Introduction to Lie Groups

Michael Taylor

Math. Dept., UNC
E-mail address: met@math.unc.edu

2010 *Mathematics Subject Classification.* 22-01

Key words and phrases. Lie groups, Lie algebras, representations, roots, weights, eigenspaces, homogeneous spaces

Contents

Preface

This text developed from lectures I have given on Lie groups, in Math 773, at UNC. Prerequisites include the basic first-year graduate courses in analysis, algebra, geometry, and topology, and an introductory course in manifold theory. Algebraic background can be found in [**42**], and basic analytic background can be found in [**40**].

The first chapter introduces the notion of a Lie group and provides a number of classical examples. These examples are matrix groups, such as groups of invertible matrices, orthogonal matrices, unitary matrices, and others. We introduce the algebra $\mathbb{H}$ of quaternions and matrices of quaternions, and certain compact Lie groups of such matrices. We also discuss the matrix exponential, which later will be extended in a fundamental way to the abstract Lie group setting. We end this chapter with a presentation of left and right invariant integrals on a Lie group.

Chapter 2 introduces the notion of a representation of a Lie group, and develops some of the elementary machinery of the representation theory of compact Lie groups. The invariant integral (which is bi-invariant for compact groups), plays an important role. We establish the Weyl orthogonality relations. We also prove the Peter-Weyl theorem, to the effect that matrix entries of irreducible unitary representations of a compact Lie group G, suitably normalized, yield an orthonormal basis of $L^2(G)$. We first get this for compact matrix groups, and then after a bit more theory, regarding $L^1(G)$ as a convolution algebra, for general compact Lie groups (which, as a corollary, are seen always to be isomorphic to compact matrix groups).

Chapter 3 introduces the concept of a Lie algebra $\mathfrak{g}$, associated to a Lie group G, and an exponential map $\mathrm{Exp} : \mathfrak{g} \to G$. In case G is a matrix group, we compare the general notion with an alternative construction of the Lie algebra, as the tangent space to G at the identity element, and compare the general exponential map to the matrix exponential. We show how Lie group representations give rise to Lie algebra representations. We also introduce the universal enveloping algebra of a Lie algebra.

Chapter 4 concentrates on the unitary groups $\mathrm{U}(n)$. Topics discussed include the classification of irreducible unitary representations of $\mathrm{U}(n)$, involving the notion of roots and weights, and some of their properties. We also treat the decomposition of $\otimes^k \mathbb{C}^n$ into irreducible spaces for $\mathrm{U}(n)$, and the duality with the symmetric group S_k that arises here, and also classical character formulas and some of their implications for harmonic analysis on $\mathrm{U}(n)$.

Chapter 5 discusses some further topics on analysis on $\mathrm{U}(n)$, involving the Laplace operator, arising from a bi-invariant metric on G, which is in the center of the universal enveloping algebra and hence acts as a scalar on each irreducible representation.

Chapter 6 extends some of the results of §§4.2–4.4 to the setting of general compact Lie groups, particularly discussing roots of their Lie algebras and weights of their representations. We also have material on the structure of simple Lie algebras, including a discussion of Cartan matrices and Dynkin diagrams.

Chapter 7 specializes again, this time to the setting of the orthogonal groups $\mathrm{SO}(n)$ and certain two-fold covers, denoted $\mathrm{Spin}(n)$, which are constructed in §7.4, via use of Clifford algebras, introduced in §7.3.

Chapter 8 takes a look at the representations π_k of $SO(n)$ on the eigenspaces V_k of the Laplace-Beltrami operator on the sphere S^{n-1}. It is seen that these representations are all irreducible. The analysis involves making contact with the study of harmonic functions, and in particular the classical theory of spherical harmonics, examined here through the lens of representation theory.

Chapter 9 pursues more general studies of actions of isometry groups of compact Riemannian manifolds, and the representations they induce on eigenspaces of the Laplace-Beltrami operator. Particular attention is paid to the class of compact, rank-one symmetric spaces, for which a number of results on spherical harmonics derived in Chapter 8 are seen to have extensions.

Chapter 10 studies the groups of quaternionic matrices $Sp(n)$, which, together with $SU(n)$ and $SO(n)$, constitute the classical compact Lie groups. It investigates the maximal tori of $Sp(n)$, the roots of its Lie algebra, and the weights of its irreducible representations. One feature of this chapter is a section devoted to quaternionic linear algebra, which differs enough from linear algebra over $\mathbb{R}$ and $\mathbb{C}$ to merit some discussion. We construct a family of representations of $Sp(n)$ that verifies the theorem of the highest weight for this group, and hence produces a representation-theoretic proof that $Sp(n)$ is simply connected.

Chapter 11 introduces the algebra $\mathbb{O}$ of octonions, obtained from the quaternions $\mathbb{H}$ by a process similar to that by which $\mathbb{H}$ is built from $\mathbb{C}$. A central object here is $\mathrm{Aut}(\mathbb{O})$, the group of automorphisms of $\mathbb{O}$. This is seen to be a 14-dimensional compact Lie group, isomorphic to a group denoted G_2, the first in a series of exceptional Lie groups.

This text ends with several appendices, presenting some background material in advanced calculus and ODE theory, linear algebra, and basic functional analysis, and also some further material, complