutive roots in (8.19) is 120°; nonconsecutive roots in (8.20) are orthogonal. The angle between the root in (8.20) and the last root in (8.19) is 135°; the root in (8.20) is orthogonal to the other roots in (8.19).

8.9 Dynkin Diagrams and the Classification

In this section, we discuss (without proof) the classification, up to equivalence, of root systems. This leads to a classification, up to equivalence, of semisimple Lie algebras. The classification of root systems is given in terms of an object called the Dynkin diagram.

Suppose $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ is a base for a root system R. Then, the Dynkin diagram for R (relative to the base Δ) is a graph having vertices $v_1, \ldots, v_r$. Between any two vertices, we place either no edge, one edge, two edges, or three edges as follows. Consider distinct indices i and j. If the corresponding roots α_i and α_j are orthogonal, then we put no edge between v_i and v_j. In the cases where α_i and α_j are not orthogonal, we put one edge between v_1 and v_j if α_i and α_j have the same length, two edges if the longer of α_i and

α_j is $\sqrt{2}$ longer than the shorter, and three edges if the longer of α_i and α_j is $\sqrt{3}$ longer than the shorter. In addition, if α_i and α_j are not orthogonal and not of the same length, then we decorate the edges between v_i and v_j with an arrow pointing from the vertex associated to the longer root toward the vertex associated to the shorter root. (Thinking of the arrow as a "greater than" sign makes it clear which way the arrow is supposed to go.) Proposition 8.6 tells us that if α_i and α_j are not orthogonal, then the only possible length ratios are 1, $\sqrt{2}$, and $\sqrt{3}$. Furthermore, Propositions 8.6 and 8.11 tell us that these three cases correspond to angles of 120°, 135°, and 150°, respectively.

Two Dynkin diagrams are said to be equivalent if there is a one-to-one, onto map of the vertices of one to the vertices of the other that preserves the number of bonds and the direction of the arrows. Recall (Theorem 8.20) that any two bases for the same root system can be mapped into one another by the action of the Weyl group. This implies that the equivalence class of the Dynkin diagram is independent of the choice of base. As we will see, not every graph arises as the Dynkin diagram of a root system, but only graphs of certain very special forms.

Theorem 8.26. *A root system is irreducible if and only if its Dynkin diagram is connected.*

Two root systems with equivalent Dynkin diagrams are equivalent.

If $R^\vee$ is the dual root system to R, then the Dynkin diagram of $R^\vee$ is the same as that of R except that the direction of each arrow is reversed.

So, the classification of irreducible root systems amounts to classifying all the connected diagrams that can arise as Dynkin diagrams of root systems.

The calculations in the previous section and in Section 6.9 allow us to read off the Dynkin diagrams for the classical Lie algebras, $\mathsf{sl}(n;\mathbb{C})$, $\mathsf{so}(n;\mathbb{C})$, and $\mathsf{sp}(n;\mathbb{C})$.

A_n. The root system A_n is the root system of $\mathsf{sl}(n+1;\mathbb{C})$, which has rank n. The Dynkin diagram for A_n is shown in Figure 8.11.

B_n. The root system B_n is the root system of $\mathsf{so}(2n+1;\mathbb{C})$, which has rank n. The Dynkin diagram for B_n is shown in Figure 8.12.

C_n. The root system C_n is is the root system of $\mathsf{sp}(n;\mathbb{C})$, which has rank n. The Dynkin diagram for C_n is shown in Figure 8.13.

D_n. The root system D_n is the root system of $\mathsf{so}(2n;\mathbb{C})$, which has rank n. The Dynkin diagram for D_n is shown in Figure 8.14.

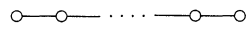


Fig. 8.11. The Dynkin diagram for A_n

Certain special things happen in low rank. In rank one, there is only one possible Dynkin diagram, reflecting that there is only one isomorphism class of complex semisimple Lie algebras in rank one. The Lie algebra $\mathsf{so}(2;\mathbb{C})$ is

Fig. 8.12. The Dynkin diagram for B_n

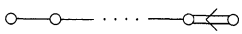


Fig. 8.13. The Dynkin diagram for C_n

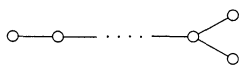


Fig. 8.14. The Dynkin diagram for D_n

not semisimple and the remaining three, $\mathsf{sl}(2;\mathbb{C})$, $\mathsf{so}(3;\mathbb{C})$, and $\mathsf{sp}(1;\mathbb{C})$, are isomorphic. In rank two, the Dynkin diagram D_2 is disconnected, reflecting that $\mathsf{so}(4;\mathbb{C}) \cong \mathsf{sl}(2;\mathbb{C}) \oplus \mathsf{sl}(2;\mathbb{C})$. Also, the Dynkin diagrams B_2 and C_2 are isomorphic, reflecting that $\mathsf{so}(5;\mathbb{C}) \cong \mathsf{sp}(2;\mathbb{C})$. In rank three, the Dynkin diagrams A_3 and D_3 are isomorphic, reflecting that $\mathsf{sl}(4;\mathbb{C}) \cong \mathsf{so}(6;\mathbb{C})$.

From the calculations in the previous section, we may observe certain things about the short and long roots in root systems where more than one length of root occurs. The long roots in B_n form a root system by themselves, namely D_n. The short roots in B_n form a root system by themselves, namely $A_1 \times \cdots \times A_1$. In C_n, it is the reverse: The long roots form $A_1 \times \cdots \times A_1$ and the short roots form D_n.

In addition to the root systems associated to the classical Lie algebras, there are five "exceptional" irreducible root systems, denoted G_2, F_4, E_6, E_7, and E_8, whose Dynkin diagrams are shown in Figure 8.15. We have constructed the root system G_2 explicitly; for constructions of the other exceptional root systems, see Section 12 of Humphreys (1972).

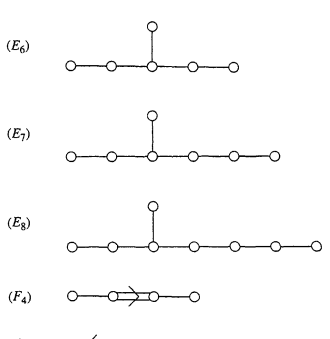


Fig. 8.15. The exceptional Dynkin diagrams

We are now ready to state (without proof) the classification theorem for irreducible root systems.

Theorem 8.27. *Every irreducible root system is isomorphic to precisely one root system from the following list:*

1. *The classical root systems* A_n, $n \geq 1$
2. *The classical root systems* B_n, $n \geq 2$
3. *The classical root systems* C_n, $n \geq 3$
4. *The classical root systems* D_n, $n \geq 4$
5. *The exceptional root systems* G_2, F_4, E_6, E_7, *and* E_8

The restrictions on the values of n are to avoid the low-rank repetitions discussed earlier. The classification of irreducible root systems leads to a classification of all root systems. Every root system can be decomposed as a direct sum of irreducible root systems, and the decomposition is unique. Thus, general root systems are classified by listing which irreducible summands occur and how many times each on occurs.

It turns out that the classification of semisimple Lie algebras is equivalent to the classification of root systems, as the following theorem explains.

Theorem 8.28.

1. *If* R_1 *and* R_2 *are the root systems for two different Cartan subalgebras of the same complex semisimple Lie algebra, then* R_1 *and* R_2 *are isomorphic.*
2. *A semisimple Lie algebra is simple if and only if its root system is irreducible.*
3. *If two complex semisimple Lie algebras have isomorphic root systems, then the semisimple Lie algebras are isomorphic.*
4. *Every root system arises as the root system of some complex semisimple Lie algebra.*

Point 4 of the theorem can be proved either by a general construction of semisimple Lie algebras, as in Section 18 of Humphreys (1972), or by a case-by-case analysis. It suffices to prove this for irreducible root systems and we already know the result for the root systems of type A, B, C, and D. So, it suffices to construct Lie algebras corresponding to the exceptional root systems, as, for example, in Jacobson (1962).

Theorems 8.27 and 8.28 lead to the following classification of complex simple Lie algebras.

Theorem 8.29. *Every complex simple Lie algebra is isomorphic to precisely one algebra from the following list:*

1. $\mathsf{sl}(n+1;\mathbb{C})$, $n \geq 1$
2. $\mathsf{so}(2n+1;\mathbb{C})$, $n \geq 2$
3. $\mathsf{sp}(n;\mathbb{C})$, $n \geq 3$

4. $\mathsf{so}(2n;\mathbb{C})$, $n \geq 4$
5. *The exceptional Lie algebras* G_2, F_4, E_6, E_7, *and* E_8

A semisimple Lie algebra is determined up to isomorphism by specifying which simple summands occur and how many times each on occurs.

8.10 The Root Lattice and the Weight Lattice

If E is a finite-dimensional real vector space, then a subset Λ of E is called a **lattice** if there is some basis $v_1, \ldots, v_r$ for E such that Λ is the space of all linear combinations of $v_1, \ldots, v_r$ with integer coefficients. The set of all integer linear combinations of roots is a lattice (called the **root lattice**) since the set of such linear combinations is the same as the set of linear combinations of the positive *simple* roots with integer coefficients, and the positive simple roots form a basis for E. The set of all integral elements is also a lattice, called the **weight lattice**, because it is the set of integer linear combinations of the fundamental weights. The weight lattice is precisely the set of elements that arise as weights of finite-dimensional representations of $\mathfrak{g}$.

Now, we have observed that every root is an integral element, and, therefore, a linear combination of roots with integer coefficients is also an integral element. This means that the root lattice is contained in the weight lattice. Are the two lattices equal? In general, no, not even for $\mathsf{sl}(2;\mathbb{C})$. In the $\mathsf{sl}(2;\mathbb{C})$, case we think of the weights as eigenvalues of the element H. These eigenvalues are integers, and every integer occurs as an eigenvalue of H in one of the finite-dimensional representations π_m of $\mathsf{sl}(2;\mathbb{C})$ described in Chapter 4. Thus, the weight lattice is isomorphic to $\mathbb{Z}$. Meanwhile, the eigenvalues of H in the adjoint representation are 0, 2, and -2. So, the roots correspond to the elements ± 2 inside $\mathbb{Z}$ and the root lattice corresponds to the set of *even* integers inside $\mathbb{Z}$.

For any root system, we may regard both the root lattice and the weight lattice as commutative subgroups of E under the operation of vector addition. The quotient group (weight lattice)/(root lattice) is then a finite commutative group. The following result is offered without proof.

Theorem 8.30. *Suppose K is a simply-connected compact Lie group with Lie algebra $\mathfrak{k}$. Let $\mathfrak{t}$ be a maximal commutative subalgebra of $\mathfrak{k}$ so that $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$ is a Cartan subalgebra of $\mathfrak{k}_{\mathbb{C}}$. Consider the root lattice and the weight lattice inside $\mathfrak{h}^*$. Then, the center of K satisfies*

$$Z(K) \cong (\textit{weight lattice})/(\textit{root lattice}).$$

It is essential in this result that K be simply connected. Note that the weight lattice and the root lattice are purely Lie-algebraic constructions and,

thus, their quotient must be canonically associated to the Lie algebra. Although there can be several different connected compact groups (with nonisomorphic centers) having Lie algebra K, there is (up to isomorphism) only one simply-connected group.

The way I have formulated Theorem 8.30 is "dual" to the usual formulation. For more information about this and for a result on the center of non-simply-connected compact groups, see Section E.4.

Instead of considering the simply-connected group K, we can consider the *adjoint group*, $\mathrm{Ad}(K) := K/Z(K)$. (It is called this since the kernel of the adjoint representation is the center of K, which means that the adjoint group is isomorphic to the image of K under the adjoint representation.) If K is compact and simply connected, then $\mathrm{Ad}(K)$ is the unique (up to isomorphism) Lie group whose center is trivial and whose Lie algebra is isomorphic to the Lie algebra of K. Then, we have

$$\pi_1(\mathrm{Ad}(K)) \cong (\text{weight lattice})/(\text{root lattice}),$$

where π_1 denotes the fundamental group (Appendix E). This explains why the quotient of the weight lattice by the root lattice is called the "fundamental group" of a root system in Humphreys (1972).

Another way to think about the relationship between the two lattices is as follows. The weight lattice is the set of possible weights of representations of the Lie algebra $\mathfrak{k}$ or, equivalently, of the simply-connected group K. The root lattice is the set of possible weights of representations of the adjoint group $\mathrm{Ad}(K)$. In particular, the roots themselves are the weights of the adjoint representation, which may be thought of as a representation of $\mathrm{Ad}(K)$.

We have already pointed out that in the case of $\mathsf{sl}(2;\mathbb{C})$, the root lattice inside the weight lattice may be thought of as the set of even integers inside the set of all integers. Thus, in this case, the quotient is isomorphic to $\mathbb{Z}/2$. This reflects that the center of the compact simply-connected group $\mathsf{SU}(2)$ is $\{I, -I\} \cong \mathbb{Z}/2$ and that the fundamental group of the adjoint group $\mathsf{SO}(3) \cong \mathsf{SU}(2)/\{I, -I\}$ is $\mathbb{Z}/2$.

Suppose that $\alpha_1, \ldots, \alpha_r$ form a base for a root system R in E, and suppose that $\mu_1, \ldots, \mu_r$ are the associated fundamental weights (having the property that $2\langle \mu_k, \alpha_l \rangle / \langle \alpha_l, \alpha_l \rangle = \delta_{kl}$, $k, l = 1, \ldots, r$). Then, $\mu_1, \ldots, \mu_r$ form a basis for E as a vector space. Thus, every element of E has a unique expansion in terms of $\mu_1, \ldots, \mu_r$ and the integral elements are precisely those elements of E for which the expansion coefficients are integers. Since each root is an integral element, we have, for each $k = 1, \ldots, r$, $\alpha_k = n_{k1}\mu_1 + \cdots + n_{kr}\mu_r$, with each n_{kl} being an integer. We can then form an $r \times r$ matrix whose entries are these integers. Then, the number of elements in (weight lattice)/(root lattice) is equal to the absolute value of the determinant of this matrix.

To see that this is so, observe that the set $a_1\mu_1 + \cdots + a_r\mu_r$, $0 \leq a_k < 1$, is a fundamental domain for the weight lattice; that is, every element of E can be written uniquely as the sum of an element of this set and an element of the weight lattice. Similarly, the set $a_1\alpha_1 + \cdots + a_r\alpha_r$ is a fundamental

domain for the root lattice. It is not hard to see that the number of elements in the quotient (weight lattice)/(root lattice) is the ratio of the volume of a fundamental domain in the root lattice to the volume of a fundamental domain in the weight lattice. This ratio of volumes is the absolute value of the determinant of the linear map that takes μ_k to α_k $(k = 1, \ldots, r)$. The matrix that represents this linear transformation in the basis $\mu_1, \ldots, \mu_r$ is the matrix whose entries are n_{kl}.

In the case of A_2 (the root system for $\mathfrak{sl}(3;\mathbb{C})$), the fundamental weights and positive simple roots are related by $\mu_1 = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$ and $\mu_2 = \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2$. Inverting this gives $\alpha_1 = 2\mu_1 - \mu_2$ and $\alpha_2 = -\mu_1 + 2\mu_2$, which can also be seen directly from Figure 5.2. Since

$$\det \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 3,$$

we conclude that (weight lattice)/(root lattice) has three elements in this case. Figure 8.16 shows the root lattice (large triangles) and the weight lattice (small triangles) for A_2. The large triangles in Figure 8.16 have three times the area of the small triangles, reflecting that the quotient of the two lattices has three elements. (In each lattice, a triangle is half of a fundamental domain.)

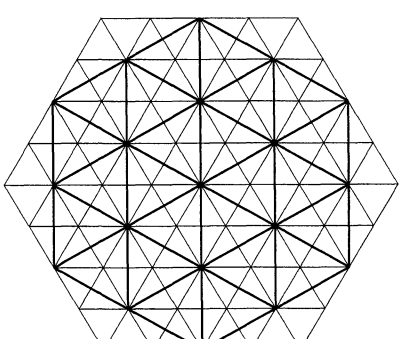


Fig. 8.16. The root lattice and the weight lattice for A_2

The reader is invited to perform the analogous calculation for the remaining rank-two root systems and verify the results of the following table.

root system	**# of elements**
$A_1 \times A_1$	4
A_2	3
B_2	2
G_2	1

Note that for G_2, each fundamental weight is a root, and thus in this case, the root lattice and weight lattice are equal. The same holds for the

exceptional root systems F_4 and E_8. For all other irreducible root systems, the root lattice is a proper sublattice of the weight lattice (see Section 3.6 of Chapter X of Helgason (1978)).

For the case of $\mathsf{sl}(n;\mathbb{C})$ (as in Section 6.9), the roots are the diagonal matrices having one diagonal entry equal to 1, one diagonal entry equal to -1, and the other diagonal entries equal to 0. The root lattice (integer linear combinations of the roots) is then the set of all diagonal matrices whose diagonal entries are integers that sum to zero.

Meanwhile, the co-roots are the same as the roots under our identifications. The integral elements are then the trace-zero diagonal matrices whose inner product with each (co-)root is an integer. This means that an integral element is a trace-zero diagonal matrix in which the *difference* of any two diagonal entries is an integer. It is not hard to show that such matrices are precisely those of the form

$$\mu = \mathrm{diag}\left(l_1 + \frac{k}{n}, l_2 + \frac{k}{n}, \cdots, l_n + \frac{k}{n}\right),$$

where k is an integer and $l_1, \ldots, l_n$ are integers satisfying $l_1 + \cdots + l_n = -k$. Here, $\mathrm{diag}(\cdot)$ denotes the diagonal matrix with the indicated diagonal entries. It is then not hard to show that the weight lattice modulo the root lattice is isomorphic to $\mathbb{Z}/n$, reflecting that

$$Z(\mathsf{SU}(n)) = \left\{ e^{2\pi i k/n} I \,\middle|\, k \in \mathbb{Z} \right\} \cong \mathbb{Z}/n.$$

8.11 Exercises

1. If R is a root system in E, consider the function $q : E \to \mathbb{R}$ given by

$$q(H) = \prod_{\alpha \in R^+} \langle \alpha, H \rangle .$$

   Using results from Section 8.7, show that q satisfies

$$q(w \cdot H) = \det(w) q(H)$$

   for all $w \in W$ and all $H \in E$.
2. Consider the diagrams of dominant integral elements for rank-two root systems in Section 8.5. In each case, verify that the roots labeled α_1 and α_2 form a base for the corresponding root system.
3. For each of the rank-two root systems, verify that the number of Weyl chambers is equal to the order of the Weyl group. (Compare Theorem 8.20.)
4. Verify directly that any two bases for B_2 can be mapped into one another by the action of the Weyl group.

5. For each of the irreducible rank-two root systems, verify that the Weyl group acts transitively on each length of root. (Compare Property 3 in Section 8.7.)
6. Consider the base $\{\alpha_1, \alpha_2\}$ for B_2 in Figure 8.6. Let μ_1 and μ_2 denote the fundamental weights for B_2, namely those satisfying

$$2\frac{\langle \alpha_k, \mu_l \rangle}{\langle \alpha_k, \alpha_k \rangle} = \delta_{kl},$$

for $k, l \in \{1, 2\}$. (The fundamental weights are circled in Figure 8.6.) Then, every dominant integral element μ for B_2 can be expressed in the form $\mu = m_1\mu_1 + m_2\mu_2$, with m_1 and m_2 being non-negative integers.
Using Theorem 7.43, compute the dimension of the irreducible representation with highest weight $m_1\mu_1 + m_2\mu_2$ explicitly as a function of m_1 and m_2.
7. Prove Property 10 in Section 8.7 using Property 5.
Hint: Suppose α is a positive simple root. Show that the positive roots β with $\beta \neq \alpha$ and $\langle \alpha, \beta \rangle \neq 0$ come in pairs $\{\beta_1, \beta_2\}$ with $\langle \alpha, \beta_1 \rangle = -\langle \alpha, \beta_2 \rangle$.
8. Suppose μ_1 and μ_2 are dominant integral elements and that $\mu_1 \succeq \mu_2$. Show, using the results of Section 8.7, that the convex hull of the W-orbit of μ_1 contains the convex hull of the W-orbit of μ_2.
9. Suppose that μ is an integral element. Using Exercise 8 and Theorem 7.41, show that there are infinitely many inequivalent irreducible representations of $\mathfrak{g}$ for which μ is a weight.
10. Compute the root lattice and the weight lattice for B_2 and verify that (weight lattice)/(root lattice) has two elements.
11. For which rank-two root systems is $-I$ an element of the Weyl group? (Compare Section 7.6.1.)
12. Verify the multiplicities of the irreducible representation of $\mathfrak{sl}(3;\mathbb{C})$ with highest weight $(1, 2)$ given in Section 5.7. Use Kostant's formula (Section 7.6) together with the invariance of the weights and multiplicities under the action of the Weyl group.

A

A Quick Introduction to Groups

For the reader who may not have had course in abstract group theory, I provide a quick review here. Only the definitions and basic properties of groups are needed in reading this book.

A.1 Definition of a Group and Basic Properties

Definition A.1. *A **group** is a set G, together with a map of $G \times G$ into G (denoted $g * h$) with the following properties:*

First, associativity: For all $g, h, k \in G$,

$$g * (h * k) = (g * h) * k. \tag{A.1}$$

Second, there exists an element e in G such that for all $g \in G$,

$$g * e = e * g = g \tag{A.2}$$

and such that for all $g \in G$, there exists $h \in G$ with

$$g * h = h * g = e. \tag{A.3}$$

*If $g * h = h * g$ for all $g, h \in G$, then the group is said to be **commutative** (or **abelian**).*

The element e is (as we shall see shortly) unique and is called the **identity element** of the group or, simply, the **identity**. Part of the definition of a group is that multiplying a group element g by the identity on *either the right or the left* must give back g.

The map of $G \times G$ into G is called the **product operation** for the group. Part of the definition of a group G is that the product operation map $G \times G$ into G (i.e., that the product of two elements of G be again an element of G). This may not always be obvious in examples. This property is referred to as **closure**.

Given a group element g, a group element h such that $g * h = h * g = e$ is called an **inverse** of g. We shall see momentarily that each group element has a *unique* inverse.

Given a set and an operation, there are four things that must be checked to show that this is a group: *closure*, *associativity*, existence of an *identity*, and existence of *inverses*.

Proposition A.2 (Uniqueness of the Identity). *Let G be a group and let $e, f \in G$ be such that for all $g \in G$,*

$$e * g = g * e = g,$$
$$f * g = g * f = g.$$

Then, $e = f$.

Proof. Since e is an identity, we have

$$e * f = f.$$

On the other hand, since f is an identity, we have

$$e * f = e.$$

Thus, $e = e * f = f$. □

Proposition A.3 (Uniqueness of Inverses). *Let G be a group, e the (unique) identity element of G, and g, h, k arbitrary elements of G. Suppose that*

$$g * h = h * g = e,$$
$$g * k = k * g = e.$$

Then, $h = k$.

Proof. We know that $g * h = g * k \ (= e)$. Multiplying on the left by h gives

$$h * (g * h) = h * (g * k).$$

By associativity, this gives

$$(h * g) * h = (h * g) * k,$$

and, so,

$$e * h = e * k,$$
$$h = k.$$

This is what we wanted to prove. □

Proposition A.4. *Let G be a group, e the identity element of G, and g an arbitrary element of G. Suppose $h \in G$ satisfies either $h * g = e$ or $g * h = e$. Then, h is the (unique) inverse of g and, thus, both $h * g = e$ and $g * h = e$ hold.*

Proof. To show that h is the inverse of g, we must show *both* that $h * g = e$ and $g * h = e$. Suppose we know, say, that $h * g = e$. Then, our goal is to show that this implies that $g * h = e$.

Since $h * g = e$,

$$g * (h * g) = g * e = g.$$

By associativity, we have

$$(g * h) * g = g.$$

Now, by the definition of a group, g has an inverse. Let k be that inverse. (Of course, in the end, we will conclude that $k = h$, but we cannot assume that now.) Multiplying on the right by k and using associativity again gives

$$\begin{aligned} ((g * h) * g) * k &= g * k = e, \\ (g * h) * (g * k) &= e, \\ (g * h) * e &= e, \\ g * h &= e. \end{aligned}$$

A similar argument shows that if $g * h = e$, then $h * g = e$. □

Note that in order to show that $h * g = e$ implies $g * h = e$, we used the fact that g has an inverse, since it is an element of a group. In more general contexts (i.e., in some systems that are not groups), one may have $h * g = e$ but not $g * h = e$. (See Exercise 10.)

Notation A.5 *For any group element g, its unique inverse will be denoted g^{-1}.*

Proposition A.6 (Properties of Inverses). *Let G be a group, e its identity, and g, h arbitrary elements of G. Then,*

$$\begin{aligned} \left(g^{-1}\right)^{-1} &= g, \\ (gh)^{-1} &= h^{-1}g^{-1}, \\ e^{-1} &= e. \end{aligned}$$

Proof. Exercise 3. □

A.2 Examples of Groups

From now on, we will denote the product of two group elements g and h simply by gh, instead of the more cumbersome $g * h$. Moreover, since we have associativity, we will omit parentheses and write ghk in place of $(gh)k$ or $g(hk)$.

A.2.1 The trivial group

The set with one element, e, is a group, with the group operation being defined as $ee = e$. This group is commutative.

Associativity is automatic, since both sides of (A.1) must be equal to e. Of course, e itself is the identity and is its own inverse. Commutativity is also automatic.

A.2.2 The integers

The set $\mathbb{Z}$ of integers forms a group with the product operation being addition. This group is commutative.

First, we check *closure*, namely that addition maps $\mathbb{Z} \times \mathbb{Z}$ into $\mathbb{Z}$ (i.e., that the sum of two integers is an integer). Since this is obvious, it remains only to check *associativity*, *identity*, and *inverses*. Addition is associative; zero is the additive identity (i.e., $0 + n = n + 0 = n$, for all $n \in \mathbb{Z}$); each integer n has an additive inverse, namely $-n$. Since addition is commutative, $\mathbb{Z}$ is a commutative group.

A.2.3 The reals and $\mathbb{R}^n$

The set $\mathbb{R}$ of real numbers also forms a group under the operation of addition. This group is commutative. Similarly, the n-dimensional Euclidean space $\mathbb{R}^n$ forms a group under the operation of vector addition. This group is also commutative.

The verification is the same as for the integers.

A.2.4 Nonzero real numbers under multiplication

The set of nonzero real numbers forms a group with respect to the operation of multiplication. This group is commutative.

Again, we check closure: The product of two nonzero real numbers is a nonzero real number. Multiplication is associative; the number 1 is the multiplicative identity; each nonzero real number x has a multiplicative inverse, namely $\frac{1}{x}$. Since multiplication of real numbers is commutative, this is a commutative group.

This group is denoted $\mathbb{R}^*$.

A.2.5 Nonzero complex numbers under multiplication

The set of nonzero complex numbers forms a group with respect to the operation of complex multiplication. This group is commutative and is denoted $\mathbb{C}^*$.

A.2.6 Complex numbers of absolute value 1 under multiplication

The set of complex numbers with absolute value 1 (i.e., of the form $e^{i\theta}$) forms a group under complex multiplication. This group is commutative.

This group is the unit circle, denoted S^1.

A.2.7 The general linear groups

For each positive integer n, the set of all $n \times n$ invertible matrices with real entries forms a group with respect to the operation of matrix multiplication.

We check closure: The product of two invertible matrices A and B is invertible, since $(AB)^{-1} = B^{-1}A^{-1}$. Matrix multiplication is associative; the identity matrix (with ones on the diagonal and zeros elsewhere) is the identity element; by definition, an invertible matrix has an inverse. Simple examples show that the group is noncommutative, except in the trivial case $n = 1$. (See Exercise 8.)

This group is called the **general linear group** (over the reals) and is denoted $\mathsf{GL}(n; \mathbb{R})$.

In the same way, we define the general linear group over the complex numbers, denoted $\mathsf{GL}(n; \mathbb{C})$.

A.2.8 Symmetric group (permutation group)

The set of one-to-one, onto maps of the set $\{1, 2, \ldots, n\}$ to itself forms a group under the operation of composition.

We check closure: The composition of two one-to-one, onto maps is again one-to-one and onto. Composition of functions is associative; the identity map (which sends 1 to 1, 2 to 2, etc.) is the identity element; a one-to-one, onto map has an inverse. Simple examples show that the group is noncommutative for $n \geq 3$. (See Exercise 9.)

This group is called the **symmetric group** and is denoted S_n. A one-to-one, onto map of $\{1, 2, \ldots, n\}$ is a permutation, and, so, S_n is also called the **permutation group**. The group S_n has $n!$ elements.

A.2.9 Integers mod n

The set $\{0, 1, \ldots, n-1\}$ forms a group under the operation of addition modulo n, where n is a positive. This group is commutative.

Explicitly, the group operation is the following. Consider a, b in the set $\{0, 1, \ldots, n-1\}$. If $a + b < n$, then we define $a + b$ **mod** n to be $a + b$. If $a + b \geq n$, then we define $a + b$ **mod** n to be $a + b - n$. (Since a and b are less than n, $a + b - n$ is less than n; thus, we have closure.) To show associativity, note that both

$$(a + b \textbf{ mod } n) + c \textbf{ mod } n$$

and

$$a + (b + c \bmod n) \bmod n$$

are equal to $a+b+c$, minus some multiple of n, and hence differ by a multiple of n. However, since both are in the set $\{0, 1, \ldots, n-1\}$, the only possible multiple on n is zero. Zero is still the identity for addition modulo n. The inverse of an element a in $\{0, 1, \ldots, n-1\}$ is $n-a$. The group is commutative because ordinary addition is commutative.

This group is referred to as "$\mathbb{Z}$ **mod** n" and is denoted $\mathbb{Z}/n$.

A.3 Subgroups, the Center, and Direct Products

Definition A.7. *A **subgroup** of a group G is a subset H of G with the following properties:*

1. *The identity is an element of H.*
2. *If $h \in H$, then $h^{-1} \in H$.*
3. *If $h_1, h_2 \in H$, then $h_1 h_2 \in H$.*

The conditions on H guarantee that H is a group, with the same product operation as G (but restricted to H). Closure is assured by Condition 3, associativity follows from associativity in G, and the existence of an identity and of inverses is assured by Conditions 1 and 2 (together with the existence of an identity and inverses in G). If H is a nonempty subset of G, then Conditions 2 and 3 imply Condition 1.

Every group G has at least two subgroups: G itself and the one-element subgroup $\{e\}$. (If G itself is the trivial group, then these two subgroups coincide.) These are called the **trivial subgroups** of G.

The set of even integers is a subgroup of $\mathbb{Z}$: Zero is even, the negative of an even integer is even, and the sum of two even integers is even.

The set H of $n \times n$ real matrices with determinant one is a subgroup of $\mathsf{GL}(n;\mathbb{R})$. The set H is a sub*set* of $\mathsf{GL}(n;\mathbb{R})$ because any matrix with determinant one is invertible. The identity matrix has determinant one, so Condition 1 is satisfied. The determinant of the inverse is the reciprocal of the determinant, so Condition 2 is satisfied; and the determinant of a product is the product of the determinants, so Condition 3 is satisfied. This group is called the **special linear group** (over the reals) and is denoted $\mathsf{SL}(n;\mathbb{R})$.

Additional examples, as well as some nonexamples, are given in Exercise 2.

Definition A.8. *The **center** of a group G is the set of all $g \in G$ such that $gh = hg$ for all $h \in G$.*

It is not hard to see that the center of any group G is a subgroup G.

Definition A.9. *Let G and H be groups and consider the Cartesian product of G and H (i.e., the set of ordered pairs (g, h) with $g \in G, h \in H$). Define a product operation on this set as follows:*

$$(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2).$$

This operation makes the Cartesian product of G and H into a group, called the ***direct product*** *of G and H and denoted $G \times H$.*

It is a simple matter to check that this operation truly makes $G \times H$ into a group. For example, the identity element of $G \times H$ is the pair (e_1, e_2), where e_1 is the identity for G and e_2 is the identity for H.

A.4 Homomorphisms and Isomorphisms

Definition A.10. *Let G and H be groups. A map $\Phi : G \to H$ is called a* ***homomorphism*** *if $\Phi(gh) = \Phi(g)\Phi(h)$ for all $g, h \in G$. If, in addition, Φ is one-to-one and onto, then Φ is called an* ***isomorphism.*** *An isomorphism of a group with itself is called an* ***automorphism.***

Proposition A.11. *Let G and H be groups, e_1 the identity element of G, and e_2 the identity element of H. If $\Phi : G \to H$ is a homomorphism, then $\Phi(e_1) = e_2$ and $\Phi(g^{-1}) = \Phi(g)^{-1}$ for all $g \in G$.*

Proof. Let g be any element of G. Then, $\Phi(g) = \Phi(ge_1) = \Phi(g)\Phi(e_1)$. Multiplying on the left by $\Phi(g)^{-1}$ gives $e_2 = \Phi(e_1)$. Now, consider $\Phi(g^{-1})$. Since $\Phi(e_1) = e_2$, we have $e_2 = \Phi(e_1) = \Phi(gg^{-1}) = \Phi(g)\Phi(g^{-1})$. From Proposition A.4, we conclude that $\Phi(g^{-1})$ is the inverse of $\Phi(g)$. □

Definition A.12. *Let G and H be groups, $\Phi : G \to H$ a homomorphism, and e_2 the identity element of H. The* ***kernel*** *of Φ is the set of all $g \in G$ for which $\Phi(g) = e_2$.*

Proposition A.13. *Let G and H be groups and $\Phi : G \to H$ a homomorphism. Then, the kernel of Φ is a subgroup of G.*

Proof. Easy. □

Actually, the kernel Φ will be a *normal* subgroup of G. See Section A.5.

Given any two groups G and H, we have the trivial homomorphism from G to H: $\Phi(g) = e$ for all $g \in G$. The kernel of this homomorphism is all of G.

In any group G, the identity map (which maps every element g to itself) is an automorphism of G, whose kernel is just $\{e\}$.

Let $G = H = \mathbb{Z}$ and define $\Phi(n) = 2n$. This is a homomorphism of $\mathbb{Z}$ to itself, but not an automorphism. The kernel of this homomorphism is just $\{0\}$.

The determinant is a homomorphism of $\mathsf{GL}(n, \mathbb{R})$ to $\mathbb{R}^*$. The kernel of this map is $\mathsf{SL}(n, \mathbb{R})$.

Additional examples are given in Exercises 7 and 12.

If there exists an isomorphism from G to H, then G and H are said to be **isomorphic**, and this relationship is denoted $G \cong H$. (See Exercise 4.) Two groups which are isomorphic should be thought of as being (for all practical purposes) the same group.

A.5 Quotient Groups

Definition A.14. *Let G be a group and N a subgroup of G. Then, N is called a **normal subgroup** of G if for each element g of G and each element n of N, the element gng^{-1} belongs to N.*

Note that if G is commutative, then every subgroup of G is automatically normal, since, in the commutative case, $gng^{-1} = n$.

The motivation for this definition comes from the following result.

Proposition A.15. *If G and H are groups and $\Phi : G \to H$ is a homomorphism, then $\ker \Phi$ is a normal subgroup of G.*

Proof. Let e_2 denote the identity element of H. Suppose that g is an element of G and n is an element of $\ker \Phi$ (i.e., that $\Phi(n) = e_2$). Then, we compute that

$$\begin{aligned} \Phi(gng^{-1}) &= \Phi(g)\Phi(n)\Phi(g^{-1}) \\ &= \Phi(g)e_2\Phi(g)^{-1} \\ &= e_2. \end{aligned}$$

This shows that gng^{-1} is, again, an element of $\ker \Phi$ and, thus, that $\ker \Phi$ is a normal subgroup of G. □

Let G be a group and N a subgroup of G (for the moment, not assumed normal). Then, define an equivalence relation on G by defining two elements g and h to be equivalent if gh^{-1} is an element of N. Let us see that this is indeed an equivalence relation (i.e., that this notion of equivalence is reflexive, symmetric, and transitive). For any g, $gg^{-1} = e$ and e is an element of N. This shows that every element of G is equivalent to itself. If gh^{-1} is an element of N, then $(gh^{-1})^{-1} = hg^{-1}$ is also an element of N. This shows that if g is equivalent to h, then h is also equivalent to g. Finally, if gh^{-1} is an element of N and hk^{-1} is an element of N, then $gh^{-1}hk^{-1} = gk^{-1}$ is an element of N. This shows that if g is equivalent to h and h is equivalent to k, then g is equivalent to k.

Proposition A.16. *Suppose that G is a group and N is a normal subgroup of G. Define two elements g and h of G to be equivalent if $gh^{-1} \in N$.*

1. *If g_1 is equivalent to g_2 and h_1 is equivalent to h_2, then g_1h_1 is equivalent to g_2h_2.*
2. *If g_1 is equivalent to g_2, then g_1^{-1} is equivalent to g_2^{-1}.*

In this proposition, it is essential that N be a *normal* subgroup of G. This proposition says that the equivalence relation "respects" the group operations of multiplication and inversion.

Proof. Assume that g_1 is equivalent to g_2 and that h_1 is equivalent to h_2. We want to show that g_1h_1 is equivalent to g_2h_2. We first show that g_1h_1 is equivalent to g_1h_2 by computing

$$g_1h_1(g_1h_2)^{-1} = g_1h_1h_2^{-1}g_1^{-1}.$$

Now, we note that $h_1h_2^{-1}$ is an element of N, since h_1 and h_2 are assumed equivalent. Then, *because* N is normal, we have that $g_1(h_1h_2^{-1})g_1^{-1}$ is also an element of N and, therefore, g_1h_1 is equivalent to g_1h_2. Next, we show that g_1h_2 is equivalent to g_2h_2 by computing that

$$g_1h_2(g_2h_2)^{-1} = g_1h_2h_2^{-1}g_2^{-1} = g_1g_2^{-1} \in N.$$

So, g_1h_1 is equivalent to g_1h_2, which is equivalent to g_2h_2. This implies (as we have shown earlier) that g_1h_1 is equivalent to g_2h_2.

Meanwhile, if g_1 is equivalent to g_2, then

$$g_1^{-1}(g_2^{-1})^{-1} = g_1^{-1}g_2 = g_2^{-1}g_2g_1^{-1}g_2.$$

However, since g_2 is equivalent to g_1, $g_2g_1^{-1}$ is an element of N, and then because N is normal, $g_2^{-1}(g_2g_1^{-1})g_2$ is also an element of N. □

Now, if g is any element of G, let $[g]$ denote the equivalence class containing G; that is, $[g]$ is the subset of G consisting of all elements equivalent to g (including g itself). Since our equivalence relation is an equivalence (reflexive, symmetric, and transitive), if g_1 is equivalent to g_2, then the equivalence class $[g_1]$ is the same as the equivalence class $[g_2]$. Every element of G belongs to precisely one equivalence class.

Definition A.17. *Let G be a group and N a normal subgroup of G. The* ***quotient group*** *G/N is the set of all equivalence classes in G, with the product operation defined by*

$$[g][h] = [gh].$$

Let us see if we understand what this means. The elements of G/N are equivalence classes, and the group product is defined by choosing one element g out of the first equivalence class, choosing one element h out of the second equivalence class, and then defining the product to be the equivalence class containing gh. We need to check that the group product is *well defined* (i.e.,

that the product does not depend on the choice of elements out of each equivalence class). However, this is precisely where the normality of N comes in. If we pick a different element g' from the first equivalence class and a different element h' out of the second, then by the proposition, $g'h'$ will be equivalent to gh. This means that the equivalence class $[gh]$ is the same as the equivalence class $[g'h']$, which shows that our product really is well defined.

The idea behind the quotient group construction is that we make a new group out of G by setting every element of N equal to the identity. This then forces ng to be equal to g for any $g \in G$ and $n \in N$. However, ng and n are equivalent, since $(ng)g^{-1} = n \in N$. So, setting elements of N equal to the identity forces elements that are equivalent (in the above sense) to be equal. The condition that N be a normal subgroup guarantees that we still have defined group operations after setting equivalent elements equal to each other.

The simplest example of a quotient group is the group of integers modulo n. In this case, we take $G = \mathbb{Z}$ and $N = n\mathbb{Z}$ (the set of integer multiples of n). It is easy to check that N is a subgroup of $\mathbb{Z}$, and since $\mathbb{Z}$ is commutative, all subgroups are normal. To form the quotient group, we say that two elements of $\mathbb{Z}$ are equivalent if their difference is in N. (We use additive notation for the group operation in $\mathbb{Z}$ and, so, the quotient gh^{-1} becomes $i - j$.) Thus, the equivalence class of an integer i is the set of all integers that are equal modulo n to i. The operation of addition makes the set of equivalence classes into a group and this group is nothing but the group of integers modulo n, as described in Section A.2. (In Section A.2, we label equivalence classes modulo n by picking the unique element of the equivalence class that is between 0 and $n - 1$.)

Another example is obtained by taking $G = \mathsf{SL}(n;\mathbb{C})$ and taking N to be the set of elements of $\mathsf{SL}(n;\mathbb{C})$ that are multiples of the identity. The elements of N are the matrices of the form $e^{2\pi ik/n}I$, $k = 0, 1, \ldots, n-1$. This is a normal subgroup of $\mathsf{SL}(n;\mathbb{C})$ because each element of N is a multiple of the identity, and, thus, for any $A \in \mathsf{SL}(n;\mathbb{C})$, we have $A(e^{2\pi ik/n}I)A^{-1} = AA^{-1}(e^{2\pi ik/n}I) = e^{2\pi ik/n}I$. The quotient group $\mathsf{SL}(n;\mathbb{C})/N$ is customarily denoted $\mathsf{PSL}(n;\mathbb{C})$, where the P stands for "projective." It can be shown that $\mathsf{PSL}(n;\mathbb{C})$ is a simple group for all $n \geq 2$; that is, $\mathsf{PSL}(n;\mathbb{C})$ has no normal subgroups other than $\{I\}$ and $\mathsf{PSL}(n;\mathbb{C})$ itself.

If G is a group and N a normal subgroup, then there is a homomorphism q of G into the quotient group G/N given by

$$q(g) = [g].$$

It follows from the definition of the product operation on G/N that q is indeed a homomorphism and, clearly, q maps G onto G/N. More generally, suppose that G and H are groups and that $\Phi : G \to H$ is a homomorphism. We have observed that the kernel of Φ is a normal subgroup of G. If Φ maps G *onto* H, then it can be shown that H is isomorphic to the quotient group $G/\ker\Phi$.

A.6 Exercises

Recall the definitions of the groups $\mathsf{GL}(n;\mathbb{R})$, S_n, $\mathbb{R}^*$, and $\mathbb{Z}/n$ from Section A.2, and the definition of the group $\mathsf{SL}(n;\mathbb{R})$ from Section A.3.

1. Show that the center of any group G is a subgroup G.
2. In (a)-(f), you are given a group G and a subset H of G. In each case, determine whether H is a subgroup of G.
   (a) $G = \mathbb{Z}$, $H = \{\text{odd integers}\}$
   (b) $G = \mathbb{Z}$, $H = \{\text{multiples of } 3\}$
   (c) $G = \mathsf{GL}(n;\mathbb{R})$, $H = \{A \in \mathsf{GL}(n;\mathbb{R}) \,|\det A \text{ is an integer}\}$
   (d) $G = \mathsf{SL}(n;\mathbb{R})$, $H = \{A \in \mathsf{SL}(n;\mathbb{R}) \,|\text{all entries of } A \text{ are integers}\}$
   Hint: Recall the formula for A^{-1} in terms of cofactors of A.
   (e) $G = \mathsf{GL}(n;\mathbb{R})$, $H = \{A \in \mathsf{GL}(n;\mathbb{R}) \,|\text{all entries of } A \text{ are rational}\}$
   (f) $G = \mathbb{Z}/9$, $H = \{0, 2, 4, 6, 8\}$
3. Verify the properties of inverses in Proposition A.6.
4. Let G and H be groups. Suppose there exists an isomorphism ϕ from G to H. Show that there exists an isomorphism from H to G.
5. Show that the set of positive real numbers is a subgroup of $\mathbb{R}^*$. Show that this group is isomorphic to the group $\mathbb{R}$.
6. Show that the set of automorphisms of any group G is itself a group, under the operation of composition. This group is the **automorphism group** of G, $\mathrm{Aut}(G)$.
7. Given any group G and any element g in G, define $\phi_g : G \to G$ by $\phi_g(h) = ghg^{-1}$. Show that ϕ_g is an automorphism of G. Show that the map $g \to \phi_g$ is a homomorphism of G into $\mathrm{Aut}(G)$ and that the kernel of this map is the center of G.
   Note: An automorphism which can be expressed as ϕ_g for some $g \in G$ is called an **inner automorphism**; any automorphism of G which is not equal to any ϕ_g is called an **outer automorphism**.
8. Give an example of two 2×2 invertible real matrices which do not commute. (This shows that $\mathsf{GL}(2,\mathbb{R})$ is not commutative.)
9. An element σ of the permutation group S_n can be written in a two-row form:
$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma_1 & \sigma_2 & \cdots & \sigma_n \end{pmatrix},$$
   where σ_i denotes $\sigma(i)$. Thus,
$$\sigma = \begin{pmatrix} 1\,2\,3 \\ 2\,3\,1 \end{pmatrix}$$
   is the element of S_3 which sends 1 to 2, 2 to 3, and 3 to 1. When multiplying (i.e., composing) two permutations, one performs the one on the right first and then the one on the left. (This is the usual convention for composing functions.)
   Compute

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

Conclude that S_3 is not commutative.

10. Consider the set $\mathbb{N} = \{0, 1, 2, \ldots\}$ of natural numbers and the set $\mathcal{F}$ of *all* functions of $\mathbb{N}$ to itself. Composition of functions defines a map of $\mathcal{F} \times \mathcal{F}$ into $\mathcal{F}$, which is associative. The identity $(\mathrm{id}(n) = n)$ has the property that $\mathrm{id} \circ f = f \circ \mathrm{id} = f$, for all f in $\mathcal{F}$. However, since we do not restrict to functions which are one-to-one and onto, not every element of $\mathcal{F}$ has an inverse. Thus, $\mathcal{F}$ is not a group.
Give an example of two functions f and g in $\mathcal{F}$ such that $f \circ g = \mathrm{id}$, but $g \circ f \neq \mathrm{id}$. (Compare with Proposition A.4.)
11. Consider the groups $\mathbb{Z}$ and $\mathbb{Z}/n$. For each a in $\mathbb{Z}$, define a **mod** n to be the unique element b of $\{0, 1, \ldots, n-1\}$ such that a can be written as $a = kn + b$, with k an integer. Show that the map $a \to a$ **mod** n is a homomorphism of $\mathbb{Z}$ into $\mathbb{Z}/n$.
12. Show that the center of any group G is a *normal* subgroup of G.

B

Linear Algebra Review

In this appendix, I collect together results from linear algebra that are used in the text. Only the simplest proofs are given here. The results quoted here are mostly standard, except for the SN decomposition, which is often skipped over on the way to the Jordan canonical form, and the discussion of weight spaces. For more information, the reader is encouraged to consult such standard linear algebra textbooks as Hoffman and Kunze (1971) or Axler (1997).

B.1 Eigenvectors, Eigenvalues, and the Characteristic Polynomial

If A is any matrix in $M_n(\mathbb{C})$, then a nonzero vector v in $\mathbb{C}^n$ is called an **eigenvector** for A if there is some complex number λ such that

$$Av = \lambda v.$$

An **eigenvalue** for A is a complex number λ for which there exists a nonzero $v \in \mathbb{C}^n$ with $Av = \lambda v$. So, λ is an eigenvalue for A if the equation $Av = \lambda v$ or, equivalently, the equation

$$(A - \lambda I)v = 0,$$

has a nonzero solution v. This happens precisely when $A - \lambda I$ fails to be invertible, which is precisely when $\det(A - \lambda I) = 0$.

For any $A \in M_n(\mathbb{C})$, we define the **characteristic polynomial** p of A to be given by

$$p(\lambda) = \det(A - \lambda I), \quad \lambda \in \mathbb{C}.$$

This is a polynomial of degree n. In light of the above discussion, the eigenvalues are precisely the zeros of the characteristic polynomial.

More generally, we may consider **vector spaces**. A vector space is a set V together with two operations: one called vector addition that takes two

elements of V and produces another element of V, and one called scalar multiplication that takes a complex number and an element of V and produces another element of V. To be a vector space, these two operations should have the same algebraic properties as vector addition and scalar multiplication in $\mathbb{C}^n$ (e.g., that vector addition is commutative and associative). There is a notion of the dimension of a vector space V, which may be infinite.

A **linear operator** on a vector space V (also called a linear transformation) is a map A of V to itself that satisfies

$$\begin{aligned} A(u+v) &= Au + Av, \\ A(\lambda u) &= \lambda Au \end{aligned}$$

for all u and v in V and all λ in $\mathbb{C}$. If A is an $n \times n$ matrix, then the map that sends a vector u in $\mathbb{C}^n$ to the vector Au (defined as the matrix product of the $n \times n$ matrix A and the $n \times 1$ matrix u) is a linear operator, and every linear operator on $\mathbb{C}^n$ arises in this way for some unique matrix A. So, we will interchangeably regard A as an $n \times n$ matrix or as a linear transformation of $\mathbb{C}^n$.

For any linear operator, we may define the notion of eigenvector and eigenvalue in precisely the same way as for $\mathbb{C}^n$. If A is a linear operator on a finite-dimensional space, then the theory of eigenvectors and eigenvalues for A is the same as for matrices. If A is a linear operator on an infinite-dimensional space, then A may not have any eigenvectors.

If A is a linear operator on a vector space V and λ is an eigenvalue for A, then the **λ-eigenspace** for A, denoted V_λ, is the set of all vectors $v \in V$ (including the zero vector) that satisfy $Av = \lambda v$. The λ-eigenspace for A is a subspace of V. The dimension of this space is called the **multiplicity** of λ. (More precisely, this is the "geometric multiplicity" of λ. In the finite-dimensional case, there is also a notion of the "algebraic multiplicity" of λ, which is the number of times that λ occurs as a root of the characteristic polynomial. The geometric multiplicity of λ cannot exceed the algebraic multiplicity.)

Proposition B.1. *Suppose that A is a linear operator on a vector space V and $v_1, \ldots, v_k$ are eigenvectors with **distinct** eigenvalues $\lambda_1, \ldots, \lambda_k$. Then, $v_1, \ldots, v_k$ are linearly independent.*

Note that here V does not have to be finite dimensional.

Proposition B.2. *Every linear operator A on a finite-dimensional complex vector space has at least one eigenvector.*

This follows from the fundamental theorem of algebra, which says that every nonconstant polynomial with complex entries has at least one complex root.

B.2 Diagonalization

Two matrices $A, B \in M_n(\mathbb{C})$ are said to be **similar** if there exists an invertible matrix C such that

$$A = CBC^{-1},$$

in which case, $B = C^{-1}AC$. The operation $B \to CBC^{-1}$ is called **conjugation** of B by C. A matrix is said to be **diagonalizable** if it is similar to a diagonal matrix. A matrix $A \in M_n(\mathbb{C})$ is diagonalizable if and only if there exist n linearly independent eigenvectors for A. Specifically, if $v_1, \ldots, v_n$ are linearly independent eigenvectors, let C be the matrix whose k^{th} column is v_k. Then, C is invertible and we will have

$$A = C \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} C^{-1},$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues associated to the eigenvectors $v_1, \ldots, v_n$, in that order.

If $A \in M_n(\mathbb{C})$ has n distinct eigenvalues (i.e., n distinct roots to the characteristic polynomial), then A is automatically diagonalizable, by Proposition B.1. If the characteristic polynomial of A has repeated roots, then A may or may not be diagonalizable.

Recall that for $A \in M_n(\mathbb{C})$, the **adjoint** of A, denoted A^*, is the conjugate-transpose of A,

$$(A^*)_{kl} = \overline{A_{lk}}.$$

A matrix A is said to be **self-adjoint** (or **Hermitian**) if $A^* = A$. A matrix A is said to be **skew self-adjoint** (or **skew Hermitian** or just **skew**) if $A^* = -A$. A matrix is said to be **unitary** if $A^* = A^{-1}$. If A is self-adjoint, skew self-adjoint, or unitary, then A is automatically diagonalizable. Furthermore, in these cases, it is possible to find an orthonormal basis of eigenvectors for A, which means that the matrix C in the definition of diagonalizability may be taken to be unitary.

If A is self-adjoint, then all of its eigenvalues are real. If A is real and self-adjoint (or, equivalently, real and symmetric), then the eigenvectors may be taken to be real as well, which means that in this case, the matrix C may be taken to be orthogonal. If A is skew, then its eigenvalues are imaginary. If A is unitary, then its eigenvalues are complex numbers of absolute value 1 (i.e., of the form $\lambda = e^{i\theta}$, with $\theta \in \mathbb{R}$).

A matrix A is said to be **normal** if A commutes with its adjoint (i.e., if $AA^* = A^*A$). If A is self-adjoint, skew, or unitary, then it is normal (since in those cases, A^* is A or $-A$ or A^{-1}, all of which commute with A). A normal matrix is automatically diagonalizable and has an orthonormal basis of eigenvectors. We summarize the results of the previous paragraphs in the following.

Theorem B.3. *Suppose that $A \in M_n(\mathbb{C})$ has the property that $A^*A = AA^*$, (e.g., if $A^* = A$, $A^* = A^{-1}$, or $A^* = -A$). Then, A is diagonalizable and it is possible to find an orthonormal basis for $\mathbb{C}^n$ consisting of eigenvectors for A.*

If A is real and symmetric, then all of the eigenvalues of A are real and it is possible to choose an orthonormal basis of eigenvectors for A in which each eigenvector is real.

Proposition B.4. *Suppose that A and B are linear operators on a finite-dimensional vector space V and suppose that $AB = BA$. Then, B maps the λ-eigenspace of A into itself, for each eigenvalue λ of A.*

Proof. Let λ be an eigenvalue of A and let V_λ be the λ-eigenspace of A. Then, let v be an element of V_λ and consider Bv. Since B commutes with A, we have

$$A(Bv) = BAv = \lambda Bv;$$

that is, applying A to Bv gives us back λ times the vector we started with, and, so, Bv is, again, an element of V_λ. □

B.3 Generalized Eigenvectors and the SN Decomposition

Not all matrices are diagonalizable. For example, consider the matrix

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

The characteristic polynomial of this is $p(\lambda) = (\lambda - 1)^2$, so the only eigenvalue of A is $\lambda = 1$. Solving the equation $Av = v$ gives

$$v = c \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

where c is an arbitrary constant. This means that we cannot find two linearly independent eigenvectors for A.

If A does not have enough linearly independent eigenvectors to be diagonalizable, then we may consider the more general concept of generalized eigenvectors. A nonzero vector $v \in \mathbb{C}^n$ is called a **generalized eigenvector** for A if there is some complex number λ and some positive integer k such that

$$(A - \lambda I)^k v = 0.$$

This can happen only if $(A - \lambda I)$ is noninvertible. This means that the number λ must be an (ordinary) eigenvalue for A. However, given an eigenvalue λ, there may be generalized eigenvectors v that are not ordinary eigenvectors (in addition to at least one ordinary eigenvector). For example, if A is the 2×2 matrix given earlier, then we may check that

$$(A - 1I)^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

so that $(0, 1)$ is a generalized eigenvector for A (but not an ordinary eigenvector). This is in addition to $(1, 0)$, which is an ordinary eigenvector.

Given any $A \in M_n(\mathbb{C})$, it is possible to find a basis $v_1, \ldots, v_n$ for $\mathbb{C}^n$ such that each v_k is a generalized eigenvector for A. This can be proved fairly easily by induction on n. Given a matrix A and an eigenvalue λ, let V_λ be defined by

$$V_\lambda = \left\{ v \in \mathbb{C}^n \,\middle|\, (A - \lambda I)^k v = 0 \text{ for some } k \right\};$$

that is, V_λ is the space of all generalized eigenvectors with eigenvalue λ, together with the zero vector, which, by definition, is not a generalized eigenvector but which satisfies $(A - \lambda I)^k v = 0$. For any λ, V_λ is a subspace of $\mathbb{C}^n$; that is, any linear combination of elements of V_λ is, again, in V_λ. It can be shown that $\mathbb{C}^n$ decomposes as a direct sum of the V_λ's, as λ ranges over all the eigenvalues of A. This means that if $\lambda_1, \ldots, \lambda_k$ denote the distinct eigenvalues for A (with $k \leq n$), then every vector v in $\mathbb{C}^n$ can be written uniquely as

$$v = v_1 + v_2 + \cdots + v_k,$$

with each v_j in V_{λ_j}. In particular, this means that every matrix has a basis of generalized eigenvectors.

Now, if v is in V_λ, then Av is also in V_λ, since $(A-\lambda I)^k Av = A(A-\lambda I)^k v = 0$. This means that the subspace V_λ is *invariant* under the matrix A. Let A_λ denote the restriction of A to the subspace V_λ, and write A_λ in the form

$$A_\lambda = \lambda I + N_\lambda$$

(i.e., we define N_λ to be $A_\lambda - \lambda I$). Then, N_λ is **nilpotent**; that is, $N_\lambda^k = 0$ for some positive integer k. We summarize the preceding discussion in the following theorem.

Theorem B.5. *Let A be an $n \times n$ complex matrix. Then, there exists a basis for $\mathbb{C}^n$ consisting of generalized eigenvectors for A. Furthermore, $\mathbb{C}^n$ is the direct sum of the generalized eigenspaces V_λ, each V_λ is invariant under A, and the restriction of A to each V_λ is of the form $\lambda I + N_\lambda$, where N_λ is nilpotent.*

Theorem B.6. *Let A be an $n \times n$ complex matrix. Then, there exists a unique pair (S, N) of matrices with the following properties: (1) $A = S + N$, (2) $SN = NS$, (3) S is diagonalizable, and (4) N is nilpotent.*

The expression $A = S + N$, with S and N as in the theorem, is called the **SN decomposition** of A. The existence of an SN decomposition follows from the previous theorem: We define S to be the operator equal to λI on each generalized eigenspace of A and we set N to be the operator equal to N_λ on each generalized eigenspace. For example, if A is the 2×2 matrix defined earlier, then we have

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

One useful result that follows fairly easily from the SN decomposition is the following.

Theorem B.7. *Every matrix is similar to an upper triangular matrix. Every nilpotent matrix is similar to an upper triangular matrix with zeros on the diagonal.*

B.4 The Jordan Canonical Form

The Jordan canonical form may be viewed as a refinement of the SN decomposition, based on a further analysis of the nilpotent matrices N_λ in Theorem B.5. Although the SN decomposition is sufficient for the purposes of this book, I discuss the Jordan canonical form here simply because it is more commonly taught in linear algebra courses. (The Jordan canonical form could be useful for a few of the exercises in Chapter 2.) To get to the Jordan form from the SN decomposition, one needs to be able to classify nilpotent matrices up to similarity.

Theorem B.8. *Every $A \in M_n(\mathbb{C})$ is similar to a block-diagonal matrix in which each block is of the form*

$$\begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}.$$

Two matrices A and B are similar if and only if they have precisely the same Jordan blocks, up to reordering.

There may be several different Jordan blocks (possibly of different sizes) for the same value of λ. In the case in which A is diagonalizable, each block is 1×1, in which case, the 1's above the diagonal do not appear. Note that each Jordan block is, in particular, of the form $\lambda I + N$, where N is nilpotent.

B.5 The Trace

If A is an $n \times n$ matrix, we define the **trace** of A to be the sum of the diagonal entries of A; that is,

$$\text{trace}(A) = \sum_{k=1}^{n} A_{kk}.$$

Note that the trace is a linear function of A (unlike the determinant).

If A and B are two $n \times n$ matrices, then

$$\text{trace}(AB) = \sum_{k=1}^{n} (AB)_{kk} = \sum_{k=1}^{n} \sum_{l=1}^{n} A_{kl} B_{lk}. \tag{B.1}$$

Similarly,

$$\text{trace}(BA) = \sum_{k=1}^{n} \sum_{l=1}^{n} B_{kl} A_{lk} = \sum_{k=1}^{n} \sum_{l=1}^{n} A_{lk} B_{kl}, \tag{B.2}$$

which is just the same sum as (B.1) with the labels for the summation variables reversed. Thus, we conclude that $\text{trace}(AB) = \text{trace}(BA)$. Then, if C is an invertible matrix and we apply this to the matrices CA and C^{-1}, we have

$$\text{trace}(CAC^{-1}) = \text{trace}(C^{-1}CA) = \text{trace}(A);$$

that is, the trace is invariant under conjugation, or, similar matrices have the same trace.

More generally, if A is a linear operator on a finite-dimensional vector space V, we can define the trace of A by picking a basis and then defining the trace of A to be the trace of the matrix that represents A in that basis. The above calculations show that the value of the trace of A is independent of the choice of basis.

B.6 Inner Products

Let $\langle \cdot, \cdot \rangle$ denote the standard inner product on $\mathbb{C}^n$, defined by

$$\langle u, v \rangle = \sum_{k=1}^{n} \overline{u_k} v_k,$$

where we follow the convention of putting the complex-conjugate on the first factor. If A is any matrix in $M_n(\mathbb{C})$, then the adjoint A^* of A has the property that

$$\langle u, Av \rangle = \langle A^* u, v \rangle \tag{B.3}$$

for all $u, v \in \mathbb{C}^n$.

If V is any vector space over $\mathbb{C}$, then an **inner product** on V is a map that associates to any two vectors u and v in V a complex number $\langle u, v \rangle$ and that has the following properties:

(1) Conjugate-symmetry: $\langle v, u \rangle = \overline{\langle u, v \rangle}$ for all $u, v \in V$.

(2) Linearity in the second factor: $\langle u, v_1 + a v_2 \rangle = \langle u, v_1 \rangle + a \langle u, v_2 \rangle$, for all $u, v_1, v_2 \in V$ and $a \in \mathbb{C}$.

(3) Positivity: For all $v \in V$, $\langle v, v \rangle$ is real and satisfies $\langle v, v \rangle \geq 0$, and $\langle v, v \rangle = 0$ only if $v = 0$.

Note that in light of the conjugate-symmetry and the linearity in the second factor, an inner product must be conjugate-linear in the first factor:

$$\langle v_1 + av_2, u\rangle = \langle v_1, u\rangle + \bar{a}\,\langle v_2, u\rangle\,.$$

An inner product on a real vector space is defined in the same way except that conjugate-symmetry is replaced by symmetry ($\langle v, u\rangle = \langle u, v\rangle$) and the constant a in Point 2 now takes only real values.

If V is a vector space with inner product, then the **norm** of a vector $v \in V$, denoted $\|v\|$, is defined by

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

The positivity condition on the inner product guarantees that $\|v\|$ is always a non-negative real number and that $\|v\| = 0$ only if $v = 0$.

As an example, consider $M_n(\mathbb{C})$, which is a complex vector space, and define an inner product on $M_n(\mathbb{C})$ by

$$\langle A, B\rangle = \operatorname{trace}(A^*B). \tag{B.4}$$

Note that

$$\operatorname{trace}(A^*B) = \operatorname{trace}((B^*A)^*) = \overline{\operatorname{trace}(B^*A)},$$

which shows that $\langle\cdot,\cdot\rangle$ is conjugate-symmetric. Linearity in the second factor follows from linearity of the trace. Finally,

$$\begin{aligned}\operatorname{trace}(A^*A) &= \sum_{k=1}^{n}(A^*A)_{kk}\\ &= \sum_{k,l=1}^{n} A^*_{kl}A_{lk}\\ &= \sum_{k,l=1}^{n} |A_{kl}|^2 \geq 0,\end{aligned}$$

and the sum is zero only if each entry of A is zero (i.e., only if A is zero). This shows that (B.4) defines an inner product. This inner product on the space of matrices is called the **Hilbert–Schmidt** inner product. The norm associated to the Hilbert–Schmidt inner product is the norm on matrices introduced in Section 2.1.

For any inner product on a finite-dimensional vector space, we can define the **adjoint** of a linear operator by the condition that $\langle u, Av\rangle = \langle A^*u, v\rangle$ for all u and v in the space. (Compare (B.3).)

Suppose that V is a finite-dimensional vector space with inner product and that W is a subspace of V. Then, the **orthogonal complement** of W, denoted $W^\perp$, is the set of all vectors v in V such that $\langle w, v\rangle = 0$ for all w in W. The basic results are (1) $(W^\perp)^\perp = W$ and (2) V decomposes as the direct sum of W and $W^\perp$. The second point means that every vector v in V can be decomposed uniquely as $v = w + u$, where $w \in W$ and $u \in W^\perp$. This, in particular, means that $\dim W + \dim W^\perp = \dim V$.

B.7 Dual Spaces

A **linear functional** on a vector space V is a linear map of V into $\mathbb{C}$. If $v_1, \ldots, v_n$ is a basis for V, then for each set of constants $a_1, \ldots, a_n$, there is a unique linear functional ϕ such that $\phi(v_k) = a_k$. If V is a finite-dimensional complex vector space, then the **dual space** to V, denoted V^*, is the set of all linear functionals on V. This is also a vector space and its dimension is the same as that of V. Since V and V^* have the same dimension, there is a temptation to think of them as being the same space. This temptation should be resisted—a failure to distinguish clearly between V and V^* is a source of much needless confusion. In certain cases, we *will* want to identify V with V^*, but I have tried to clearly indicate in all such cases that we are making such an identification and how it is made. (The identification is usually made by means of an inner product; see below.)

If W is a subspace of a vector space V, then the **annihilator subspace** of W, denoted $W\hat{}$, is the set of all ϕ in V^* such that $\phi(w) = 0$ for all w in W. Then, $W\hat{}$ is a subspace of V^*. If V is finite dimensional, then $\dim W + \dim W\hat{} = \dim V$ and the map $W \to W\hat{}$ provides a one-to-one correspondence between subspaces of V and subspaces of V^*.

Suppose that V is a finite-dimensional vector space with an inner product. Then, for each u in V, we can define a linear functional $\phi^u \in V^*$ by the formula

$$\phi^u(v) = \langle u, v \rangle .$$

Recall that we take the inner product to be linear in the second spot, so that ϕ^u is indeed a linear functional on V. It can be shown that every linear functional on V arises in this way for some unique vector u in V. (If ϕ is zero then $u = 0$. If ϕ is nonzero then we choose u to be in the orthogonal complement of the kernel of ϕ, where this orthogonal complement is one dimensional, and adjust the normalization of u until $\phi^u = \phi$.) Thus, the map $u \to \phi^u$ gives a one-to-one, onto correspondence between V and V^*. However, this correspondence is not linear! This is because u goes into the conjugate-linear spot in the inner product, as it must, since v needs to go into the linear spot in order for $\phi^u(v)$ to be a linear functional in v. Actually, the map $u \to \phi^u$ is conjugate-linear: $\phi^{\lambda u} = \bar{\lambda}\phi^u$.

So (in the finite-dimensional case), an inner product gives us a way to identify V and V^*. However, this identification is not intrinsic; it depends on the choice of the inner product. Furthermore, the identification is conjugate-linear rather than linear.

B.8 Simultaneous Diagonalization

Definition B.9. *Suppose that V is a vector space and $\mathcal{A}$ is some collection of linear operators on V. Then a* ***simultaneous eigenvector*** *for $\mathcal{A}$ is a*

nonzero vector $v \in V$ such that for all $A \in \mathcal{A}$, there exists a constant λ_A with $Av = \lambda_A v$. The numbers λ_A are the ***simultaneous eigenvalues*** *associated to v.*

Consider, for example, the space $\mathcal{D}$ of all diagonal $n \times n$ matrices. Then, for each $k = 1, \ldots, n$, the standard basis element e_k is a simultaneous eigenvector for $\mathcal{D}$. For each diagonal matrix A, the simultaneous eigenvalue associated to e_k is the k^{th} diagonal entry of A.

Proposition B.10. *If $\mathcal{A}$ is a commuting family of linear operators on a finite-dimensional complex vector space, then $\mathcal{A}$ has at least one simultaneous eigenvector.*

It is essential here that the elements of $\mathcal{A}$ commute; noncommuting families of operators typically have no simultaneous eigenvectors.

In most cases, the collection $\mathcal{A}$ of operators on V is a *subspace* of $End(V)$, the space of all linear operators from V to itself. In that case, if v is a simultaneous eigenvector for $\mathcal{A}$ then the eigenvalues λ_A for v depend linearly on A. (After all, if $A_1 v = \lambda_1 v$ and $A_2 v = \lambda_2 v$, then $(A_1 + cA_2)v = (\lambda_1 + c\lambda_2)v$.) This leads to the following definition.

Definition B.11. *Suppose that V is a vector space and $\mathcal{A}$ is a vector space of linear operators on V. Define a* ***weight*** *for $\mathcal{A}$ to be a linear functional μ on $\mathcal{A}$ such that there exists a nonzero vector $v \in V$ with*

$$Av = \mu(A)v$$

for all A in $\mathcal{A}$. For a given weight μ, the set of all vectors $v \in V$ satisfying $Av = \mu(A)v$ for all A in $\mathcal{A}$ is called the ***weight space*** *associated to the weight μ.*

That is to say, a weight is a set of simultaneous eigenvalues for the operators in $\mathcal{A}$. If V is finite dimensional and the elements of $\mathcal{A}$ all commute with one another, then there will exist at least one weight for $\mathcal{A}$.

If $\mathcal{A}$ is finite dimensional and comes equipped with an inner product, then it is often convenient to use the inner product to identify $\mathcal{A}$ and $\mathcal{A}^*$ in the definition of a weight. From this point of view, we define a weight to be an element μ of $\mathcal{A}$ (not $\mathcal{A}^*$) such that there exists a nonzero v in V with

$$Av = \langle \mu, v \rangle \, v$$

for all $A \in \mathcal{A}$.

Definition B.12. *Suppose that V is a finite-dimensional vector space and $\mathcal{A}$ is some collection of linear operators on V. Then the elements of $\mathcal{A}$ are said to be* ***simultaneously diagonalizable*** *if there exists a basis $v_1, \ldots, v_n$ for V such that each v_k is a simultaneous eigenvector for $\mathcal{A}$.*

If $\mathcal{A}$ is a vector space of linear operators on V, then saying that the elements of $\mathcal{A}$ are simultaneously diagonalizable is equivalent to saying that V can be decomposed as a direct sum of weight spaces of $\mathcal{A}$.

If a collection $\mathcal{A}$ of operators is simultaneously diagonalizable, then the elements of $\mathcal{A}$ must commute, since they commute when applied to each v_k. Conversely, if each $A \in \mathcal{A}$ is diagonalizable by itself and if the elements of $\mathcal{A}$ commute, then (it can be shown), the elements of $\mathcal{A}$ are simultaneously diagonalizable. We record these results in the following proposition.

Proposition B.13. *If $\mathcal{A}$ is a simultaneously diagonalizable family of linear operators on a finite-dimensional vector space V, then the elements of $\mathcal{A}$ commute. If $\mathcal{A}$ is a commuting collection of linear operators on a finite-dimensional vector space V and each $A \in \mathcal{A}$ is diagonalizable, then the elements of $\mathcal{A}$ are simultaneously diagonalizable.*

We close this appendix with an analog of Proposition B.1 for simultaneous eigenvectors.

Proposition B.14. *Suppose V is a vector space and $\mathcal{A}$ is a vector space of linear operators on V. Suppose $\mu_1, \ldots, \mu_m$ are distinct weights for $\mathcal{A}$ and $v_1, \ldots, v_m$ are elements of the corresponding weight spaces. If $v_1 + \cdots + v_m = 0$, then $v_k = 0$ for all $k = 1, \ldots, m$.*

C

More on Lie Groups

In this appendix, I briefly summarize (without proofs) the notion of a differentiable manifold and the notion of a general (not necessarily matrix) Lie group. I then explain briefly the standard approach to the Lie algebra and exponential mapping for general Lie groups. This means that the Lie algebra is the space of left-invariant vector fields and the exponential mapping is defined in terms of the flow along such vector fields. Although this approach is not used in the rest of the book, I cover it in order to help the reader make contact with the approach used in other books. Anyone who is going to delve deeply into the theory of Lie groups needs to learn this approach eventually. For more information on the manifold approach to Lie groups, see standard references such as Warner (1983) or Varadarajan (1974).

I will also give two examples of Lie groups that are not matrix Lie groups. The first is the group described in Section 1.8, which may also be described as a quotient of the Heisenberg group by a discrete subgroup of its center. The second is the universal cover of $\mathsf{SL}(n;\mathbb{R})$.

C.1 Manifolds

C.1.1 Definition

A **topological manifold** $\mathcal{M}$ of dimension n is a topological space (assumed second-countable and Hausdorff) that is locally homeomorphic to $\mathbb{R}^n$. This means that for each point m in $\mathcal{M}$, there is a neighborhood U of m and a one-to-one, continuous map ϕ of U into $\mathbb{R}^n$ onto some open set $\phi(U)$ in $\mathbb{R}^n$ such that the inverse map $\phi^{-1} : \phi(U) \to U$ is also continuous. We may say that a manifold is a topological space that looks locally like a little piece of $\mathbb{R}^n$. We think of the map ϕ as defining local coordinate functions $x_1, \ldots, x_n$, where each x_k is the continuous function from U into $\mathbb{R}$ given by $x_k(m) = \phi(m)_k$ (the k^{th} component of $\phi(m)$). If ψ is another homeomorphism of another neighborhood V of m, and $y_k(m) = \psi(m)_k$ is the associated coordinate system,

then both coordinate systems are defined in the neighborhood $U \cap V$ of m. We can then think of the y's as functions of the x's; more precisely, we may consider the map $\psi \circ \phi^{-1}$ that maps the set $\phi(U \cap V)$ onto the set $\psi(U \cap V)$. This is the "change of coordinates" map; that is, for all m in $U \cap V$, we have

$$(y_1(m), \ldots, y_n(m)) = \left(\psi \circ \phi^{-1}\right)(x_1(m), \ldots, x_n(m)).$$

This change of coordinates map is continuous (since both ψ and ϕ^{-1} are continuous).

A **smooth manifold** of dimension n is a topological manifold $\mathcal{M}$ together with a distinguished family of local coordinate systems (U_α, ϕ_α) with the following properties. (Here, α ranges over some indexing set.) First, every point in $\mathcal{M}$ is contained in at least one of the U_α's. Second, for any two of these coordinates systems (U_α, ϕ_α) and (U_β, ϕ_β), the change-of-coordinates map $\phi_\beta \circ \phi_\alpha^{-1}$ is a *smooth* map of the set $\phi_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$ onto the set $\phi_\beta(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$. In more concrete terms, this means that to make a smooth manifold, we start with a topological manifold and then choose a collection of local coordinate systems that cover the whole manifold and such that whenever two coordinate systems are defined in overlapping regions, the expression for one set of coordinates in terms of the other is always smooth. Note that we must *choose* these coordinate systems in order to give a smooth structure to the topological manifold $\mathcal{M}$. For some topological manifolds, it is impossible to make such a choice: Some manifolds do not admit a smooth structure. When a smooth structure exists, it is not unique. (In some cases it is unique "up to diffeomorphism," but it is never actually unique.)

Once a smooth structure is chosen, we define a **smooth local coordinate system** to be any local coordinate system (U, ϕ) (not necessarily one of the U_α's) such that $\phi \circ \phi_\alpha^{-1}$ is smooth for each (U_α, ϕ_α). A function $f : \mathcal{M} \to \mathbb{R}$ is called **smooth** if for each smooth local coordinate system (U, ϕ), the function $f \circ \phi^{-1}$ is a smooth function on the set $\phi(U)$. Another way to say this is that f is smooth if it is smooth in each smooth local coordinate system. If (U, ϕ) is a smooth local coordinate system and $x_1, \ldots, x_n$ are the associated coordinate functions $x_k(m) = \phi(m)_k$, then it is common to write $f(x_1, \ldots, x_n)$ to mean $f(\phi^{-1}(x_1, \ldots, x_n))$. In that case,

$$\frac{\partial f}{\partial x_k}(m)$$

means more pedantically the value of the k^{th} partial derivative of $f \circ \phi^{-1}$ evaluated at the point $\phi(m) = (x_1(m), \ldots, x_k(m))$.

C.1.2 Tangent space

One way to construct a manifold is as a submanifold of some Euclidean space $\mathbb{R}^n$. We may think, for example, of a smooth surface $\mathcal{S}$ inside $\mathbb{R}^3$. In that case, the tangent space at a point $m \in \mathcal{S}$ is the set of vectors v in $\mathbb{R}^3$ that

can be expressed as $v = d\gamma/dt|_{t=0}$, where $\gamma(t)$ is a smooth curve lying in $\mathcal{S}$ and satisfying $\gamma(0) = m$. For each point m in $\mathcal{S}$, the tangent space at $\mathcal{S}$ is a two-dimensional subspace of $\mathbb{R}^3$.

This description is "extrinsic"; that is, it depends on having $\mathcal{S}$ embedded inside $\mathbb{R}^3$. We want an "intrinsic" description of the tangent space, one that does not depend on having our manifold embedded inside some Euclidean space. We look, then, for some aspect of the tangent space that can be described without reference to the embedding. One possibility is to think about the directional derivative in the direction of a tangent vector v. If f is a smooth function defined on $\mathcal{S}$, we define the directional derivative of f at the point m and in the direction of the vector v to be

$$(D_v f)(m) = \left.\frac{d}{dt} f(\gamma(t))\right|_{t=0},$$

where γ is any smooth curve lying in $\mathcal{S}$ with $\gamma(0) = 0$ and $d\gamma/dt|_{t=0} = v$. (The value of the directional derivative is independent of the choice of γ.) Note that the directional derivative associates a number to each smooth function f. Also, derivatives of this sort satisfy the usual product rule for derivatives.

For a general manifold, not necessarily embedded in $\mathbb{R}^m$, we define the notion of tangent space by abstracting the notion of the directional derivative. The **tangent space at** m to $\mathcal{M}$, denoted $T_m(\mathcal{M})$, is the set of all linear maps X from $C^\infty(\mathcal{M})$ into $\mathbb{R}$ satisfying (1) the "product rule":

$$X(fg) = X(f)g(m) + f(m)X(g)$$

for all f and g in $C^\infty(\mathcal{M})$; (2) "localization": If f is equal to g in a neighborhood of m, then $X(f) = X(g)$. This is easily seen to be a real vector space. An element of $T_m(\mathcal{M})$ is called a **tangent vector** at m. If $x_1, \ldots, x_n$ is a local coordinate system, then one can prove that each tangent vector X at m can be expressed uniquely as

$$X(f) = \sum_{k=1}^{n} a_k \frac{\partial f}{\partial x_k}(m) \tag{C.1}$$

for some real constants $a_1, \ldots, a_n$. This means that if $\mathcal{M}$ is a manifold of dimension n, then for each m in $\mathcal{M}$, $T_m(\mathcal{M})$ is a real vector space of dimension n.

C.1.3 Differentials of smooth mappings

A map Φ from a manifold $\mathcal{M}$ of dimension n_1 to a manifold $\mathcal{N}$ of dimension n_2 is called **smooth** if it is smooth in local coordinates; that is, Φ is smooth if, for every coordinate system ϕ_α on $\mathcal{M}$ and every coordinate system ϕ_β on $\mathcal{N}$, $\phi_\beta \circ \Phi \circ \phi_\alpha^{-1}$ is a smooth map from an open subset of $\mathbb{R}^{n_1}$ into $\mathbb{R}^{n_2}$. Given a smooth map, one can define the differential (or derivative) of Φ at each point

m in $\mathcal{M}$, denoted $\Phi_{*,m}$. The differential $\Phi_{*,m}$ is the linear map of $T_m(\mathcal{M})$ into $T_{\Phi(m)}(\mathcal{N})$ given by

$$\Phi_{*,m}(X)(f) = X(f \circ \Phi),$$

where X is a tangent vector at m to $\mathcal{M}$ and f is a smooth (real-valued) function on $\mathcal{N}$. It is straightforward to check that $\Phi_{*,m}(X)$ is, indeed, a tangent vector to $\mathcal{N}$ at $\Phi(m)$ (i.e., that it satisfies the Leibniz rule) and that the map $\Phi_{*,m}$ is linear. In local coordinates, Φ will look like a map from $\mathbb{R}^{n_1}$ to $\mathbb{R}^{n_2}$ and $\Phi_{*,m}$ will then be essentially the matrix of partial derivatives of Φ. The differential of Φ is sometimes written as $d\Phi$.

If $\Phi : \mathcal{M} \to \mathcal{N}$ and $\Psi : \mathcal{N} \to \mathcal{P}$ are smooth maps, then $\Psi \circ \Phi : \mathcal{M} \to \mathcal{P}$ is also smooth. The chain rule in this setting takes the form

$$(\Psi \circ \Phi)_{*,m} = \Psi_{*,\Phi(m)} \circ \Phi_{*,m}. \tag{C.2}$$

Suppose that $\gamma : (a, b) \to \mathcal{M}$ is a smooth curve. Then for each $t \in (a, b)$ we will let $d\gamma/dt$ denote the element of $T_{\gamma(t)}(\mathcal{M})$ with the property that

$$\frac{d\gamma}{dt}(f) = \frac{df(\gamma(t))}{dt} \tag{C.3}$$

for all smooth functions f on $\mathcal{M}$. (Recall that we are thinking of tangent vectors as things that operate on functions.) In a smooth local coordinate system $x_1, \ldots, x_n$, we can find smooth functions $x_1(\cdot), \ldots, x_n(\cdot)$ of one variable such that $\gamma(t)$ is the point whose coordinates are $x_1(t), \ldots, x_n(t)$. (Actually, $x_k(t)$ is nothing but $x_k(\gamma(t))$. Here we make a typical abuse of notation by allowing x_k to denote both the coordinate function on $\mathcal{M}$ and the associated function on $\mathbb{R}$ obtained by evaluating x_k on γ.) In that case, the chain rule tells us that $df(\gamma(t))/dt = \sum \partial f/\partial x_k \, dx_k/dt$. Thus, (C.3) becomes

$$\frac{d\gamma}{dt} = \sum_{k=1}^{n} \frac{dx_k}{dt} \frac{\partial}{\partial x_k}. \tag{C.4}$$

C.1.4 Vector fields

A **vector field** is a map X that associates to each point m in $\mathcal{M}$ a tangent vector $X_m \in T_m(\mathcal{M})$. Given a local coordinate system $x_1, \ldots, x_n$ a vector field can be expressed (in the domain of definition of that coordinate system) as

$$X_m(f) = \sum_{k=1}^{n} a_k(m) \frac{\partial f}{\partial x_k}, \tag{C.5}$$

where the a_k's are real-valued functions. (Here we are simply using the representation (C.1) at each point.) A vector field is called **smooth** if the coefficient functions a_k are smooth in each local coordinate system. We can apply a vector field to a smooth function f by applying X_m to f at each point m. The

result $X(f)$is then another function, which will be smooth if X is a smooth vector field. So, a smooth vector field is a map from $C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ that satisfies the product rule in the form

$$X(fg) = fX(g) + X(f)g. \tag{C.6}$$

Note, here, that $X(fg)$ is a function, not a number, and that on the right-hand side, we do not evaluate f or g at any point. Equation (C.6) can be restated as saying that a vector field is a **derivation** of the algebra of smooth functions.

One should maintain a geometric picture of a vector field as a collection of arrows, one at each point in the manifold. Nevertheless, one can also think of the vector field as a differential operator (mapping the space of smooth functions to itself), the one obtained by differentiating a function at each point in the direction of the tangent vector at that point.

Looking at (C.5) we see that a vector field can be regarded as a *first-order* differential operator. If we multiply (i.e., compose) two vector fields, we will get a second-order differential operator; this is not a vector field. However, if X and Y are vector fields and we compute their commutator $XY - YX$, then the second-order terms in XY will cancel with the second-order terms in YX and the result will again be first-order differential operator (i.e., a vector field). Alternatively, one can check that if X and Y satisfy the product rule (C.6), then so does $XY - YX$. The space of smooth vector fields then becomes an infinite-dimensional Lie algebra with the bracket defined by $[X, Y] = XY - YX$. (This bracket satisfies the Jacobi identity because the composition of differential operators is associative.)

C.1.5 The flow along a vector field

If X is a vector field and $\gamma : (a, b) \to \mathcal{M}$ is a smooth curve in $\mathcal{M}$, then γ is called an **integral curve** for X if for each $t \in (a, b)$, we have $d\gamma/dt = X_{\gamma(t)}$. In a smooth local coordinate system $x_1, \ldots, x_n$, $\gamma(t)$ will be represented a family of functions $x_1(t), \ldots, x_n(t)$ and the vector field X will be represented in the form (C.5) with each a_k being a smooth function of $x_1, \ldots, x_n$. In light of (C.4), the equation $d\gamma/dt = X_{\gamma(t)}$ becomes, in local coordinates,

$$\frac{dx_k(t)}{dt} = a_k(x_1(t), \ldots, x_n(t)).$$

This is a system of first-order ordinary differential equations (not necessarily linear). Applying standard results giving uniqueness and local existence for solutions of such systems, we obtain the following results.

Theorem C.1 (Local Existence). *Given a smooth vector field X and a point $m \in \mathcal{M}$, there exist $\varepsilon > 0$ and a smooth curve $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$ such that $\gamma(0) = m$ and such that γ is an integral curve for X.*

Theorem C.2 (Uniqueness). *Suppose that X is a smooth vector field and that $\gamma_1 : (-a_1, b_1) \to \mathcal{M}$ and $\gamma_2 : (-a_2, b_2) \to \mathcal{M}$ are two integral curves for X satisfying $\gamma_1(0) = \gamma_2(0) = m$. Then, $\gamma_1(t) = \gamma_2(t)$ for all t between $-\min(a_1, a_2)$ and $\min(b_1, b_2)$; that is, the two curves agree on the interval where both of them are defined.*

Note that existence is merely local: In general, it may be impossible to find an integral curve $\gamma(t)$ through a point m that is defined for all $t \in \mathbb{R}$. In the case $\mathcal{M} = \mathbb{R}$, this amounts to asserting that first-order ordinary differential equations may not have solutions defined for all time. (Consider, for example, the separable equation $dy/dt = y^2$, the solutions of which are $y(t) = (c-t)^{-1}$.)

A vector field X is called **complete** if $\gamma(t)$ can be defined for all t for all initial points m. Any vector field on a compact manifold is always complete. On a noncompact manifold, some vector fields will be complete and some will not be. If X is a complete vector field, then one can define the associated **flow** on $\mathcal{M}$. This is a family of maps $\Phi_t : \mathcal{M} \to \mathcal{M}$ defined so that if γ is an integral curve for X with $\gamma(0) = m$, then $\Phi_t(m) = \gamma(t)$. This means that $\Phi_t(m)$ is defined by starting at m and "flowing" along the vector field X for time t. (If X is not complete, one can still define a sort of flow, but then each Φ_t is defined only on part of $\mathcal{M}$.) If X is a smooth complete vector field, then each Φ_t is a smooth map of $\mathcal{M}$ to itself, and the maps satisfy $\Phi_t \circ \Phi_s = \Phi_{t+s}$.

C.1.6 Submanifolds of vector spaces

If V is a finite-dimensional real vector space, then we may make V into a smooth manifold by using a single, globally defined linear coordinate system. Given vectors u and v in V, we can define the directional derivative of a function f at the point u in the direction of the vector v as

$$(D_v f)(u) := \frac{d}{dt} f(u + tv)\bigg|_{t=0}.$$

For each v, the directional derivative D_v satisfies the Leibniz rule. Thus, each vector v gives rise to an element of $T_u(V)$. It is not hard to show that every tangent vector at u can be expressed in this form. Thus, we have a natural way to identify $T_u(V)$ with V itself, for each $u \in V$.

Suppose V is a real vector space of dimension n. A subset $\mathcal{M}$ of V is called a **smooth embedded submanifold** of dimension k if given any m_0 in $\mathcal{M}$, there exists a smooth coordinate system (ϕ, U) defined in a neighborhood U of m_0 such that for any $m \in U$, m is in $U \cap \mathcal{M}$ if and only if $\phi(m)$ is in $\mathbb{R}^k \subset \mathbb{R}^n$. (Here, we think of $\mathbb{R}^k$ as the subset of $\mathbb{R}^n$ where the last $n-k$ coordinates are zero.) This says that locally, in a suitable coordinate system, $\mathcal{M}$ looks like $\mathbb{R}^k$ sitting inside $\mathbb{R}^n$. If $\mathcal{M}$ is a smooth embedded submanifold of dimension k, then we can make $\mathcal{M}$ into a smooth manifold of dimension k as follows. We use as our basic coordinate neighborhoods the sets of the form

$U \cap \mathcal{M}$. We use as our coordinate map on such a set the restriction of ϕ to $U \cap \mathcal{M}$. (This restriction maps $U \cap \mathcal{M}$ to $\mathbb{R}^k \subset \mathbb{R}^n$.)

If $\mathcal{M}$ is a smooth embedded submanifold of V, then the inclusion map i of $\mathcal{M}$ into V is a smooth map. (Here, i is defined by $i(m) = m$ for $m \in \mathcal{M}$.) The differential $i_* : T_m(\mathcal{M}) \to T_m(V)$ is injective, and it is customary to identify $T_m(\mathcal{M})$ with its image in $T_m(V)$, which is a k-dimensional subspace of the n-dimensional space $T_m(V)$. To say this more explicitly, if X is a tangent vector to $\mathcal{M}$ at m (viewed as a map of $C^\infty(\mathcal{M})$ to $\mathbb{R}$), then we can make X into a tangent vector $\tilde{X}$ at m to V by defining

$$\tilde{X}(f) = X(f|_{\mathcal{M}}).$$

This allows us to think of the tangent space to $\mathcal{M}$ at m as a subspace of the tangent space to V at m. However, we are identifying the tangent space at m to V with V itself. Thus, the tangent space to $\mathcal{M}$ at m is identified with a subspace of V. (This subspace depends on the point m.)

It is not hard to show that if $\mathcal{M}$ is a smooth embedded submanifold of V, then for each m, $T_m(\mathcal{M})$ (viewed as a subspace of V), is the usual geometric tangent space, as follows.

Proposition C.3. *Let $\mathcal{M}$ be a smooth embedded submanifold of a finite-dimensional real vector space V. Then, for each $m \in \mathcal{M}$, the tangent space to $\mathcal{M}$ at m (regarded as a subspace of V as above) is the set of all u in V such that there exists a smooth curve γ in $\mathcal{M}$ with $\gamma(0) = m$ and $d\gamma/dt = u$.*

C.1.7 Complex manifolds

A complex manifold is a smooth manifold of dimension $2n$ such that the basic coordinate patches (U_α, ϕ_α) have the property that the change-of-coordinates map $\phi_\beta \circ \phi_\alpha^{-1}$ is *holomorphic* for each α and β. Here, $\mathbb{R}^{2n}$ is identified with $\mathbb{C}^n$ and holomorphic means the same as complex analytic. Then, any other local coordinate system ϕ (not necessarily one of the ϕ_α's) is said to be holomorphic if the change-of-coordinates map between ϕ and each of the ϕ_α's is holomorphic. A map between two complex manifolds is said to be holomorphic if it is holomorphic in each holomorphic local coordinate system. If V is a complex vector space, then a subset $\mathcal{M}$ of V is called an **embedded complex submanifold** of dimension k if, given any m_0 in $\mathcal{M}$, there exists a holomorphic local coordinate system (ϕ, U) defined in a neighborhood U of m_0 such that for any $m \in U$, m is in $U \cap \mathcal{M}$ if and only if $\phi(m)$ is in $\mathbb{C}^k \subset \mathbb{C}^n$.

C.2 Lie Groups

C.2.1 Definition

A Lie group is a smooth manifold that is also a group. More precisely, we have the following.

Definition C.4. *A **Lie group** is a smooth manifold G together with a smooth map from $G \times G \to G$ that makes G into a group and such that the inverse map $g \to g^{-1}$ is a smooth map of G to itself.*

The simplest example is $G = \mathbb{R}^n$, with the product map given by $(x, y) \to x + y$. A more interesting example was given in Section 1.8. See also Section C.3. It is shown in Chapter 2 that every matrix Lie group is a Lie group. (See also Subsection C.2.6.)

C.2.2 The Lie algebra

If G is a Lie group and g an element of G, we define a map $L_g : G \to G$ by $L_g(h) = gh$. This is the "left multiplication by g" map, which is smooth since the product map of $G \times G$ to itself is assumed smooth. Then, the differential $(L_g)_*$ of L_g at a point h will be a linear map of $T_h(G)$ to $T_{gh}(G)$. A vector field X on G is called **left-invariant** if X satisfies

$$(L_g)_*(X_h) = X_{gh}.$$

Let $T_e(G)$ denote the tangent space at the identity. Then, given any vector $v \in T_e(G)$, there is a unique left-invariant vector field X^v with $X^v_e = v$, which can be constructed by defining

$$X^v_g = (L_g)_*(v).$$

To show that the vector field constructed in this way is left-invariant, one needs to note that $L_g \circ L_h = L_{gh}$, from which it follows (by the chain rule (C.2)) that $(L_{gh})_{*,e} = (L_g)_{*,h}(L_h)_{*,e}$. It should be evident that every left-invariant vector field arise in this way (with v equal to the value of the left-invariant vector field at the identity). The set of all left-invariant vector fields is a real vector space whose dimension is the same as that of G, and it is isomorphic as a vector space to $T_e(G)$ by means of evaluation at the identity.

Recall that if we think of vector fields as first-order differential operators, then the commutator of two vector fields is, again, a vector field. It is not difficult to show that the commutator of two left-invariant vector fields is, again, a left-invariant vector field.

Definition C.5. *The **Lie algebra** $\mathfrak{g}$ of a Lie group G is the tangent space at the identity with the bracket operation defined by*

$$[v, w] = [X^v, X^w]_e.$$

If we identify the space of left-invariant vector fields with $T_e(G)$ by means of the map $v \longleftrightarrow X^v$, then $\mathfrak{g}$ is just the space of left-invariant vector fields, which forms a Lie algebra under the commutator of vector fields.

C.2.3 The exponential mapping

The exponential mapping for a general Lie group is defined in terms of the flow along left-invariant vector fields. This definition is justified by the following result.

Proposition C.6. *If G is a Lie group, then every left-invariant vector field on G is complete.*

Definition C.7. *Let G be a Lie group and let $\mathfrak{g} = T_e(G)$ be its Lie algebra. For each $v \in \mathfrak{g}$, let X^v be the associated left-invariant vector field and let Φ_t^v be the associated flow. Then, the* ***exponential mapping*** *is the map $\exp : \mathfrak{g} \to G$ defined by*

$$\exp(v) = \Phi_1^v(e).$$

This means that to compute $\exp(v)$, we first construct the left-invariant vector field X^v and we then find an integral curve γ^v to X^v that starts at the identity. Then, $\exp(v) = \gamma^v(1)$. To say this yet again, the exponential mapping is the time-one flow along a left-invariant vector field starting at the identity.

It can be shown that the exponential mapping is a smooth map of $\mathfrak{g}$ into G and that differential of exp at the origin is the identity map of $\mathfrak{g}$ to itself. (Here, we identity both $T_0(\mathfrak{g})$ and $T_e(G)$ with $\mathfrak{g}$.) It then follows from the inverse function theorem that the exponential mapping takes all sufficiently small neighborhoods of the origin in $\mathfrak{g}$ diffeomorphically onto neighborhoods of the identity in G. The properties of the exponential mapping that we have proved for matrix Lie groups continue to hold for general Lie groups. For, example, $\exp(v+w) = \exp v \exp w$ whenever $[v, w] = 0$, the Lie product formula holds, and the Baker–Campbell–Hausdorff formula holds.

C.2.4 Homomorphisms

Suppose $\Phi : G \to H$ is a smooth map of a Lie group G to a Lie group H that is also a group homomorphism. Then, Φ is called a **Lie group homomorphism**. (In the matrix case, we originally required only that Lie group homomorphisms be continuous. However, we proved (Section 2.7) that every continuous homomorphism between matrix Lie groups is actually smooth.)

If $\Phi : G \to H$ is a Lie group homomorphism, then the differential $\Phi_{*,g}$ maps $T_g(G)$ into $T_{\Phi(g)}(H)$. In particular, $\Phi_{*,e}$ is a linear map of $\mathfrak{g} = T_e(G)$ into $\mathfrak{h} = T_e(H)$.

Proposition C.8. *Let $\Phi : G \to H$ be a Lie group homomorphism and set $\phi = \Phi_{*,e}$, so that ϕ is a linear map of $\mathfrak{g}$ into $\mathfrak{h}$. Then, ϕ is a Lie algebra homomorphism and satisfies*

$$\exp(\phi(v)) = \Phi(\exp v)$$

for all $v \in \mathfrak{g}$.

Note that the way ϕ is defined is consistent with the way we defined things in the matrix case (Theorem 2.21).

For each g in G, let $C_g : G \to G$ be the "conjugation by g" map; that is, $C_g(h) = ghg^{-1}$. For each g, C_g is a Lie group homomorphism of G to itself. Thus, the differential of C_g at the identity is a Lie algebra homomorphism of $\mathfrak{g}$ to itself. This map is denoted Ad_g:

$$\mathrm{Ad}_g = (C_g)_{*,e}.$$

This can be computed in a more concrete way as

$$\mathrm{Ad}_g(X) = \frac{d}{dt} g e^{tX} g^{-1} \bigg|_{t=0}.$$

Note, here, that $ge^{tX}g^{-1}$ is a smooth curve in G that passes through the identity at $t = 0$, and, thus, the derivative of this curve at $t = 0$ is an element of $\mathfrak{g} = T_e(G)$.

C.2.5 Quotient groups and covering groups

Given a connected smooth manifold $\mathcal{M}$, there is a construction (which I will not attempt to describe here) that yields a simply-connected manifold $\tilde{\mathcal{M}}$ together with a map $\Phi : \tilde{\mathcal{M}} \to \mathcal{M}$ with the following property: Each $m \in \mathcal{M}$ has a neighborhood U such that $\Phi^{-1}(U)$ is the disjoint union of open sets V_α, each of which is mapped by Φ diffeomorphically onto U. Such a pair $(\tilde{\mathcal{M}}, \Phi)$ is called a **universal cover** of $\mathcal{M}$ and is "unique up to canonical diffeomorphism." If $\mathcal{M} = G$ is a connected Lie group, then the universal cover $\tilde{G}$ can be given a group structure in a canonical way and the map Φ in this case is a Lie group homomorphism of $\tilde{G}$ onto G. The associated Lie algebra of $\phi : \tilde{\mathfrak{g}} \to \mathfrak{g}$ is an isomorphism; therefore, we often say that $\tilde{G}$ and G have "the same" Lie algebra.

This construction illustrates the advantage of working in the general Lie group setting: Every Lie group has a universal cover that is, again, a Lie group and that can be constructed in a canonical way. By contrast, the universal cover $\tilde{G}$ of a *matrix* Lie group G may not be a matrix Lie group, and even if it is, there is no canonical procedure for finding a matrix representation of $\tilde{G}$. For example, the universal cover of $\mathsf{SL}(n;\mathbb{R})$, $n \geq 2$, is not a matrix Lie group, as shown in the next section.

Meanwhile, suppose that G is a Lie group and that N is a closed normal subgroup of G. Then, there is a unique manifold structure on the quotient group G/N that makes G/N into a Lie group and such that the quotient map $G \to G/N$ is then a Lie group homomorphism. That this procedure can be carried out for *any* Lie group G and *any* closed normal subgroup N again illustrates the advantage of working with general Lie groups. By contrast, if G is a matrix Lie group, G/N may not be, as the first example in the next section shows. Furthermore, even if G/N happens to be a matrix Lie group, there is no canonical procedure for finding a matrix representation of it.

C.2.6 Matrix Lie groups as Lie groups

We have proved that every matrix Lie group is a smooth embedded submanifold of the vector space $V = M_n(\mathbb{C})$. (Here we think of $M_n(\mathbb{C})$ as a *real* vector space of dimension $2n^2$. A matrix Lie group is in general only a real embedded submanifold of $M_n(\mathbb{C})$ and not necessarily a complex embedded submanifold.) Since the matrix product and matrix inverse are smooth on the open subset $\mathsf{GL}(n;\mathbb{C})$ of $M_n(\mathbb{C})$, this shows that every matrix Lie group is a Lie group. (The restrictions of smooth mappings to smooth embedded submanifolds are smooth.)

Meanwhile, the Lie algebra $\mathfrak{g}$ of a matrix Lie group G (as we have defined $\mathfrak{g}$ in Chapter 2) is just the tangent space to G at the identity. To see this, note that every X in $\mathfrak{g}$ is the derivative of a smooth curve through the identity, namely the curve $\gamma(t) = e^{tX}$. Conversely, using the local logarithm , one can show that if X is the derivative of any smooth curve in G passing through I at $t = 0$, then X is in $\mathfrak{g}$. (See Corollary 2.35 in Section 2.7.)

It remains to show that the exponential map as we have defined it in the matrix case agrees with the exponential map as we have defined for general Lie groups. If $X \in \mathfrak{g}$, what we need to show is that the curve

$$\gamma(t) = e^{tX}$$

is an integral curve for the left-invariant vector field whose value at the identity is X. This means that we must show that

$$\frac{d}{dt} e^{tX} = (L_{e^{tX}})_* (X).$$

To do this, we note that

$$\begin{aligned} \frac{d}{dt} e^{tX} &= \frac{d}{da} e^{(t+a)X} \Big|_{a=0} = \frac{d}{da} e^{tX} e^{aX} \Big|_{a=0} \\ &= (L_{e^{tX}})_* \frac{d}{da} e^{aX} \Big|_{a=0} = (L_{e^{tX}})_* (X). \end{aligned}$$

The second-to-last equality is essentially the chain rule.

We conclude, then, that every matrix Lie group is a Lie group and that the way the Lie algebra and the exponential mapping are defined in the matrix case is consistent with the way the Lie algebra and the exponential mapping are defined for general Lie groups.

C.2.7 Complex Lie groups

A **complex Lie group** is a complex manifold endowed with a group structure in such a way that the product and inverse maps are holomorphic.

Suppose that G is a Lie group of dimension $2n$ with Lie algebra $\mathfrak{g}$. If there exists a real-linear map $J : \mathfrak{g} \to \mathfrak{g}$ such that (1) $J^2 = -I$ and (2)

$[JX, Y] = J[X, Y]$ for all X and Y in $\mathfrak{g}$, then we say that J is a complex structure on the Lie algebra $\mathfrak{g}$. In that case, we can regard $\mathfrak{g}$ as a complex Lie algebra by defining the "multiplication by i" map to be J. Condition 1 tells us that J makes $\mathfrak{g}$ into a complex vector space. Condition 2 then implies that the bracket is complex-bilinear.

If G is a complex Lie group, then there is a natural way of defining a map J on $\mathfrak{g}$ so that $\mathfrak{g}$ becomes a complex Lie algebra and so that the exponential mapping is holomorphic from $\mathfrak{g}$ into G. Conversely, suppose G is an even-dimensional Lie group and we can find a map $J : \mathfrak{g} \to \mathfrak{g}$ that makes $\mathfrak{g}$ into a complex Lie group. Then G can be given the structure of a complex manifold in such a way that G becomes a complex Lie group and the exponential mapping is holomorphic. (There may be many different possible complex structures on G that make G into a complex Lie group. In that case, different complex structures on G will correspond to different complex structures J on $\mathfrak{g}$.) To oversimplify slightly, we may say that G is a complex Lie group if and only if $\mathfrak{g}$ is a complex Lie algebra. See Varadarajan (1974) for more information.

If $G \subset \mathrm{GL}(n;\mathbb{C})$ is a matrix Lie group and its Lie algebra $\mathfrak{g}$ is a complex subalgebra of $\mathfrak{gl}(n;\mathbb{C})$, then $\mathfrak{g}$ has a complex structure, namely the usual multiplication by i map. In that case, it can be shown that G is an embedded complex submanifold of $\mathrm{GL}(n;\mathbb{C})$ and, thus, a complex Lie group. This shows that the definition of a complex matrix Lie group given in Chapter 2 (Definition 2.20) is sensible: A complex matrix Lie group is indeed a complex Lie group.

C.3 Examples of Nonmatrix Lie Groups

We explained in the previous section that every matrix Lie group is (as the name suggests) a Lie group. In this section, we will show that the converse is not true: Not every Lie group is isomorphic to a matrix Lie group.

Our first example is the Lie group G introduced in Section 1.8, namely $G = \mathbb{R} \times \mathbb{R} \times S^1$, with the group product defined by

$$(x_1, y_1, u_1) \cdot (x_2, y_2, u_2) = (x_1 + x_2, y_1 + y_2, e^{ix_1 y_2} u_1 u_2).$$

Meanwhile, let H be the Heisenberg group (i.e., the group of 3×3 real matrices that are upper triangular matrices with ones on the diagonal). Consider the map $\Phi : H \to G$ given by

$$\Phi \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} = (a, c, e^{ib}).$$

Direct computation shows that Φ is a homomorphism. The kernel of Φ is the discrete normal subgroup N of H given by

$$N = \left\{ \begin{pmatrix} 1 & 0 & 2\pi n \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \middle| \, n \in \mathbb{Z} \right\}.$$

Now, suppose that Ψ is any finite-dimensional representation of G (i.e., a continuous map of G into some $\mathsf{GL}(n;\mathbb{C})$). Then, we can define an associated representation Σ of H by $\Sigma = \Psi \circ \Phi$. Clearly, the kernel of any such representation of H must include the kernel N of Φ. Now, let $Z(H)$ denote the center of H, which is easily shown to be

$$Z(H) = \left\{ \begin{pmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \middle| \, b \in \mathbb{R} \right\}.$$

Theorem C.9. *Let Σ be any finite-dimensional representation of H. If $\ker \Sigma \supset N$, then $\ker \Sigma \supset Z(H)$.*

Once this is established, we will be able to conclude that there are no faithful finite-dimensional representations of G. After all, if Ψ is any finite-dimensional representation of G, then the kernel of $\Sigma = \Psi \circ \Phi$ will contain N and, thus, $Z(H)$, by the theorem. Thus, for all $b \in \mathbb{R}$,

$$\Sigma \begin{pmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \Psi(0, 0, e^{ib}) = I.$$

This means that the kernel of Ψ contains all elements of the form $(0, 0, u)$ and Ψ is not faithful. So, we obtain the following result.

Corollary C.10. *The Lie group G has no faithful finite-dimensional representations. In particular, G is not isomorphic to any matrix Lie group.*

We now proceed with the proof of Theorem C.9.

Proof. Let σ be the associated representation of the Lie algebra $\mathfrak{h}$ of H. Let A, B, and C be the basis elements for $\mathfrak{h}$ given by

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

These satisfy the commutation relations $[A, C] = B$ and $[A, B] = [C, B] = 0$. Thus, $[\sigma(A), \sigma(C)] = \sigma(B)$ and $[\sigma(A), \sigma(B)] = [\sigma(C), \sigma(B)] = 0$.

I now claim that $\sigma(B)$ must be nilpotent. In light of the SN decomposition, this is equivalent to showing that all of the eigenvalues of $\sigma(B)$ are zero. So, let λ be an eigenvalue for $\sigma(B)$ and let V_λ be the associated eigenspace. Certainly, V_λ is invariant under $\sigma(B)$ since $\sigma(B) = \lambda I$ on V_λ. Furthermore, since $\sigma(A)$ and $\sigma(C)$ commute with $\sigma(B)$, they must also leave V_λ invariant

(Proposition B.4). The restrictions of these operators to V_λ must still satisfy the same commutation relations as they do on the whole space. This means that the restriction of $\sigma(B)$ to V_λ is the commutator of two operators on V_λ (i.e., operators that map V_λ to itself), namely the restrictions of $\sigma(A)$ and $\sigma(C)$ to V_λ. Since $\sigma(B)|_{V_\lambda}$ is a commutator, its trace must be zero (since the trace of UV is the same as the trace of VU for any operators U and V). On the other hand, $\sigma(B)|_{V_\lambda} = \lambda I$. So, $0 = \text{trace}(\sigma(B)|_{V_\lambda}) = \lambda \dim V_\lambda$. If λ is actually an eigenvalue, then $\dim V_\lambda \neq 0$ and, thus, we must have $\lambda = 0$. Since λ was an arbitrary eigenvalue of $\sigma(B)$, we conclude that 0 is the only eigenvalue of $\sigma(B)$ and, thus, $\sigma(B)$ is nilpotent.

Lemma C.11. *If X is a nonzero nilpotent matrix, then for all nonzero real numbers t, $e^{tX} \neq I$.*

Proof. Since X is nilpotent, the power series for e^{tX} terminates after a finite number of terms. Thus, each entry of e^{tX} depends polynomially on t; that is, there exist polynomials $p_{kl}(t)$ such that $(e^{tX})_{kl} = p_{kl}(t)$. Now, suppose that there is some nonzero t_0 such that $e^{t_0 X} = I$. Then, $e^{nt_0 X} = I^n = I$ for all integers n. In terms of the polynomials p_{kl}, this means that $p_{kl}(nt_0) = \delta_{kl}$ for all n. However, a polynomial that takes on a certain value infinitely many times must be constant. Thus, as soon as there is one nonzero t_0 for which $e^{t_0 X} = I$, we must have $e^{tX} = I$ for all t (assuming, still, that X is nilpotent!). This, however, would then imply that $X = d/dt(\exp tX)|_{t=0} = 0$. So, if X is nonzero and nilpotent, e^{tX} must be different from the identity for all nonzero t. □

Now, we note that the 3×3 matrix B satisfies $B^2 = 0$, and, so,

$$e^{tB} = \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Suppose now that Σ is a representation of H and that $\ker \Sigma \supset N$. Then,

$$e^{2\pi n \sigma(B)} = \Sigma\left(e^{2\pi n B}\right) = I$$

for all integers n. Since $\sigma(B)$ is nilpotent, the only way this can happen (according to the lemma) is if $\sigma(B)$ is zero. Thus, for all real numbers t,

$$\Sigma\left(e^{tB}\right) = e^{t\sigma(B)} = I$$

and $\ker \Sigma \supset Z(H)$. □

We now turn to our second example of a nonmatrix Lie group. We make use of the following topological result: For all $n \geq 2$, $\mathsf{SL}(n;\mathbb{R})$ is not simply connected and $\mathsf{SL}(n;\mathbb{C})$ is simply connected. This result was recorded in the

table in Chapter 1; the method of proof is described in Appendix E. We also make use of the concept of the universal cover of a Lie group, as described in the previous section. The universal cover of $\mathsf{SL}(n;\mathbb{R})$, denoted $\widetilde{\mathsf{SL}}(n;\mathbb{R})$, is a simply-connected Lie group together with a Lie group homomorphism $\Phi : \widetilde{\mathsf{SL}}(n;\mathbb{R}) \to \mathsf{SL}(n;\mathbb{R})$ such that the associated Lie algebra homomorphism $\phi : \widetilde{\mathsf{sl}}(n;\mathbb{R}) \to \mathsf{sl}(n;\mathbb{R})$ is an isomorphism. It is customary to permanently identify $\widetilde{\mathsf{sl}}(n;\mathbb{R})$ with $\mathsf{sl}(n;\mathbb{R})$ by means of the map ϕ and say (in a slight abuse of terminology) that the Lie algebra of $\widetilde{\mathsf{SL}}(n;\mathbb{R})$ *is* $\mathsf{sl}(n;\mathbb{R})$.

Theorem C.12. *There does not exist any faithful finite-dimensional representation of* $\widetilde{\mathsf{SL}}(n;\mathbb{R})$. *Thus,* $\widetilde{\mathsf{SL}}(n;\mathbb{R})$ *is a Lie group that is not isomorphic to any matrix Lie group.*

Proof. We will show that if Π is any finite-dimensional representation of $\widetilde{\mathsf{SL}}(n;\mathbb{R})$, then the kernel of Π contains the kernel of the homomorphism $\Phi : \widetilde{\mathsf{SL}}(n;\mathbb{R}) \to \mathsf{SL}(n;\mathbb{R})$. Since $\mathsf{SL}(n;\mathbb{R})$ is not simply connected, Φ has a nontrivial kernel and, thus, Π is not faithful (i.e., has a nontrivial kernel).

So, now let Π be a finite-dimensional representation of $\widetilde{\mathsf{SL}}(n;\mathbb{R})$, acting on a finite-dimensional complex vector space V. We then have an associated Lie algebra representation π of $\mathsf{sl}(n;\mathbb{R})$, which is (isomorphic to) the Lie algebra of $\widetilde{\mathsf{SL}}(n;\mathbb{R})$. We may extend π by complex-linearity to a representation of $\mathsf{sl}(n;\mathbb{C})$, also denoted π. Then, since $\mathsf{SL}(n;\mathbb{C})$ is simply connected, we may exponentiate π to a representation Π' of $\mathsf{SL}(n;\mathbb{C})$. Finally, we may restrict Π' to $\mathsf{SL}(n;\mathbb{R})$.

Now, construct a new representation Σ of $\widetilde{\mathsf{SL}}(n;\mathbb{R})$ by setting $\Sigma = \Pi' \circ \Phi$. I claim that Σ coincides with the original representation Π of $\widetilde{\mathsf{SL}}(n;\mathbb{R})$. To see this, consider the associated Lie algebra representation $\sigma = \pi' \circ \phi$. Now, we are regarding ϕ as the identity map of $\mathsf{sl}(n;\mathbb{R})$ to itself. Meanwhile, by construction, the Lie algebra map π' associated to Π' is π. So, in fact, $\sigma = \pi$. Since $\widetilde{\mathsf{SL}}(n;\mathbb{R})$ is connected this implies that the associated group homomorphisms Σ and Π are equal. (We have proved such a result for matrix Lie groups; the same result with much the same proof holds for all Lie groups.)

So, every representation Π of $\widetilde{\mathsf{SL}}(n;\mathbb{R})$ is of the form $\Pi = \Pi' \circ \Phi$ for some representation Π' of $\mathsf{SL}(n;\mathbb{R})$. Thus, $\ker \Pi \supset \ker \Phi \neq \{e\}$ and Π is not faithful. □

These two examples illustrate the limitations of matrix Lie groups with respect to the operations of quotients and universal covers. The group G is isomorphic to the quotient group H/N, and although H is a matrix Lie group, H/N is not. Similarly, $\mathsf{SL}(n;\mathbb{R})$ is a matrix Lie group, but its universal cover is not.

C.4 Differential Forms and Haar Measure

Recall from Section 4.10 the notion of Haar measure. A **left Haar measure** on a Lie group G is a nonzero measure that is locally finite and left-invariant. The simplest way to prove the existence of such a measure is to use differential forms.

Differential forms are an important aspect of the theory of differentiable manifolds. I have no space here to describe this notion in any detail, but the idea is roughly this. A k-form η on an n-dimensional manifold $\mathcal{M}$ is an object that can be expressed in local coordinates as

$$\eta(m) = \sum a_{i_1 i_2 \cdots i_k}(m)\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}.$$

Here, $\wedge$ is the "wedge product," which is defined in such a way that the expression $dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}$ changes sign with the interchange of any two factors. In coordinate-independent language, a k-form is something that takes values at each point in the k^{th} exterior power of the cotangent space at m, where the cotangent space is defined as the dual space to the tangent space at m. The significance of the concept of k-forms is that there is a natural (coordinate-independent) way of integrating a k-form over (oriented) k-dimensional submanifolds of $\mathcal{M}$. In particular, if $\mathcal{M}$ is oriented and η is an n-form, then it makes sense to integrate η over $\mathcal{M}$ itself.

Given an n-dimensional oriented manifold and an nowhere-vanishing oriented n-form η, we can make a measure on $\mathcal{M}$ by defining the integral of f against μ to be the integral of the n-form $f\eta$. It is not hard to show that on an n-dimensional Lie group G, there exists a nowhere-vanishing n-form that is invariant under left translations and that this form is unique up to a constant. Integrating functions against this form (with the correct orientation) gives a left-invariant measure (i.e., a left Haar measure).

Suppose that μ is a left Haar measure on G and g is an element of G. Suppose we define a new measure $R_g(\mu)$ by setting $R_g(\mu)(E) = \mu(R_g E)$, where R_g denotes right-translation by g. Then, it is not hard to see that $R_g(\mu)$ is again left-invariant (because left-translations commute with right-translations) and again given by integration against a left-invariant n-form. However, the left-invariant n-form describing $R_g(\mu)$ may not be equal to the left-invariant n-form describing μ; it may be a constant multiple of that n-form. So, given any $g \in G$, there is a constant $\chi(g)$ such that $R_g(\mu) = \chi(g)\mu$, where μ is a left Haar measure. The function $\chi(g)$ is called the **modular function** of G. A group G is called **unimodular** if the modular function is identically equal to one. If G is unimodular, then $R_g(\mu) = \mu$ (i.e., the left Haar measure is also right-invariant).

Using the description of Haar measure in terms of left-invariant n-forms, one can show the following result.

Proposition C.13. *If G is a connected Lie group, then G is unimodular if and only if*

$$\det(\mathrm{Ad}_g) = 1$$

for all $g \in G$ or, equivalently, if and only if

$$\mathrm{trace}(\mathrm{ad}_X) = 0$$

for all $X \in \mathfrak{g}$.

Here, Ad_g and ad_X are viewed as linear maps of the Lie algebra $\mathfrak{g}$ to itself. The reason for this result is as follows. A right-translation can be expressed as a combination of a left-translation and the adjoint action: $R_g(h) = \mathrm{Ad}_g L_g(h)$. So, G is unimodular if and only if the left Haar measure is invariant under the adjoint action. Since everything in sight is left-invariant, it suffices to check the invariance of a left-invariant n-form under Ad_g at the identity. So, the relevant question is how Ad_g acts on the n^{th} exterior power of the cotangent space at the identity, which is by the determinant of Ad_g.

This proposition implies that compact Lie groups are unimodular, since, then, there exists an inner product with respect to which ad_X is skew-symmetric, and the trace of a real skew matrix must be zero. More generally, it can be shown that all semisimple groups are unimodular. Meanwhile, the Heisenberg group is also unimodular because in this case ad_X is nilpotent for all X in the Lie algebra. The simplest example of a group that is not unimodular is the following two-dimensional matrix group:

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & \frac{1}{a} \end{pmatrix} \middle| \, a \in (0, \infty), \, b \in \mathbb{R} \right\}.$$

The reader is invited to verify that there exists elements X of the Lie algebra $\mathfrak{g}$ of G such that $\mathrm{trace}(\mathrm{ad}_X) \neq 0$.

D

Clebsch–Gordan Theory for SU(2) and the Wigner–Eckart Theorem

D.1 Tensor Products of $\mathsf{sl}(2;\mathbb{C})$ Representations

Recall from Section 4.6 the notion of the tensor product of representations of a group or Lie algebra. We consider this in the case of the irreducible representations of the group $\mathsf{SU}(2)$ or, equivalently, the irreducible complex-linear representations of $\mathsf{sl}(2;\mathbb{C})$. These irreducible representations were classified in Section 4.4. For each non-negative integer m, we have an irreducible representation (π_m, V_m) of $\mathsf{sl}(2;\mathbb{C})$ of dimension $m+1$, and every irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ is isomorphic to one of these. We regard the tensor product $V_m \otimes V_n$ as a representation of $\mathsf{sl}(2;\mathbb{C})$. (Recall that it is also possible to view $V_m \otimes V_n$ as a representation of $\mathsf{sl}(2;\mathbb{C}) \oplus \mathsf{sl}(2;\mathbb{C})$.) The action of $\mathsf{sl}(2;\mathbb{C})$ on $V_m \otimes V_n$ is given by

$$(\pi_m \otimes \pi_n)(X) = \pi_m(X) \otimes I + I \otimes \pi_n(X). \tag{D.1}$$

We use the standard basis $\{X, Y, H\}$ for $\mathsf{sl}(2;\mathbb{C})$.

By the averaging method of Section 4.10, we can find on each space V_m an inner product that is invariant under the action of the compact group $\mathsf{SU}(2)$. (In the case of $V_1 \cong \mathbb{C}^2$, we can use the standard inner product on $\mathbb{C}^2$.) With respect to such an inner product, the orthogonal complement of a subspace invariant under $\mathsf{SU}(2)$ (or $\mathsf{su}(2)$ or $\mathsf{sl}(2;\mathbb{C})$) is again invariant under $\mathsf{SU}(2)$ (or $\mathsf{su}(2)$ or $\mathsf{sl}(2;\mathbb{C})$). Since the element H of $\mathsf{sl}(2;\mathbb{C})$ is in $i\mathsf{su}(2)$, $\pi_m(H)$ will be self-adjoint with respect to this inner product. This means that eigenvectors of $\pi_m(H)$ with distinct eigenvectors must be orthogonal. Once we have chosen $\mathsf{SU}(2)$-invariant inner products on V_m and V_n, there is a unique inner product on $V_m \otimes V_n$ with the property that $\langle u_1 \otimes v_1, u_2 \otimes v_2 \rangle = \langle u_1, u_2 \rangle \langle v_1, v_2 \rangle$. (This can be proved using the universal property of tensor products.) The inner product on $V_m \otimes V_n$ is also invariant under the action of $\mathsf{SU}(2)$. We assume in the rest of this section that an inner product of this sort has been chosen on each $V_m \otimes V_n$.

In general, $V_m \otimes V_n$ will not be an irreducible representation of $\mathsf{sl}(2;\mathbb{C})$; the goal of this section is to describe how $V_m \otimes V_n$ decomposes as a direct sum

of irreducible invariant subspaces. This decomposition is referred to as the Clebsch–Gordan theory or (in the physics literature) as addition of angular momentum. Here, we use the mathematician's labeling of the representations of $\mathsf{sl}(2;\mathbb{C})$; physicists normally label the representations by the "spin" $l = m/2$, so that the possible values of l are $l = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$.

Let us consider first the case of $V_1 \otimes V_1$, where $V_1 = \mathbb{C}^2$, the standard representation of $\mathsf{sl}(2;\mathbb{C})$. If $\{e_1, e_2\}$ is the standard basis for $\mathbb{C}^2$, then the vectors of the form $e_k \otimes e_l$, $1 \leq k, l \leq 2$, form a basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$. Since e_1 and e_2 are eigenvalues for $\pi_1(H)$ with eigenvalues 1 and -1, respectively, then, by (D.1), the basis elements for $\mathbb{C}^2 \otimes \mathbb{C}^2$ are eigenvectors for the action of H with eigenvalues 2, 0, 0, and -2, respectively. Since 2 is the largest eigenvalue for H, the corresponding eigenvector $e_1 \otimes e_1$ must be annihilated by X (i.e., by the operator $\pi_1(X) \otimes I + I \otimes \pi_1(X)$). If we apply Y repeatedly to $e_1 \otimes e_1$, we obtain $e_1 \otimes e_2 + e_2 \otimes e_1$ and then $2e_2 \otimes e_2$ and then zero. The space spanned by these vectors is invariant under $\mathsf{sl}(2;\mathbb{C})$ and irreducible, isomorphic to the three-dimensional representation V_2. The orthogonal complement of this space in $\mathbb{C}^2 \otimes \mathbb{C}^2$, namely the span of $e_1 \otimes e_2 - e_2 \otimes e_1$, is also invariant, and $\mathsf{sl}(2;\mathbb{C})$ acts trivially on this space. So, we have

$$\mathbb{C}^2 \otimes \mathbb{C}^2 = \text{span}\{e_1 \otimes e_1, e_1 \otimes e_2 + e_2 \otimes e_1, e_2 \otimes e_2\} \oplus \text{span}\{e_1 \otimes e_2 - e_2 \otimes e_1\}.$$

Thus the four-dimensional space $V_1 \otimes V_1$ is isomorphic, as an $\mathsf{sl}(2;\mathbb{C})$ representation, to $V_2 \oplus V_0$.

Theorem D.1. *For any non-negative integer k, let V_k denote the irreducible representation of $\mathsf{sl}(2;\mathbb{C})$ of dimension $k+1$. For two non-negative integers m and n, consider $V_m \otimes V_n$ as a representation of $\mathsf{sl}(2;\mathbb{C})$. Assume $m \geq n$. Then,*

$$V_m \otimes V_n \cong V_{m+n} \oplus V_{m+n-2} \oplus \cdots \oplus V_{m-n+2} \oplus V_{m-n},$$

where $\cong$ denotes an equivalence of $\mathsf{sl}(2;\mathbb{C})$ representations.

Note that this theorem is consistent with the special case worked out earlier: $V_1 \otimes V_1 \cong V_2 \oplus V_0$. Note that each each irreducible representation occurring in the decomposition of $V_m \otimes V_n$ occurs only once. This is a special feature of the theory of $\mathsf{sl}(2;\mathbb{C})$ representations and the analogous statement does *not* hold for tensor products of representations of other Lie algebras.

Proof. Let us take a basis for each of the two spaces that is labeled by the eigenvalues for H. So, we have a basis $u_m, u_{m-2}, \ldots, u_{2-m}, u_{-m}$ for V_m and $v_n, v_{n-2}, \ldots, v_{2-n}, v_{-n}$ for V_n. Then, the vectors of the form $u_k \otimes v_l$ form a basis for $V_m \otimes V_n$, and we compute that

$$[\pi_m(H) \otimes I + I \otimes \pi_n(H)]u_k \otimes v_l = (k+l)u_k \otimes v_l.$$

So, each of our basis elements is an eigenvector for the action of H in $V_m \otimes V_n$. Let us work out the eigenspaces for H in $V_m \otimes V_n$. The eigenspace with

eigenvalue $m+n$ is one dimensional, spanned by $u_m \otimes v_n$. If $n > 0$, then the eigenspace with eigenvalue $m+n-2$ has dimension 2, spanned by $u_{m-2} \otimes v_n$ and $u_m \otimes v_{n-2}$. Each time we decrease the eigenvalue of H by 2 we increase the dimension of the corresponding eigenspace by 1, until we reach the eigenvalue $m-n$, which is spanned by the vectors $u_{m-2n} \otimes v_n$, $u_{m-2n+2} \otimes v_{n-2}$, and so on up to $u_m \otimes v_{-n}$. This space has dimension $n+1$. As we continue to decrease the eigenvalue of H in increments of 2, the dimensions remain constant until we reach eigenvalue $n-m$, at which point the dimensions begin decreasing by 1 until we reach the eigenvalue $-m-n$, for which the corresponding eigenspace has dimension one, spanned by $u_{-m} \otimes v_{-n}$. This pattern is illustrated by the following table, which lists, for the case of $V_4 \otimes V_2$, each eigenvalue for H and a basis for the corresponding eigenspace. I leave it to the reader to verify that this pattern holds true in general.

Eigenvalue for H	**Basis**
6	$u_4 \otimes v_2$
4	$u_2 \otimes v_2$, $u_4 \otimes v_0$
2	$u_0 \otimes v_2$, $u_2 \otimes v_0$, $u_4 \otimes v_{-2}$
0	$u_{-2} \otimes v_2$, $u_0 \otimes v_0$, $u_2 \otimes v_{-2}$
-2	$u_{-4} \otimes v_2$, $u_{-2} \otimes v_0$, $u_0 \otimes v_{-2}$
-4	$u_{-4} \otimes v_0$, $u_{-2} \otimes v_{-2}$
-6	$u_{-4} \otimes v_{-2}$

Now, consider the vector $u_m \otimes v_n$, which is annihilated by X and is an eigenvector for H with eigenvalue $m+n$. Applying Y repeatedly gives a chain of eigenvectors for H with eigenvalues decreasing by 2 until they reach $-m-n$. According to Theorem 4.12, the span of these vectors is invariant under $\mathsf{sl}(2;\mathbb{C})$ and irreducible, isomorphic to V_{m+n}.

The orthogonal complement of the invariant subspace W obtained in the previous paragraph is also invariant. Since W contains each of the eigenvalues of H with multiplicity one, $W^\perp$ will have the multiplicity of each H-eigenvalue lowered by 1. So, $m+n$ is not an eigenvalue for H in $W^\perp$; the largest remaining eigenvalue is $m+n-2$ and this has multiplicity one. So, if we start with an eigenvector for H in $W^\perp$ with eigenvalue $m+n-2$, this will be annihilated by X and will generate an irreducible invariant subspace isomorphic to V_{m+n-2}.

We now continue on in the same way, at each stage looking at the orthogonal complement of the sum of all the invariant subspaces we have obtained in the previous stages. Each step reduces the multiplicity of each H-eigenvalue by 1 and thereby reduces the largest remaining H-eigenvalue by 2. This process will continue until there is nothing left, which will occur after V_{m-n}.

In the case of $V_4 \otimes V_2$, we will get a seven-dimensional invariant subspace isomorphic to V_6, then a five-dimensional invariant subspace isomorphic to V_4, and then a three-dimensional invariant subspace isomorphic to V_2. Actually working out what these invariant subspaces are is complicated, but, in principle, possible. □

D.2 The Wigner–Eckart Theorem

Recall that the Lie algebras $\mathsf{su}(2)$ and $\mathsf{so}(3)$ are isomorphic. Specifically, we use the bases $\{E_1, E_2, E_3\}$ for $\mathsf{su}(2)$ and $\{F_1, F_2, F_3\}$ for $\mathsf{so}(3)$ described in Section 4.9. Then, the unique linear map $\phi : \mathsf{su}(2) \to \mathsf{so}(3)$ such that $\phi(E_k) = F_k$, $k = 1, 2, 3$, is a Lie algebra isomorphism, as a simple calculation will confirm. So, the representations of $\mathsf{so}(3)$ are in one-to-one correspondence with the representations of $\mathsf{su}(2)$, which, in turn, are in one-to-one correspondence with the complex-linear representations of $\mathsf{sl}(2;\mathbb{C})$. So, the analysis of the decomposition of tensor products of $\mathsf{sl}(2;\mathbb{C})$ representations in the previous section applies also to $\mathsf{so}(3)$ representations.

Now, suppose that π is a $\mathsf{so}(3)$ representation acting on a finite-dimensional vector space V. Let $\mathrm{End}(V)$ denote the space of endomorphisms of V (i.e., the space of linear operators of V into itself). Then, we can define an associated representation $\tilde{\pi}$ of $\mathsf{so}(3)$ acting on $\mathrm{End}(V)$ by the formula

$$\tilde{\pi}(X)(C) = [\pi(X), C], \quad X \in \mathsf{so}(3),\ C \in \mathrm{End}(V);$$

that is, $\tilde{\pi}$ is the map $X \to \mathrm{ad}_{\pi(X)}$. Since the maps $X \to \pi(X)$ and $C \to \mathrm{ad}_C$ are Lie algebra homomorphisms, $\tilde{\pi}$ is also a homomorphism and, thus, a representation of $\mathsf{so}(3)$ acting on $\mathrm{End}(V)$. (This is the standard way of "promoting" the action of a Lie algebra on a vector space V to an action on $\mathrm{End}(V)$.)

Recall that the elements of $\mathsf{so}(3)$ are 3×3 real skew-symmetric matrices. These matrices then act on $\mathbb{R}^3$; this is the standard representation of $\mathsf{so}(3)$.

Definition D.2. *Suppose that π is a representation of $\mathsf{so}(3)$ acting on a finite-dimensional space V and $\tilde{\pi}$ is the associated representation of $\mathsf{so}(3)$ acting on $\mathrm{End}(V)$. Then, a linear map $A : \mathbb{R}^3 \to \mathrm{End}(V)$ is called a* ***vector operator*** *(acting on V) if A is an intertwining map of $\mathsf{so}(3)$ representations, that is, if*

$$A(Xv) = [\pi(X), A(v)] \tag{D.2}$$

for all $X \in \mathsf{so}(3)$ and all $v \in \mathbb{R}^3$.

Let us try to understand more concretely what this means. Since A is assumed to be linear, it is determined by its values on the basis elements e_1, e_2, e_3 for $\mathbb{R}^3$. So, let $A_k = A(e_k)$, $k = 1, 2, 3$. In the physics literature, one thinks of A as a "triple of operators" $\mathbf{A} = (A_1, A_2, A_3)$, in which case the linear map $A : \mathbb{R}^3 \to \mathrm{End}(V)$ is the same as

$$\mathbf{A} \cdot v = A_1 v_1 + A_2 v_2 + A_3 v_3.$$

Of course, not every triple of operators gives rise to a vector operator; the map A must satisfy (D.2). It suffices to check (D.2) for X ranging over a basis of $\mathsf{so}(3)$ and v ranging over a basis of $\mathbb{R}^3$. So, using the basis $\{F_1, F_2, F_3\}$ for

$\mathsf{so}(3)$ and the standard basis $\{e_1, e_2, e_3\}$, (D.2) is equivalent (given a linear map $A : \mathbb{R}^3 \to \mathrm{End}(V)$) to the assertion that

$$A(F_k e_l) = [\pi(F_k), A(e_l)] \tag{D.3}$$

for $k = 1, 2, 3$ and $l = 1, 2, 3$.

Now, I claim that the 3×3 matrices F_1, F_2, and F_3 from Section 4.9 satisfy

$$F_k e_l = \sum_{m=1}^{3} \varepsilon_{klm} e_m, \tag{D.4}$$

where ε_{klm} is defined by

$$\varepsilon_{klm} = \begin{cases} 0 \text{ if any two of } k, l, m \text{ are equal} \\ 1 \text{ if } (k, l, m) \text{ is a cyclic permutation of } (1, 2, 3) \\ -1 \text{ if } (k, l, m) \text{ is an non-cyclic permutation of } (1, 2, 3). \end{cases}$$

Concretely, the last two conditions mean that ε_{123}, ε_{231}, and ε_{312} are equal to 1 and ε_{132}, ε_{213}, and ε_{321} are equal to -1. Let us check (D.4), say, in the case $k = l = 1$. We note that $\varepsilon_{11m} = 0$ for all m and, so, (D.4) says that $F_1 e_1 = 0$, which is true. Let us also check (D.4) in the case $k = 1$ and $l = 2$. We note that ε_{12m} is nonzero only when $m = 3$, in which case its value is 1, so (D.4) says that $F_1 e_2 = e_3$, which is also true. I leave it to the reader to check the correctness of (D.4) in the remaining cases.

Proposition D.3. *Suppose π is a representation of $\mathsf{so}(3)$ acting on a space V and define J_k, $k = 1, 2, 3$, by $J_k = \pi(F_k)$. Now, suppose that $A : \mathbb{R}^3 \to \mathrm{End}(V)$ is a linear map and define A_k, $k = 1, 2, 3$, by $A_k = A(e_k)$. Then, A is a vector operator if and only if*

$$[J_k, A_l] = \sum_{m=1}^{3} \varepsilon_{klm} A_m \tag{D.5}$$

for all $k, l \in \{1, 2, 3\}$.

Proof. This is just (D.3) written out (in the reverse order) using the expression (D.4) for F_k and setting $J_k = \pi(F_k)$ and $A_k = A(e_k)$. □

Condition (D.5) differs from the one in the physics literature by a factor of i, reflecting a factor of i difference between the mathematics and physics literatures regarding the definition of the Lie algebra $\mathsf{so}(3)$.

Proposition D.4. *Suppose that π is any finite-dimensional representation of $\mathsf{so}(3)$ acting on a space V. Let J the unique linear map from $\mathbb{R}^3$ into $\mathrm{End}(V)$ satisfying $J(e_k) = \pi(F_k)$, $k = 1, 2, 3$. Then, J is a vector operator.*

That is to say, if we identify $\mathbb{R}^3$ with $\mathsf{so}(3)$ by identifying the basis e_1, e_2, e_3 for $\mathbb{R}^3$ with the basis F_1, F_2, F_3 for $\mathsf{so}(3)$, then π itself is a vector operator; in physics notation, $\mathbf{J} = (J_1, J_2, J_3)$ is a vector operator.

Proof. The commutation relations among the F_k's (computed in Section 4.9) may be expressed as

$$[F_k, F_l] = \sum_{m=1}^{3} \varepsilon_{klm} F_m. \tag{D.6}$$

Since π is a representation, the J_k's (defined as $\pi(F_k)$) have the same commutation relations as the F_k's, $[J_k, J_l] = \sum_{m=1}^{3} \varepsilon_{klm} J_m$. This means that (D.5) is satisfied if $A_k = J_k$, which shows that J is a vector operator. □

Let us now assume that we have an inner product $\langle \cdot, \cdot \rangle$ on V that is invariant under the action of $\mathsf{so}(3)$, in the sense that $\langle \pi(X)u, v \rangle = -\langle u, \pi(X)v \rangle$ for all $X \in \mathsf{so}(3)$ and $u, v \in V$. We call this a "unitary" representation of $\mathsf{so}(3)$. (After all, $\mathsf{so}(3)$ is mapping into $\mathsf{u}(V)$, the skew self-adjoint operators on V, which is the Lie algebra of the group $\mathsf{U}(V)$ of unitary operators on V.)

Theorem D.5 (Wigner–Eckart). *Suppose that V is a finite-dimensional unitary representation of $\mathsf{so}(3)$ and that $A, B : \mathbb{R}^3 \to \mathrm{End}(V)$ are vector operators. Set $A_k = A(e_k)$ and $B_k = B(F_k)$, $k = 1, 2, 3$. Suppose that W_1 and W_2 are irreducible $\mathsf{so}(3)$-invariant subspaces and suppose that $\langle w, A_k w' \rangle$ is nonzero for some $w \in W_1$, $w' \in W_2$, and $k \in \{1, 2, 3\}$. Then, there exists a constant c such that*

$$\langle w, B_k w' \rangle = c \langle w, A_k w' \rangle$$

for all $w \in W_1$, $w' \in W_2$, and $k \in \{1, 2, 3\}$.

The constant c depends on the choice of W_1, W_2, A, and B, but is independent of w, w', and k.

In application in physics, V is usually infinite dimensional (but W_1 and W_2 are still finite dimensional). The result still holds in that case, subject to certain technical conditions.

Suppose that $W_1 \cong V_m$ and $W_2 \cong V_n$. There is then a standard way of choosing a basis $u_m, u_{m-2}, \ldots, u_{2-m}, u_{-m}$ for W_1 and $v_n, v_{n-2}, \ldots, v_{2-n}, v_{-n}$ for W_2. The Wigner–Eckart theorem says that the matrix entries $\langle u_l, B_k v_{l'} \rangle$ are determined up to a constant merely by the fact that B is a vector operator. There exist certain "universal" coefficients $\alpha(m, n, k, l, l')$, given in terms of the Clebsch–Gordan coefficients, that can be computed once and for all and that capture these matrix entries up to a constant. So, the Wigner–Eckart theorem can be expressed as saying that if B is a vector operator acting on V, and W_1 and W_2 are irreducible $\mathsf{so}(3)$-invariant subspaces isomorphic to V_m and V_n, respectively, then

$$\langle u_l, B_k u_{l'} \rangle = c\, \alpha(m, n, k, l, l').$$

Note that the subspaces W_1 and W_2 are assumed invariant under the action of $\mathsf{so}(3)$, but are *not* necessarily invariant under the A_k's or B_k's.

Before turning to the proof of the Wigner–Eckart Theorem, we consider a useful isomorphism. Suppose now π_1 and π_2 are representations of some Lie

algebra $\mathfrak{g}$ acting on finite-dimensional vector space W_1 and W_2, respectively. Let $\mathrm{Hom}(W_2, W_1)$ denote the space of linear maps of W_2 into W_1. Define a representation of $\mathfrak{g}$ acting on $\mathrm{Hom}(W_2, W_1)$ as follows. If $X \in \mathfrak{g}$ and $C \in \mathrm{Hom}(W_2, W_1)$, then let the action of X on C be given by

$$\pi_1(X)C - C\pi_2(X). \tag{D.7}$$

It is easily verified that this formula does, indeed, make $\mathrm{Hom}(W_2, W_1)$ into a representation of $\mathfrak{g}$.

Lemma D.6. *Suppose that π_1 and π_2 are representations of a Lie algebra $\mathfrak{g}$ acting on finite-dimensional spaces W_1 and W_2, respectively. Define an action of $\mathfrak{g}$ on $\mathrm{Hom}(W_2, W_1)$ by (D.7). Then,*

$$\mathrm{Hom}(W_2, W_1) \cong W_2^* \otimes W_1,$$

where $\cong$ denotes equivalence of representations of $\mathfrak{g}$.

Proof. Consider the unique linear map $\Psi : W_2^* \otimes W_1 \to \mathrm{Hom}(W_2, W_1)$ such that for all $\phi \in W_2^*$ and $u \in W_1$,

$$\Psi(\phi \otimes u)(v) = \phi(v)u.$$

(The universal property of the tensor product guarantees that there is, in fact, a unique such map.) It is easily verified, using bases for W_1 and W_2, that Ψ is an invertible linear map. We now verify that Ψ is an intertwining map for the actions of $\mathfrak{g}$ on $W_2^* \otimes W_1$ and on $\mathrm{Hom}(W_2, W_1)$. Consider elements ϕ in W_2^* and u in W_1. Suppose we first let an element X of $\mathfrak{g}$ act on $\phi \otimes u$. Recalling from Chapter 4 the way one takes duals and tensor products of representations, we get

$$-(\phi \circ \pi_2(X)) \otimes u + \phi \otimes (\pi_1(X)u).$$

If we then apply the map Ψ to this element and apply the resulting operator to a vector $v \in W_2$, we get

$$-\phi(\pi_2(X)v)u + \phi(v)\pi_1(X)u. \tag{D.8}$$

Suppose, on the other hand, that we *first* apply Ψ to $\phi \otimes u$, *then* let X act on the resulting operator by means of (D.7), and then apply the result to the vector v. A simple calculation shows that the result is, again, the quantity in (D.8). Thus, Ψ intertwines the actions of $\mathfrak{g}$ on elements of the form $\phi \otimes u$. Since every element of $W_2^* \otimes W_1$ is a linear combination of elements of this form, we conclude that Ψ is an intertwining map. □

We now turn to the proof of the Wigner–Eckart Theorem. The ingredients are Schur's Lemma, Lemma D.6, and Theorem D.1. The part of Theorem D.1 that we need is that each irreducible representation occurring in the decomposition of $V_m \otimes V_n$ occurs only once.

Proof. Let P_1 denote the orthogonal projection operator from V onto W_1. Since the action of $\mathsf{so}(3)$ on V is unitary and W_1 is an invariant subspace, it is not hard to show that P_1 commutes with the action of $\mathsf{so}(3)$ (i.e., that P_1 commutes with J_k, $k = 1, 2, 3$). Now, define a linear map $\tilde{A}_k : W_2 \to W_1$ by the formula

$$\tilde{A}_k w' = P_1 A_k w', \quad w' \in W_1.$$

Since A is a vector operator and P_1 commutes with each J_k, it is not hard to show that the operators $\tilde{A}_k$ satisfy $[J_k, \tilde{A}_l] = \sum_{m=1}^{3} \varepsilon_{klm} \tilde{A}_m$. This means that $\tilde{A}$ is a "vector operator from W_2 to W_1." Specifically, let $\tilde{A}$ be the unique linear map from $\mathbb{R}^3$ into $\mathrm{Hom}(W_2, W_1)$ such that $\tilde{A}(e_k) = \tilde{A}_k$, $k = 1, 2, 3$. Then, $\tilde{A}$ is an intertwining map for the action of $\mathsf{so}(3)$. We can extend $\tilde{A}$ by complex linearity to an intertwining map of $\mathbb{C}^3$ into $\mathrm{Hom}(W_2, W_1)$, where $\mathbb{C}^3$ (the complexification of the standard representation) is a three-dimensional irreducible representation of $\mathsf{so}(3)$, hence isomorphic to V_2.

Now, W_2^* and W_1 are irreducible representations of $\mathsf{so}(3)$, and so they are isomorphic to V_m and V_n, respectively, for some positive integers m and n. Thus, by Lemma D.6, $\mathrm{Hom}(W_2, W_1)$ is isomorphic to $V_m \otimes V_n$ as a representation of $\mathsf{so}(3)$. Then, by Theorem D.1, $\mathrm{Hom}(W_2, W_1) \cong V_m \otimes V_n$ decomposes as a direct sum of irreducible representations of $\mathsf{so}(3)$ *in such a way that the three-dimensional representation $V_2 \cong C^3$ occurs at most once.* If V_2 does not appear in the decomposition of $V_m \otimes V_n$, then by Schur's Lemma, $\tilde{A}$ must be the zero map. If V_2 does occur in the decomposition, then by Schur's Lemma, $\tilde{A}$ is determined up to a constant by the fact that it is an intertwining map. Thus, if $\tilde{A} \neq 0$ and if $\tilde{B}$ is defined by analogy to $\tilde{A}$, then $\tilde{B}$ must be a constant multiple of $\tilde{A}$; that is, if $\tilde{A} \neq 0$, then there is a constant c such that $\tilde{B}_k = c\tilde{A}_k$, $k = 1, 2, 3$.

We are now almost done. After all, the orthogonal projection operator $P_1 : V \to W_1$ is self-adjoint, as is easily verified. Thus, for all $w \in W_1$ and $w' \in W_2$,

$$\langle w, \tilde{A}_k w' \rangle = \langle w, P_1 A_k w' \rangle = \langle P_1 w, A_k w' \rangle = \langle w, A_k w' \rangle, \tag{D.9}$$

and similarly for $\tilde{B}_k$. Therefore, if $\langle w, A_k w' \rangle$ is non-zero for some w and w', then $\tilde{A} \neq 0$ and, so, $\tilde{B}_k = c\tilde{A}_k$. This, by (D.9), implies the Wigner–Eckart Theorem. □

D.3 More on Vector Operators

Let us now look more closely at the concept of a vector operator. Suppose at first that associated to the representation π of the Lie algebra $\mathsf{so}(3)$ there is a representation Π of the group $\mathsf{SO}(3)$. (Since $\mathsf{SO}(3)$ is not simply connected, this will not always be the case.) Then, there is a representation $\tilde{\Pi}$ of $\mathsf{SO}(3)$ associated to representation $\tilde{\pi}$ of $\mathsf{so}(3)$ acting on $\mathrm{End}(V)$, namely

$$\tilde{\Pi}(R)(A) = \Pi(R)A\Pi(R)^{-1}, \quad R \in \mathsf{SO}(3),\ A \in \mathrm{End}(V);$$

that is, $\tilde{\Pi}(R) = \mathrm{Ad}_{\Pi(R)}$; this is simply the group counterpart to expression for $\tilde{\pi}$, $\tilde{\pi}(X) = \mathrm{ad}_{\pi(X)}$. If $\psi : \mathbb{R}^3 \to \mathrm{End}(V)$ is a linear map, then ψ is an intertwining map for the $\mathsf{so}(3)$ actions on $\mathbb{R}^3$ and $\mathrm{End}(V)$ if and only if it is an intertwining map for the associated $\mathsf{SO}(3)$ actions. Thus, ψ is a vector operator if and only if

$$\psi(Rv) = \Pi(R)\psi(v)\Pi(R)^{-1} \tag{D.10}$$

for all $R \in \mathsf{SO}(3)$ and $v \in \mathbb{R}^3$. This condition is, in physics terminology, the "finite" counterpart to the "infinitesimal" condition (D.5).

In quantum mechanics, one considers a quantum Hilbert space V equipped with "angular momentum" operators J_1, J_2, and J_3 which satisfy the $\mathsf{so}(3)$ commutation relations and thus give rise to a representation π of $\mathsf{so}(3)$. (In most cases, V is infinite dimensional, but this should be regarded as a technicality; subject to certain technical conditions, the Wigner–Eckart theorem remains true if V is infinite dimensional, provided that $W \subset V$ is finite dimensional and irreducible.) Assuming there is an associated representation Π of $\mathsf{SO}(3)$ on V, an operator of the form $\Pi(R)$ represents the action of a rotation R of $\mathbb{R}^3$ on the space of quantum states. One also has "position" operators X_1, X_2, and X_3 that describe (in the quantum realm) the x_1-, x_2-, and x_3-components of the position of a particle. One also expects quantum mechanics to be a rotationally invariant theory, which means that the x_1-component of the position should not be fundamentally different from the x_2-component of the position. What that means in this context is that if R is a rotation that takes e_1 to e_2, then we should have

$$X_2 = \Pi(R)X_1\Pi(R)^{-1}.$$

This says that X_2 differs from X_1 simply by the action $\tilde{\Pi}(R)$ on $\mathrm{End}(V)$. More generally, if v is any unit vector in $\mathbb{R}^3$, the operator corresponding to the v-component of the position will be $\mathbf{X} \cdot v = X_1v_1 + X_2v_2 + X_3v_3$ and we expect that

$$\mathbf{X} \cdot (Rv) = \Pi(R)(\mathbf{X} \cdot v)\Pi(R)^{-1};$$

that is, we expect that $\mathbf{X}$ should be a vector operator. Other common vector operators in quantum mechanics are the angular momentum operators themselves and the momentum operators.

Now, the discussion in the two previous paragraphs assumes that there is a representation Π of $\mathsf{SO}(3)$ associated to the representation π of $\mathsf{so}(3)$. In practice, this may not be the case, for example, when dealing with particles with spin $\frac{1}{2}$. However, we can use the isomorphism $\phi : \mathsf{su}(2) \to \mathsf{so}(3)$ (Section 4.9) to turn π into a representation σ of $\mathsf{su}(2)$. Then, because $\mathsf{SU}(2)$ is simply connected, we can form a $\mathsf{SU}(2)$ representation Σ. Assume that Σ has the property that

$$\Sigma(-I) = \pm I. \tag{D.11}$$

If Σ is irreducible, then by Schur's Lemma, $\Sigma(-I)$ must be a multiple of the identity, and since $(-I)^2 = I$, this multiple must be $\pm I$. So, (D.11) automatically holds when Σ (or equivalently π) is irreducible. However, in practice, we do not want to assume that π is irreducible and, so, (D.11) is a nontrivial restriction.

If (D.11) holds, then we can "define" a representation Π of $\mathsf{SO}(3)$ by setting

$$\Pi(R) = \Sigma(U), \tag{D.12}$$

where $U \in \mathsf{SU}(2)$ is chosen so that $\Phi(U) = R$. (Here $\Phi : \mathsf{SU}(2) \to \mathsf{SO}(3)$ is the 2-to-1 group homomorphism corresponding to the Lie algebra homomorphism ϕ.) For any $R \in \mathsf{SO}(3)$, such a U exists and is unique up to a sign, and, so, (D.12) serves to define $\Pi(R)$ up to a sign. We think of $\Pi(R)$ as actually representing either of the two operators $\Sigma(U)$ and $\Sigma(-U)$. If $\Sigma(-I) = -I$, then there is no consistent way to choose the signs in Π and, so, Π is a "double-valued representation" of $\mathsf{SO}(3)$. However, even if Π is double-valued, the associated representation $\tilde{\Pi}$ is single-valued, since

$$\Pi(R)A\Pi(R)^{-1} = (-\Pi(R))A(-\Pi(R))^{-1}.$$

So, the "global" form (D.10) of the definition of a vector operator makes sense provided that (D.11) holds, even if Π does not exist as a single-valued representation of $\mathsf{SO}(3)$.

A "representation" of a group that is defined only up to constant multiples of the identity is what is called a **projective representation**. More precisely, a projective representation of G is a continuous homomorphism of G into the quotient group $\mathsf{GL}(V)/\mathbb{C}^*I$, where $\mathbb{C}^*I$ is the group of all nonzero constant multiples of the identity, which is a normal subgroup of $\mathsf{GL}(V)$. In quantum mechanics, a constant multiple of a vector $v \in V$ represents the same physical state as v itself, so a projective representation of $\mathsf{SO}(3)$ still gives a well-defined action of $\mathsf{SO}(3)$ on the set of physical states.

E

Computing Fundamental Groups of Matrix Lie Groups

E.1 The Fundamental Group

Let X be any path-connected Hausdorff topological space and let x_0 be a fixed point in X (the "basepoint"). We consider **loops in X based at x_0** (i.e., continuous maps $\gamma : [0,1] \to X$ with the property that $\gamma(0) = \gamma(1) = x$). The choice of the basepoint makes no substantive difference to the constructions that follow. From now on, "loop" will mean "loop based at x_0." Ultimately, we are interested in the case that X is a matrix Lie group.

If γ_1 and γ_2 are two loops, then we define the **concatenation** of γ_1 and γ_2 to be the loop $\gamma_1 \cdot \gamma_2$ given by

$$\gamma_1 \cdot \gamma_2(t) = \begin{cases} \gamma_1(2t), & 0 \leq t \leq \frac{1}{2} \\ \gamma_2(2t-1), & \frac{1}{2} \leq t \leq 1; \end{cases}$$

that is, $\gamma_1 \cdot \gamma_2$ traverses γ_1 as t goes from 0 to $1/2$ and then traverses γ_2 as t goes from $1/2$ to 1. (At $t = 1/2$, we are at the basepoint.)

Two loops γ_1 and γ_2 are said to be **homotopic** if one can be "continuously deformed" into the other. More precisely, this means that there exists a continuous map $A : [0,1] \times [0,1] \to G$ such that $A(0,t) = \gamma_1(t)$ and $A(1,t) = \gamma_2(t)$ for all $t \in [0,1]$ and such that $A(s,0) = A(s,1) = x_0$ for all $s \in [0,1]$. One should think of $A(s,t)$ as a family of loops parameterized by s. One important example of homotopic loops is the case when γ_2 is simply a reparameterization of γ_1 with the same orientation. Suppose $f : [0,1] \to [0,1]$ is a continuous, nondecreasing function with $f(0) = 0$ and $f(1) = 1$. Then, for any loop γ_1, the loop γ_2 given by $\gamma_2(t) = \gamma_1(f(t))$ will be homotopic to γ_1.

A loop is said to be **null homotopic** if it is homotopic to the constant loop (i.e., the loop γ^0 for which $\gamma^0(t) = x_0$ for all $t \in [0,1]$). If all loops in X are null homotopic, then X is said to be **simply connected**. (This is not quite identical to but is equivalent to the definition given in Section 1.5.) The notion of homotopy is an equivalence relation. This means (1) that every loop is homotopic to itself, (2) that γ_1 is homotopic to γ_2 if and only if γ_2 is

homotopic to γ_1, and (3) that if γ_1 is homotopic to γ_2 and γ_2 is homotopic to γ_3, then γ_1 is homotopic to γ_3.

The **homotopy class** of a loop γ is the set of all loops that are homotopic to γ. Each loop belongs to one and only one equivalence class. The concatenation operation "respects homotopy." This means that if γ_1 is homotopic to γ_2 and δ_1 is homotopic to δ_2, then $\gamma_1 \cdot \delta_1$ is homotopic to $\gamma_2 \cdot \delta_2$. As a result, it makes sense to define the concatenation operation on equivalence classes.

The **fundamental group** of X, denoted $\pi_1(X)$, is the set of homotopy classes equipped with the operation of concatenation. We need to check that this operation indeed makes the set of homotopy classes into a group. So, we must check associativity, the existence of an identity, and the existence of inverses. For associativity, we note that although $(\gamma_1 \cdot \gamma_2) \cdot \gamma_3$ is *not* the same as $\gamma_1 \cdot (\gamma_2 \cdot \gamma_3)$, the second of these two loops is a reparameterization of the first. Thus, the homotopy class of $(\gamma_1 \cdot \gamma_2) \cdot \gamma_3$ *is* the same as the homotopy class of $\gamma_1 \cdot (\gamma_2 \cdot \gamma_3)$ and, so, concatenation is associative at the level of homotopy classes. The identity in the fundamental group is the constant loop γ^0. This is not an identity at the level of loops but is at the level of homotopy classes; that is, $\gamma \cdot \gamma^0$ and $\gamma^0 \cdot \gamma$ are not equal to γ, but they are both homotopic to γ, since both are reparameterizations of γ. Finally, for inverses, the inverse to a homotopy class $[\gamma]$ is the homotopy class $[\gamma']$ where $\gamma'(t) = \gamma(1-t)$. It can be shown that $\gamma \cdot \gamma'$ and $\gamma' \cdot \gamma$ are both null homotopic. A topological space X is simply connected precisely if its fundamental group is the trivial group (consisting of just the identity element).

Some standard examples of fundamental groups are as follows: $\mathbb{R}^n$ is simply connected for all n, S^n is simply connected for $n \geq 2$, and the fundamental group of S^1 is isomorphic to $\mathbb{Z}$. For more information on fundamental groups and related topics, see Munkres (1975) or Hatcher (2002).

E.2 The Universal Cover

If X is a sufficiently nice topological space (the precise conditions need not concern us here except to say that a connected matrix Lie group is nice enough), then one can construct something called a **universal cover** of X. This is a simply-connected topological space $\tilde{X}$ together with a continuous map $P : \tilde{X} \to X$ with the following property: Each $x \in X$ has a neighborhood U such that $P^{-1}(U)$ is a disjoint union of open sets V_α such that P maps each V_α homeomorphically onto U. The map P is called the **projection map**. If we think of P as projecting $\tilde{X}$ "down" onto X, then $P^{-1}(U)$ is the set of points in $\tilde{X}$ lying above U. This set consists of a family of disjoint open sets V_α, each of which is identical to (i.e., homeomorphic to) U.

If X is a Lie group G, then the universal cover $\tilde{G}$ can be given the structure of a Lie group in such a way that the projection map $P : \tilde{G} \to G$ is a smooth homomorphism. In that case $\tilde{G}$ is called a **universal covering group** of G and the associated Lie algebra map $p : \tilde{\mathfrak{g}} \to \mathfrak{g}$ is a Lie algebra isomorphism. We

have the following uniqueness result for universal covering groups. Suppose that H_1 and H_2 are simply-connected Lie groups and $P_1 : H_1 \to G$ and $P_2 : H_2 \to G$ are Lie group homomorphisms such that the associated Lie algebra maps $p_1 : \mathfrak{h}_1 \to \mathfrak{g}$ and $p_2 : \mathfrak{h}_2 \to \mathfrak{g}$ are isomorphisms. Then, there exists a Lie group isomorphism $\Phi : H_1 \to H_2$ such that $P_1(h) = P_2(\Phi(h))$ for all h in H_1. With this uniqueness result in mind (compare Corollary 3.8), we speak of *the* universal covering group of G.

If $(\tilde{G}, P)$ is the universal covering group of G, then the fundamental group of G has the property that

$$\pi_1(G) \cong \ker P.$$

So, if one can find the universal covering group and associated projection map explicitly, then this allows one to compute the fundamental group of G. (Usually, there is no direct way of producing the universal cover and so one has to compute $\pi_1(G)$ by other means.) The kernel of any homomorphism is a closed normal subgroup. Since the Lie algebra map associated to P is an isomorphism, the kernel of P must be discrete and, therefore, a subgroup of the center of $\tilde{G}$ (Exercise 11 from Chapter 1). This shows that $\pi_1(G)$ is commutative for any Lie group G. (This can also be shown directly.) For general topological spaces, the fundamental group may be noncommutative.

Even if G is a matrix Lie group, $\tilde{G}$ may not be a matrix Lie group, as the example $G = \mathsf{SL}(n;\mathbb{R})$ in Appendix C demonstrates. Even if $\tilde{G}$ does happen to be a matrix Lie group, there is no canonical way to represent $\tilde{G}$ as a matrix Lie group. This shows the advantage of working with general Lie groups instead of just matrix Lie groups. On the other hand, if G is a (connected) matrix Lie group and one can somehow find a simply-connected matrix Lie group H and a homomorphism $P : H \to G$ such that the associated Lie algebra map is an isomorphism, then (H, P) is *the* universal covering group of G and one never needs to concern oneself with nonmatrix Lie groups. For example, if G is $\mathsf{SO}(3)$, then we can take H to be $\mathsf{SU}(2)$ and P to be the map given in Section 1.6, whose kernel is $\{I, -I\}$. This shows that $\pi_1(\mathsf{SO}(3))$ is isomorphic to $\mathbb{Z}/2$.

E.3 Fundamental Groups of Compact Lie Groups I

For any nice topological space, one can define **higher homotopy groups** $\pi_k(X)$, $k = 1, 2, 3, \ldots$. The precise definition need not concern us here. The relevant points are that $\pi_1(X)$ is the fundamental group as defined in the previous sections and that $\pi_k(X)$ is trivial (i.e., contains only the identity) if and only if every continuous map of the k-sphere S^k into X can be shrunk continuously to a point. We will make use of the following standard topological result:

Proposition E.1. *For a d-sphere S^d, $\pi_k(S^d)$ is trivial if $k < d$.*

This result is plausible because for $k < d$, the image of a "typical" continuous map of S^k into S^d will not be all of S^d. However, if the image of the map omits even one point in S^d, then we can remove that point and what is left of the sphere can be contracted continuously to a point.

Definition E.2. *Suppose that B and F are Hausdorff topological spaces. A **fiber bundle** with base B and fiber F is a Hausdorff topological space X together with a continuous map $p : X \to B$, called the **projection map**, having the following properties. First, for each b in B, the preimage of $p^{-1}(b)$ of b in X is homeomorphic to F. Second, given any b in B, there is a neighborhood U of b such that $p^{-1}(U)$ is homeomorphic to $U \times F$ in such a way that the projection map is simply projection onto the first factor.*

The second condition can be stated more pedantically as follows. For each $b \in B$, there should exist a neighborhood U of B and a homeomorphism Φ of $p^{-1}(U)$ with $U \times F$ having the property that $p(x) = p_1(\Phi(x))$, where $p_1 : U \times F \to U$ is the map $p_1(u, f) = u$. The sets of the form $p^{-1}(b)$ are called the **fibers** of the fiber bundle.

The simplest sort of fiber bundle is the product space $X = B \times F$, with the projection map being simply the projection onto the first factor. Such a fiber bundle is called **trivial**. The second condition in the definition of a fiber bundle is called **local triviality** and it says that any fiber bundle must look locally like a trivial bundle. In general, X need not be globally homeomorphic to $B \times F$.

If X were a trivial fiber bundle, then the fundamental group of X would be simply the product of the fundamental group of the base B and the fundamental group of the fiber F. In particular, if X were a trivial fiber bundle and $\pi_1(B)$ were trivial, then $\pi_1(X)$ would be isomorphic to $\pi_1(F)$. The following result says that if $\pi_1(B)$ and $\pi_2(B)$ are trivial, then the same conclusion holds even if X is nontrivial.

Theorem E.3. *Suppose that X is a fiber bundle with base B and fiber F. If $\pi_1(B)$ and $\pi_2(B)$ are trivial (i.e., contain only the identity), then $\pi_1(X)$ is isomorphic to $\pi_1(F)$.*

This result is a consequence of the long exact sequence of homotopy groups for fiber bundles.

Theorem E.4. *Suppose G acts smoothly and transitively on a smooth manifold M and let m_0 be an arbitrary point in M. Define $p : G \to M$ by $p(g) = g \cdot m_0$. Let H be the stabilizer of the point m_0—that is,*

$$H = \{ g \in G | \, g \cdot m_0 = m_0 \} .$$

Then, G is a fiber bundle with base M and fiber H.

This is a standard result in Lie group theory. See, for example, Warner (1983).

Example 1: $\mathsf{SO}(n)$ *acting on the unit sphere in* $\mathbb{R}^n$. We think of the $(n-1)$-dimensional sphere S^{n-1} as the set of vectors of length 1 inside $\mathbb{R}^n$. If $\|x\|^2 = \langle x, x\rangle = 1$, then $\|Rx\|^2 = \langle Rx, Rx\rangle = 1$ for all R in $\mathsf{SO}(n)$. Thus, $\mathsf{SO}(n)$ acts on S^{n-1} and the action is easily seen to be smooth. It it not hard to show that $\mathsf{SO}(n)$ acts transitively on S^{n-1}, for all $n \geq 2$. Let us take as our basepoint the vector $e_1 = (1, 0, \ldots, 0)$ in S^{n-1}. If $R \in \mathsf{SO}(n)$ and $Re_1 = e_1$, then R must map the orthogonal complement of e_1, namely the span of $e_2, \ldots, e_n$, into itself. So, R must be of the form

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

with $R' \in \mathsf{SO}(n-1)$. Thus, the stabilizer of e_1 is (isomorphic to) $\mathsf{SO}(n-1)$.

Thus, by Theorem E.4, $\mathsf{SO}(n)$ is a fiber bundle with base S^{n-1}, fiber $\mathsf{SO}(n-1)$, and projection map given by $R \to Re_1$. Suppose that n is at least 4, so that $n-1$ is at least 3. Then, by Proposition E.1, $\pi_1(S^{n-1})$ and $\pi_2(S^{n-1})$ are trivial and, so, Theorem E.3 tells us that $\pi_1(\mathsf{SO}(n))$ is isomorphic to $\pi_1(\mathsf{SO}(n-1))$, for $n \geq 4$. Thus, $\pi_1(\mathsf{SO}(n))$ is isomorphic to $\pi_1(\mathsf{SO}(3))$ for all $n \geq 4$.

This method does not tell us what $\pi_1(\mathsf{SO}(3))$ is, but this can be computed "by hand" by showing that $\mathsf{SO}(3)$ is homeomorphic to $\mathbb{RP}^3$ (Section 1.5), by showing that $\mathsf{SO}(3) \cong \mathsf{SU}(2)/\{I, -I\}$, or by the method of the next section. In any case the conclusion is that $\pi_1(\mathsf{SO}(3)) \cong \mathbb{Z}/2$. We thus conclude that $\pi_1(\mathsf{SO}(n)) \cong \mathbb{Z}/2$ for all $n \geq 3$.

Meanwhile, if $n = 3$, then $\pi_2(S^{n-1})$ is not trivial and, so, Theorem E.3 does *not* tell us that $\pi_1(\mathsf{SO}(2))$ is the same as $\pi_1(\mathsf{SO}(3))$. Indeed, $\mathsf{SO}(2)$ is diffeomorphic to S^1 and, so, $\pi_1(\mathsf{SO}(2)) \cong \mathbb{Z}$.

Example 2: $\mathsf{SU}(n)$ *acting on unit sphere in* $\mathbb{C}^n$. The group $\mathsf{SU}(n)$ acts on the unit sphere S^{2n-1} in $\mathbb{C}^n = \mathbb{R}^{2n}$. The action is transitive for all $n \geq 2$ and the stabilizer of a point is $\mathsf{SU}(n-1)$. So, $\mathsf{SU}(n)$ is a fiber bundle with base S^{2n-1} and fiber $\mathsf{SU}(n-1)$. As long as n is at least 2, then the sphere in question will have dimension at least 3 and, so, $\pi_1(S^{2n-1})$ and $\pi_2(S^{2n-1})$ will be trivial. We conclude, therefore, that $\pi_1(\mathsf{SU}(n)) \cong \pi_1(\mathsf{SU}(n-1))$ for all $n \geq 2$. In particular, $\pi_1(\mathsf{SU}(2))$ is isomorphic to $\pi_1(\mathsf{SU}(1))$, and since $\mathsf{SU}(1)$ is just a point, its fundamental group is trivial. So, we conclude (as we already knew) that $\mathsf{SU}(2)$ is simply connected. Applying this argument then for $n = 3, 4, \ldots$ shows that $\mathsf{SU}(n)$ is simply connected for all n.

Example 3: $\mathsf{U}(n)$ *acting on unit sphere in* $\mathbb{C}^n$. The group $\mathsf{U}(n)$ also acts on S^{2n-1}. The same argument as for $\mathsf{SU}(n)$ shows that all the $\mathsf{U}(n)$'s have the same fundamental group. However, $\mathsf{U}(1)$ is diffeomorphic to S^1 and, so, $\pi_1(\mathsf{U}(1)) \cong \mathbb{Z}$. Thus, $\pi_1(\mathsf{U}(n)) \cong \mathbb{Z}$ for all n.

Example 4: $\mathsf{Sp}(n)$ acting on unit sphere in $\mathbb{C}^{2n}$. The group $\mathsf{Sp}(n) := \mathsf{Sp}(n;\mathbb{C}) \cap \mathsf{U}(2n)$ is contained in $\mathsf{U}(2n)$ and therefore preserves the unit sphere S^{4n-1} in $\mathbb{C}^{2n}$. In fact, it acts transitively on this sphere, and the stabilizer of a point is isomorphic to $\mathsf{Sp}(n-1)$. Already for $n = 2$, this sphere has dimension 7, so we conclude that $\pi_1(\mathsf{Sp}(n)) \cong \pi_1(\mathsf{Sp}(n-1))$ for all $n \geq 2$. However, $\mathsf{Sp}(1;\mathbb{C}) = \mathsf{SL}(2;\mathbb{C})$ and so $\mathsf{Sp}(1) = \mathsf{SL}(2;\mathbb{C}) \cap \mathsf{U}(2) = \mathsf{SU}(2)$, which is simply connected. Thus, $\mathsf{Sp}(n)$ is simply connected for all $n \geq 1$.

These examples complete the table of fundamental groups of compact groups given in Chapter 1.

E.4 Fundamental Groups of Compact Lie Groups II

We consider now a more algebraic approach to computing $\pi_1(K)$, where K is a compact connected Lie group. We also discuss in this section a method of computing the center of K, which is not really relevant to the computation of π_1 but which uses the same constructions.

Let K be a connected compact Lie group with Lie algebra $\mathfrak{k}$. Although we will use the machinery of roots and weights, it is not necessary here to assume that $\mathfrak{k}$ is semisimple. (Recall that the Lie algebra of a compact group is automatically reductive (i.e., the direct sum of a semisimple Lie algebra and a commutative Lie algebra).) We choose once and for all an inner product on $\mathfrak{k}$ that is invariant under the adjoint action of K and a maximal commutative subalgebra $\mathfrak{t}$ of $\mathfrak{k}$. Then, let T be the connected Lie subgroup of K whose Lie algebra is $\mathfrak{t}$. Then T is a closed subgroup of K and is called a **maximal torus** in K. It is customary in the setting of compact groups to work with the **real roots** (as in Section 7.4). The real roots are the nonzero elements α of $\mathfrak{t}$ for which there exists a nonzero X in $\mathfrak{k}_{\mathbb{C}}$ such that

$$[H, X] = i\langle \alpha, H \rangle X$$

for all H in $\mathfrak{t}$. Although $\mathfrak{k}$ (or equivalently $\mathfrak{k}_{\mathbb{C}}$) is not assumed semisimple, all of the properties of roots from the semisimple case hold *except* that the roots do not necessarily span $\mathfrak{t}$.

For each root α, we have the root spaces $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ inside the complexified Lie algebra $\mathfrak{k}_{\mathbb{C}}$. These are both one dimensional and they, together with their commutator, span a three-dimensional subalgebra $\mathfrak{s}^\alpha$ of $\mathfrak{k}_{\mathbb{C}}$ that is isomorphic to $\mathsf{sl}(2;\mathbb{C})$. An examination of the proof of this result (Theorem 6.20) will show that the intersection of $\mathfrak{s}^\alpha$ with the real Lie algebra $\mathfrak{k}$ is a three-dimensional subalgebra $\mathfrak{s}^\alpha_{\mathbb{R}}$ of $\mathfrak{k}$ that is isomorphic to $\mathsf{su}(2)$. Specifically, if X_α, Y_α, and H_α are as in Theorem 6.20, then we set

$$J_\alpha = \frac{1}{i} H_\alpha,$$
$$K_\alpha = \frac{1}{2}(X_\alpha - Y_\alpha),$$
$$L_\alpha = \frac{i}{2}(X_\alpha + Y_\alpha).$$

Then, $J_\alpha \in \mathfrak{t}$ and K_α and L_α are in $\mathfrak{k}$ and they have the usual $\mathsf{su}(2)$ commutation relations: $[J_\alpha, K_\alpha] = L_\alpha$, $[K_\alpha, L_\alpha] = J_\alpha$, and $[L_\alpha, J_\alpha] = K_\alpha$. We will call the element J_α the **real co-root** corresponding to the root α. Then, the real roots and the real co-roots are related in the same way as the roots and co-roots in Chapter 6, namely

$$J_\alpha = 2\frac{\alpha}{\langle \alpha, \alpha \rangle}. \tag{E.1}$$

Definition E.5. *Let Γ denote the set of all H in $\mathfrak{t}$ for which $\exp(2\pi H) = I$. We call Γ the **kernel of the exponential mapping** (for $\mathfrak{t}$). Let J denote the set of all linear combinations of the real co-roots J_α with integer coefficients.*

It is a slight abuse of terminology to call Γ the kernel of the exponential mapping, since the map it is the kernel of is actually the exponential mapping composed with multiplication by 2π in $\mathfrak{t}$.

Proposition E.6. *If α is a real root and J_α the associated real co-root, then $\exp(2\pi J_\alpha) = I$. Thus, $J \subset \Gamma$.*

Proof. The proposition holds in the $\mathsf{SU}(2)$ case, since in that case, J_α is the matrix

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{E.2}$$

For the general case, we define a homomorphism ϕ from $\mathsf{su}(2)$ into $\mathfrak{k}$ by mapping the usual basis for $\mathsf{su}(2)$ to the elements J_α, K_α, and L_α. Since $\mathsf{SU}(2)$ is simply connected, Theorem 3.7 tells us that there is a Lie group homomorphism $\Phi_\alpha : \mathsf{SU}(2) \to K$ with the property that $\Phi(\exp X) = \exp \phi(X)$ for all $X \in \mathsf{su}(2)$. Applying this with X equal to the matrix J_α in (E.2), we see that

$$\begin{aligned} \exp(2\pi J_\alpha) &= \exp(\phi(2\pi X)) = \Phi(\exp(2\pi X)) \\ &= \Phi(I) = I. \end{aligned}$$

□

We are now ready to state the main theorem.

Theorem E.7. *Let K be a compact connected Lie group with Lie algebra $\mathfrak{k}$ and let $\mathfrak{t}$ be a maximal commutative subalgebra of $\mathfrak{k}$. Let Γ be the kernel of the exponential mapping for $\mathfrak{t}$ and let J be the lattice generated by the real co-roots. Then,*

$$\pi_1(K) \cong \Gamma / J.$$

Let us see how this works out in the case of $\mathsf{SU}(2)$ and $\mathsf{SO}(3)$. For the case of $\mathsf{SU}(2)$, we may consider the maximal commutative subalgebra $\mathfrak{t}$ consisting of the diagonal matrices inside $\mathsf{su}(2)$. Then, there are two roots (negatives of each other), with corresponding root vectors X and Y, denoted α and $-\alpha$. The real co-root corresponding to α is the element

$$J_\alpha = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

and the real co-root corresponding to $-\alpha$ is $-J_\alpha$. We now observe that

$$e^{2\pi t J_\alpha} = \begin{pmatrix} e^{2\pi i t} & 0 \\ 0 & e^{-2\pi i t} \end{pmatrix}.$$

This means that the kernel of the exponential mapping is precisely the set of integer multiples of J_α. So, in this case, Γ and J coincide and we conclude (again) that $\mathsf{SU}(2)$ is simply connected.

Meanwhile in $\mathsf{SO}(3)$, we use the isomorphism ϕ of $\mathsf{su}(2)$ and $\mathsf{so}(3)$ described in Section 4.9. We take our maximal commutative subalgebra as the image of $\mathfrak{t}$ under ϕ, in which case the real co-roots will be $\pm\phi(H_\alpha)$, where

$$\phi(H_\alpha) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 0 \end{pmatrix}.$$

Now, we compute that

$$e^{2\pi t \phi(H_\alpha)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(4\pi t) & -\sin(4\pi t) \\ 0 & \sin(4\pi t) & \cos(4\pi t) \end{pmatrix}.$$

We see that Γ consists of all *integer or half-integer* multiples of H_α, and, so, Γ/J has two elements and is isomorphic to $\mathbb{Z}/2$. Thus, $\pi_1(\mathsf{SO}(3)) \cong \mathbb{Z}/2$.

We may also consider the case of $\mathsf{U}(1)$, which is compact but not semisimple. In this case, there are no roots and, therefore, no co-roots. This means that $J = \{0\}$. However, Γ is still isomorphic to $\mathbb{Z}$, and, so, $\pi_1(U) \cong \mathbb{Z}/\{0\} \cong \mathbb{Z}$. In general, if K is compact but not semisimple, then $\pi_1(K)$ will be infinite.

We have already encountered in Section 7.4 the distinction between algebraically integral elements and analytically integral elements. Theorem E.7 allows us to understand the relationship between the two sets of integral elements. So, we consider the set of **algebraically integral real elements**. These are the elements μ of $\mathfrak{t}$ with the property that

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z} \tag{E.3}$$

for all real roots α. We also consider the set of **analytically integral real elements**. These are the elements μ of $\mathfrak{t}$ with the property that

$$\langle \mu, H \rangle \in \mathbb{Z}$$

for all H in $\mathfrak{t}$ with the property that $\exp(2\pi H) = I$.

From now, on we omit the word "real" and assume that all roots and integral elements are of the real variety. We have the following important consequence of Theorem E.7.

Corollary E.8. *The set of algebraically integral elements coincides with the set of analytically integral elements if and only if K is simply connected.*

In general, the analytically integral elements are those that can occur as weights of representations of the group K, whereas the algebraically integral elements are those that occur as weights of representations of the Lie algebra $\mathfrak{k}$. In the simply-connected case, the representations of K are in one-to-one correspondence with the representations of $\mathfrak{k}$, and, correspondingly, the analytically integral elements are (in that case) the same as the algebraically integral elements.

Proof. The analytically integral elements are by definition those elements of $\mathfrak{t}$ whose inner product with each element of Γ is an integer; that is, the lattice of analytically integral elements is the "dual lattice" to the kernel of the exponential mapping. Meanwhile, comparing the defining condition (E.3) for the algebraically integral elements with the formula (E.1) for the co-roots, we see that the algebraically integral elements are those whose inner product with each co-root is an integer, and, hence, whose inner product with each element of J is an integer. (This means that the lattice of algebraically integral elements is the "dual lattice" to the lattice generated by the co-roots.) If K is simply connected, then by Theorem E.7, J and Γ coincide and, so, the algebraically integral and analytically integral elements coincide.

If K is not simply connected then, by Theorem E.7, J must be a proper subset of Γ. Now, if $\alpha_1, \ldots, \alpha_r$ are the positive simple roots (where r is the dimension of the part of $\mathfrak{t}$ spanned by the α's), then $J_{\alpha_1}, \ldots, J_{\alpha_r}$ are linearly independent (over $\mathbb{R}$), and the elements of J are precisely the linear combinations of these elements with integer coefficients. Then, consider an element H of Γ that is not in J. The (unique) expansion of H in terms of $J_{\alpha_1}, \ldots, J_{\alpha_r}$ must have some noninteger coefficients. Then, at least one of the fundamental weights (which are algebraically integral elements) will take noninteger values on Γ and will, therefore, be an algebraically integral element that is not analytically integral. □

We have considered so far two subsets of $\mathfrak{t}$: the kernel Γ of the exponential mapping and the set J of all integer linear combinations of the co-roots. There is one other subset of $\mathfrak{t}$ that is relevant, namely the set

$$\Lambda = \{ H \in \mathfrak{t} | \, \langle \alpha, H \rangle \in \mathbb{Z}, \text{ for all real roots } \alpha \}.$$

It can be shown that every element of the kernel of the exponential mapping has this property, so that we have the inclusions

$$J \subset \Gamma \subset \Lambda.$$

Theorem E.9. *Let K be a connected compact group with Lie algebra $\mathfrak{k}$ and let $\mathfrak{t}$ be a maximal commutative subalgebra of $\mathfrak{k}$. Let Λ denote the set of $H \in \mathfrak{t}$ such that $\langle \alpha, H \rangle$ is an integer for all real roots α and let Γ denote the kernel of the exponential mapping. Then, $\Gamma \subset \Lambda$ and*

$$\Lambda/\Gamma \cong Z(K),$$

where $Z(K)$ denotes the center of K.

Note here that K does not have to be simply connected or semisimple. If K is commutative, then there are no roots and, so, in that case, Λ is simply all of $\mathfrak{t}$. In that case, we have $\Lambda/\Gamma = \mathfrak{t}/\Gamma = K$ (i.e., $K = Z(K)$). If K is not semisimple, then the center of K will be at least one dimensional; for example, the center of $\mathrm{U}(n)$ is the set of matrices of the form $e^{i\theta}I$ and has dimension one.

Let us see how Theorem E.9 compares to Theorem 8.30, which applies in the case that K is simply connected. From Theorem E.7, we have that when K is simply connected, Γ is the same as J. So, Theorem E.9 tells us that, in the simply-connected case, $Z(K)$ is isomorphic to Λ/J. Meanwhile, Theorem 8.30 says that $Z(K)$ is isomorphic to the set of (algebraically) integral elements modulo the root lattice.

Let us now see how to reconcile these two results. We have already said that a lattice Γ inside a finite-dimensional real inner-product space E is the set of integer linear combinations of a basis for E. The **dual lattice** Γ' is defined as

$$\Gamma' = \{ H \in E | \, \langle \gamma, H \rangle \in \mathbb{Z} \text{ for all } \gamma \in \Gamma \}.$$

This is, again, a lattice; if $\gamma_1, \ldots, \gamma_r$ is the generating set for Γ, then a generating set for Γ' is the set $\gamma^1, \ldots, \gamma^r$ satisfying

$$\langle \gamma_j, \gamma^k \rangle = \delta_{jk}. \tag{E.4}$$

(To show that such γ^k's do exist, take γ^k inside the one-dimensional orthogonal complement of the span of $\{\gamma_j\}_{j \neq k}$.) The notion of duality is symmetric: If Γ' is the dual lattice to Γ, then Γ is the dual lattice to Γ'. This is a consequence of the symmetry of (E.4) between the γ_j's and the γ^k's.

If Γ_1 and Γ_2 are lattices with $\Gamma_1 \subset \Gamma_2$, then we may regard both lattices as commutative groups under vector addition and form the quotient group Γ_2/Γ_1. This will be a finite commutative group. We make use of the following elementary result from lattice theory. (See, for example, Proposition 1.3.8 in Martinet (2003).)

Proposition E.10. *If $\Gamma_1 \subset \Gamma_2$, then $\Gamma_2' \subset \Gamma_1'$ and*

$$\Gamma_2/\Gamma_1 \cong \Gamma_1'/\Gamma_2'.$$

The lattice of algebraically integral elements is precisely the dual lattice to the lattice J generated by the co-roots. After all, if $\langle \mu, H_\alpha \rangle$ is an integer for each co-root α, then $\langle \mu, H \rangle$ is an integer whenever H is an integer combination of co-roots. Meanwhile, Λ is precisely the dual to the lattice generated by the roots, and, thus, by the symmetry of the duality relationship, the dual lattice to Λ is the root lattice. So,

$$\begin{aligned} J' &= \text{ algebraically integral elements,} \\ \Lambda' &= \text{ root lattice.} \end{aligned}$$

Proposition E.10 then tells us that

$$\Lambda/J \cong (\text{algebraically integral elements})/(\text{root lattice}).$$

Thus, we see that Theorem 8.30 and Theorem E.9 have the same content in the simply-connected case.

Note that "dual" to the inclusions $J \subset \Gamma \subset \Lambda$, we have inclusions in the reverse directions:

$$\begin{aligned} \text{Root lattice} \subset & \text{ analytically integral elements} \\ \subset & \text{ algebraically integral elements.} \end{aligned}$$

We conclude this section with a brief discussion of how one goes about proving Theorem E.7. We consider the inclusion map $i : T \to K$. This map induces a homomorphism $i_* : \pi_1(T) \to \pi_1(K)$. Concretely, this means simply that every loop in T is also a loop in K and that two loops in T that are homotopic in T are certainly homotopic in K. The reverse is not true: Two loops in T that are not homotopic in T may, nevertheless, be homotopic in K. Thus, the map $i_* : \pi_1(T) \to \pi_1(K)$ may not be injective.

Although $i_* : \pi_1(T) \to \pi_1(K)$ is typically not injective, it can be shown that it is surjective. This says that every loop in K (say, based at the identity) is homotopic to a loop in T. One way to prove this is to use differential geometry. We choose on K a Riemannian metric that is invariant under both the left and the right action of K. Such a metric exists because K is compact. Then, it is a standard calculation (Helgason (1978)) that with respect to such a metric, the geodesics through the identity are precisely the maps of the form $t \to \exp(tX)$, for X in the Lie algebra $\mathfrak{k}$. It is also a standard result from differential geometry that every homotopy class of loops in a compact Riemannian manifold contains a geodesic. So, in each homotopy class of loops based at the identity, we can find one of the form $t \to \exp(tX)$, where for this curve to be a loop, it must be that $\exp(X) = I$. Now, it is a standard result about compact groups that every element of the Lie algebra is conjugate to an element of $\mathfrak{t}$. So, there exists g in K such that $gXg^{-1} \in \mathfrak{t}$. Since K is connected, we can find a path $g(s)$ connecting the identity to g, and then the one-parameter family of loops

$$(s,t) \to g(s)e^{tX}g(s)^{-1}$$

is a homotopy of the loop $t \to \exp(tX)$ with the loop $g\exp(tX)g^{-1} = \exp(t\,gXg^{-1})$, which lies in T.

We conclude, then, that every loop in K is homotopic to a loop in T. This means that the map $i_* : \pi_1(T) \to \pi_1(K)$ is surjective. It remains, then, to compute $\pi_1(T)$ and to compute the kernel i_* and then we will have $\pi_1(K) \cong \pi_1(T)/\ker i_*$. Fortunately, T is just a torus and its fundamental group is easily shown to be isomorphic to Γ. Specifically, if $H \in \Gamma$ (i.e., $\exp(2\pi H) = I$), then we may consider the loop $t \to \exp(2\pi tH)$ in T. It is not hard to show that every loop in T is homotopic to one and only one loop of this form, essentially because $T \cong \mathfrak{t}/\Gamma$ and $\mathfrak{t}$ is simply connected. We must then determine which loops of this form are homotopically trivial in K.

If H is equal to one of the real co-roots J_α, then the loop $t \to \exp(2\pi tJ_\alpha)$ is homotopically trivial. To see this, we again consider the homomorphism $\Phi_\alpha : \mathsf{SU}(2) \to K$ described earlier in this section. Then, the loop $t \to \exp(2\pi tJ_\alpha)$ is the image of the loop $t \to \exp(2\pi tJ)$ in $\mathsf{SU}(2)$. However, $\mathsf{SU}(2)$ is simply connected and, so, the loop $t \to \exp(2\pi tJ)$ is homotopically trivial. This means that the loop $t \to \exp(2\pi tJ_\alpha)$ is the continuous image of a homotopically trivial loop and is therefore homotopically trivial. This argument shows that the kernel of i_* contains the lattice J. To show that there is nothing else in the kernel requires some additional effort—see Bröcker and tom Dieck (1985).

E.5 Fundamental Groups of Noncompact Lie Groups

In Section 1.7, we described the polar decomposition for $\mathsf{SL}(n;\mathbb{R})$, namely that every matrix A in $\mathsf{SL}(n;\mathbb{R})$ can be decomposed uniquely as

$$A = RP,$$

with R in $\mathsf{SO}(n)$ and P a positive symmetric matrix with real entries and determinant one. We can carry this one step further and write

$$P = e^X,$$

where X is a real symmetric matrix with trace zero. (To obtain X from P, diagonalize P, and then take the logarithm of the eigenvalues, which must be real and positive and whose product is equal to 1.) Let $\mathfrak{p}$ denote the subspace of $\mathsf{sl}(n;\mathbb{R})$ consisting of symmetric matrices (with trace zero). It can be shown that the map

$$\begin{aligned} \mathsf{SO}(n) \times \mathfrak{p} &\to \mathsf{SL}(n;\mathbb{R}), \\ (R, X) &\to Re^X \end{aligned}$$

is a homeomorphism. Thus,

$$\pi_1(\mathsf{SL}(n;\mathbb{R})) \cong \pi_1(\mathsf{SO}(n)) \times \pi_1(\mathfrak{p}) \cong \pi_1(\mathsf{SO}(n)),$$

since $\mathfrak{p} \cong \mathbb{R}^d$ is simply connected. (Here, $d = \dim \mathfrak{p}$.) Thus, the computation of the fundamental group of the noncompact group $\mathsf{SL}(n;\mathbb{R})$ can be reduced to the computation of the fundamental group of the compact group $\mathsf{SO}(n)$, which can be done by the methods of the previous two sections.

Similar arguments apply to the other noncompact matrix Lie groups G that we have studied, provided that G is semisimple or reductive. In each such case, the Lie algebra $\mathfrak{g}$ of G will decompose as $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$, where $\mathfrak{k}$ is the space of skew-self-adjoint elements in $\mathfrak{g}$ and $\mathfrak{p}$ is the space of self-adjoint elements in $\mathfrak{g}$. Then, the subgroup K of G with Lie algebra $\mathfrak{k}$ will be a compact subgroup of $\mathsf{U}(n)$ (called a **maximal compact subgroup**) and we will get a decomposition $g = x \exp X$, with $x \in K$ and $X \in \mathfrak{p}$. This shows that G is homeomorphic to $K \times \mathfrak{p}$ and, therefore, that $\pi_1(G) \cong \pi_1(K)$. The following list shows the resulting isomorphisms. Once the fundamental groups of the compact groups have been computed using the results of the previous two sections, this allows us to fill in the table of fundamental groups of noncompact groups from Chapter 1.

$$\begin{aligned}
\pi_1(\mathsf{GL}(n;\mathbb{R})^+) &\cong \pi_1(\mathsf{SO}(n)),\\
\pi_1(\mathsf{GL}(n;\mathbb{C})) &\cong \pi_1(\mathsf{U}(n)),\\
\pi_1(\mathsf{SL}(n;\mathbb{C})) &\cong \pi_1(\mathsf{SU}(n)),\\
\pi_1(\mathsf{SO}(n;\mathbb{C})) &\cong \pi_1(\mathsf{SO}(n)),\\
\pi_1(\mathsf{SO}_e(n,1)) &\cong \pi_1(\mathsf{SO}(n)),\\
\pi_1(\mathsf{Sp}(n;\mathbb{R})) &\cong \pi_1(\mathsf{U}(n)),\\
\pi_1(\mathsf{Sp}(n;\mathbb{C})) &\cong \pi_1(\mathsf{Sp}(n)).
\end{aligned}$$

References

1. **Sheldon Axler**, *Linear Algebra Done Right.* Second edition, Springer-Verlag, 1997.
2. **Andrew Baker**, *Matrix Groups: An Introduction to Lie Group Theory.* Springer-Verlag, 2002.
3. **Martin R. Bridson and Simon M. Salamon** (editors), *Invitations to Geometry and Topology*, Oxford University Press, 2002.
4. **Theodor Bröcker and Tammo tom Dieck**, *Representations of Compact Lie Groups.* Springer-Verlag, 1985.
5. **Morton L. Curtis,** *Matrix Groups.* Springer-Verlag, 1979.
6. **J. J. Duistermaat and J. A. C. Kolk**, *Lie Groups*, Springer-Verlag, 2000.
7. **Hans Freudenthal and H. de Vries**, *Linear Lie Groups*, Academic Press, 1969.
8. **William Fulton and Joe Harris**, *Representation theory: A First Course*, Springer-Verlag, 1991.
9. **Allen Hatcher,** *Algebraic topology.* Cambridge University Press, 2002.
10. **Kenneth Hoffman and Ray Kunze**, *Linear Algebra.* Second edition, Prentice-Hall, 1971.
11. **Roger Howe**, "Very Basic Lie Theory," *American Mathematical Monthly*, **90** (1983), 600–623.
12. **Sigurdur Helgason**, *Differential Geometry, Lie Groups, and Symmetric Spaces.* Academic Press, 1978.
13. **James E. Humphreys**, *Introduction to Lie Algebras and Representation Theory.* Springer-Verlag, 1972.
14. **N. Jacobson**, *Lie Algebras.* John Wiley and Sons, 1962.
15. **Jens Carsten Jantzen**, *Representations of Algebraic Groups*, Academic Press, 1987.
16. **Anthony W. Knapp**, *Representation Theory of Semisimple Groups: An Overview Based on Examples.* Princetion University Press, 1986.
17. **Anthony W. Knapp**, *Lie Groups, Lie Algebras, and Cohomology.* Princeton University Press, 1988.
18. **Anthony W. Knapp**, *Lie Groups: Beyond an Introduction.* Birkhauser, 1996.
19. **John M. Lee**, *Introduction to Smooth Manifolds*, Springer-Verlag, 2003.
20. **Jacques Martinet**, *Perfect Lattices in Euclidean Spaces*, Springer-Verlag, 2003.

21. **W. Miller**, *Symmetry Groups and Their Applications.* Academic Press, 1972.
22. **James R. Munkres**, *Topology: A First Course.* Prentice-Hall, 1975.
23. **Andrew Pressley and Graeme Segal**, *Loop Groups*, Oxford University Press, 1986.
24. **M. Reed and B. Simon**, *Methods of Modern Mathematical Physics. III. Scattering theory.* Academic Press, New York-London, 1979.
25. **M. Reed and B. Simon**, *Methods of Modern Mathematical Physics. I. Functional analysis.* Second edition, Academic Press, New York, 1980.
26. **Wulf Rossmann**, *Lie Groups: An Introduction Through Linear Groups.* Oxford University Press, 2002.
27. **Walter Rudin**, *Principles of Mathematical Analysis.* Third edition, McGraw-Hill, 1976.
28. **Walter Rudin** *Real and Complex Analysis.* Third edition, McGraw-Hill, 1987.
29. **Jean-Pierre Serre**, *Complex Semisimple Lie Algebras.* Springer-Verlag, 1987.
30. **Barry Simon**, *Representations of Finite and Compact Lie Groups.* American Mathematical Society, 1996.
31. **Frank W. Warner**, *Foundations of Differentiable Manifolds and Lie Groups.* Springer-Verlag, 1983.
32. **V.S. Varadarajan**, *Lie Groups, Lie Algebras, and Their Representations.* Springer-Verlag, 1974.

Index

A_2, 131, 244, 256
A_3, 262
A_n, 183, 270
Ad_A, 47
ad_X, 48
abelian, *see* commutative
absolute convergence, 28
abstract root systems, 243
adjoint
  of a linear operator, 298
  of a matrix, 6, 293
adjoint mapping, 47
adjoint representation, 96
Ado's theorem, 55
algebraically integral elements, 216
analytic subgroups, *see* connected Lie subgroups
analytically integral elements, 215
angles allowed in root systems, 246
angular momentum, 101, 329
averaging method, 120

B_2, 244, 257
B_3, 262
B_n, 266, 270
$\mathfrak{b}^+$, $\mathfrak{b}^-$, 220
B^+, B^-, 220
Baker–Campbell–Hausdorff formula, 63
base of a root system, 181, 249
basepoint, 331
bilinear form
  skew-symmetric, 8
  symmetric, 6, 7
Borel subalgebra, 220
Borel–Weil theory, 218
bracket, 44

C_3, 262
C_n, 268, 270
Campbell–Hausdorff formula, *see* Baker–Campbell–Hausdorff formula
canonical commutation relations, 122
canonical form, *see* Jordan canonical form
Cartan subalgebra, 143, 162
Cauchy sequence, 28
center
  discrete subgroup of, 25, 90
  of a compact Lie group, 273
  of a group, 114, 284
  of a Lie algebra, 61, 88, 185
centralizer, 187
chamber, Weyl, *see* Weyl chamber
change of coordinates, 304
character formula, *see* Weyl character formula
character of a representation, 211
characteristic polynomial, 291
class function, 211, 213
classical groups and Lie algebras, 7, 265
classification
  of irreducible root systems, 272
  of semisimple Lie algebras, 272
Clebsch–Gordan theory, 112, 322
Clifford algebra, 81
closed subgroup, 4
closed, bounded sets, 11

co-root, 167, 172, 174, 248
commutative group, 279
  representations of, 114
commutator, 44
compact group
  complete reducibility of, 120
  fundamental group of, 333, 336
compact real form, 159
compactness, 11
  of G/B^-, 225, 228
complete reducibility, 118, 156
complete vector field, 308, 311
completeness, *see* convergence of Cauchy sequences
complex
  Lie algebra, 44, 53
  Lie group, 313
  manifold, 309
  matrix Lie group, 44, 218
complexification of a real Lie algebra, 56
component, 12
conjugacy class, 211
conjugate-linear, 171, 298
connected Lie subgroups, 82
connectedness, 12
contragredient representation, *see* dual representation
convergence
  absolute, 28
  of a sequence of matrices, 3
  of Cauchy sequences, 28
convex hull, 149, 153, 232
coordinate system, *see* local coordinates
covering group, *see* universal covering group
covering space, *see* universal cover
cross product, 53, 60

D_n, 265, 270
δ, *see* half-sum of positive roots
δ_{jk}, 5
decomposition, polar, *see* polar decomposition
denominator, Weyl, *see* Weyl denominator
derivation
  of a Lie algebra, 61
  of $C^\infty(\mathcal{M})$, 307
derivative of exponential mapping, 70
diagonalizable matrices, denseness of, 35
diagonalization, 293
differential
  of a Lie group homomorphism, 311
  of a smooth mapping, 305
differential forms, 318
differential of exponential mapping, *see* derivative of exponential mapping
dimension formula, *see* Weyl dimension formula
direct product
  of groups, 285
  of matrix Lie groups, 56, 110
direct sum
  of Lie algebras, 56
  of root systems, 245
discrete subgroup, *see* center, discrete subgroup of
double cover, 81
dual
  lattice, 339, 340
  representation, 112, 152, 230
  root system, 181, 248, 258
  space, 112, 299
Dynkin diagram, 269

$\mathsf{E}(n)$, *see* Euclidean group
E_6, E_7, E_8, 271
ε_{klm}, 325
eigenspace, 292
eigenvalue, 291
eigenvector, 291
embedded submanifold, 308
equivalence
  of representations, 92
  of root systems, 246
equivalence class, 287
Euclidean group, 9, 42
exceptional root systems, 271
exponential
  of a locally nilpotent operator, 205
  of a matrix, 27
exponential mapping, 48

F_4, 271
faithful, 91
fiber bundle, 334

flow
  along a left-invariant vector field, 311
  along a vector field, 307
Freudenthal's formula, 232
fundamental domain, 274
fundamental group, 16, 331
fundamental representations, 139
fundamental weights, 147, 255
fundamental Weyl chamber, 194, 253

G_2, 244
$\mathfrak{g}_\alpha$, 165
$\mathsf{GL}(n;\mathbb{R})$, $\mathsf{GL}(n;\mathbb{C})$, *see* general linear group
general linear group, 3, 39
generalized eigenvector, 294
generalized orthogonal group, 7, 41
group, 279
group versus Lie algebra homomorphisms, 76

H_α, *see* co-root
Haar measure, 120, 318
half-sum of positive roots, 212, 233, 235, 236, 264
Heisenberg group, 8, 41, 63, 100
higher, 132, 197
highest weight, 133, 197, 201
highest weight cyclic representation, 135, 197
Hilbert–Schmidt
  inner product, 298
  norm, 28
holomorphic
  function, 223
  representation, 219
homomorphism
  of groups, 285
  of Lie algebras, 54
  of Lie groups, 311
  of matrix Lie groups, 17
homotopic, 331
homotopy group, 333
hyperplane, 178, 244, 250

ideal, 61, 157
identity component, 12
identity element, 279
indecomposable
  Lie algebra, 157
  positive root, 250
infinite-dimensional representations, 201
infinitesimal elements, 39
inner product, 297
integers mod n, 283
integral element, 130, 192, 254
integration over group, *see* Haar measure
intertwining map, 92
invariant
  inner product, 120, 121, 146, 163
  subspace, 91
inverse, 280
inverse root, *see* co-root
irreducible
  representation, 91
  root system, 245
isomorphism
  of groups, 285
  of Lie algebras, 54
  of matrix Lie groups, 17

Jacobi identity, 53
Jacobian, 237
joint eigenvector, *see* simultaneous eigenvector
Jordan canonical form, 296

kernel
  of a homomorphism, 285
  of exponential mapping, 337
Kostant's formula, 232

L^2, 99, 209
lattice, 273
left-invariant vector field, 310
left-translation, 120, 223
length ratios allowed in root systems, 246
Lie algebra
  general, 53
  of a matrix Lie group, 38
Lie algebra homomorphism, *see* homomorphism of Lie algebras
Lie group, 20, 309
Lie group homomorphism, *see* homomorphism of Lie groups
Lie product formula, 35

Lie subgroup, *see* connected Lie subgroup
linear functional, 299
linear Lie group, *see* matrix Lie group
linear operator, 292
local coordinates, 303
locally nilpotent operator, 205
logarithm
  for a matrix Lie group, 51
  of a matrix, 32
loop, 331
Lorentz group, 7, 41
lower, 132, 197
lowering operator, 103

$M_n(\mathbb{C})$, 3
manifold, 21, 303
matrix Lie group, 3
maximal commutative subalgebra, 162, 163
maximal compact subgroup, 343
maximal torus, 187, 211, 217, 240, 336
module, *see* representation
morphism, *see* intertwining map
multiplicity, 130, 150, 194, 233

$\mathfrak{n}^+$, $\mathfrak{n}^-$, 200, 220
N^+, N^-, 220
negative root, 181, 249
nilpotent, 31, 60, 190, 295
nilpotent, locally, *see* locally nilpotent operator
nonmatrix Lie groups, 314
nontrivial
  ideal, 157
  invariant subspace, 91
norm of a matrix, 27
normal subgroup, 286
normalizer, 187

$\mathsf{O}(n)$, *see* orthogonal group
$\mathsf{O}(n;k)$, *see* generalized orthogonal group
one-parameter subgroup, 37
orbit, 150, 153
orthogonal, 266
  group, 5, 40, 265
  matrix, 5
orthogonal complement, 298
orthonormal basis, 293

π_1, *see* fundamental group
path-connected, 12, 22
permutation group, 184, 283
Peter–Weyl theorem, 210
physicists' convention, 39
Poincaré group, 10, 43
polar decomposition, 19, 342
polynomials
  action of $\mathsf{SU}(2)$ on, 97
  action of $\mathsf{SU}(3)$ on, 152
positive matrix, 19, 20
positive root, 181, 249
positive simple root, 131, 181, 249, 250
product rule, 44, 59, 305
projective representation, 330

quotient group
  for Lie groups, 312, 317
  general, 287
quotient representation, 203

R^+, *see* positive root
$\mathbb{RP}^3$, 17
raising operator, 103
rank, 164, 243
rank three, 262
rank two, 256
real
  root, 212, 336
  weight, 212
reductive Lie algebra, 157
representation, 91
right-translation, 223
root, 129, 130, 164, 171
  space, 165
  system, 180, 243
  vector, 165
  beer, 325
rotation group, *see* special orthogonal group

S^1, 9
$\mathfrak{s}^\alpha$, 169
$\mathsf{SL}(n;\mathbb{R})$, $\mathsf{SL}(n;\mathbb{C})$, *see* special linear groups
S_n, *see* permutation group
$\mathsf{SO}(3)$
  fundamental group of, 16, 335, 338

representations of, 116
$\mathsf{SO}(n)$, *see* special orthogonal group
$\mathsf{SU}(2)$
relationship to $\mathsf{SO}(3)$, 18, 116
representations of, 97, 101, 115
simple connectedness of, 15, 338
$\mathsf{SU}(n)$, *see* special unitary group
Schur's Lemma, 113
Schwarz inequality, 28, 58
self-adjoint, 293
semisimple Lie algebra, 157
simple connectedness, 15, 76, 216, 221, 339
simple Lie algebra, 157
simple root, *see* positive simple root
simultaneous
diagonalization, 300
eigenvector, 299
skew self-adjoint, 293
SN decomposition, 32, 295
special linear group, 5, 40
special orthogonal group, 5, 40
special unitary group, 6, 40
spin, 116, 322
spin groups, 81, 160
square root
of a positive matrix, 19
uniqueness of, 38
standard representation, 95
Stone-Von Neumann theorem, 101
strictly dominant, 254
structure constants, 56
subalgebra, 54
subgroup, 284
subgroup, connected, *see* connected Lie subgroup
submanifold, *see* embedded submanifold
symmetric group, *see* permutation group
symplectic group, 7, 41, 268

tangent space, 305
at the identity, 52, 310
tensor product
of representations, 110, 123, 321
of vector spaces, 108
theorem of the highest weight, 133, 197
trace of a matrix, 296
transpose of a matrix, 5
trivial representation, 96
Trotter product formula, 36

$\mathsf{U}(n)$, *see* unitary group
unimodular, 121, 318
unipotent, 60
unitarian trick, 120
unitary group, 6, 40
universal cover, 80, 312, 332
universal property of tensor products, 108
upper triangular, 296

vector operator, 324, 328
Verma module, 200

w_α, 178, 243
W-orbit, *see* orbit
weight
diagram, 149, 259, 263
lattice, 273
weight, integral, *see* integral element
Weyl chamber, 194, 253
Weyl character formula, 211, 236
Weyl denominator, 213, 237
Weyl dimension formula, 235
Weyl group
of a root system, 244
of a semisimple Lie algebra, 174
of rank-two root systems, 257
of $\mathsf{sl}(3;\mathbb{C})$, 143
of $\mathsf{sl}(n;\mathbb{C})$, 184
Weyl integral formula, 237
Wigner–Eckart theorem, 326

Zome, 262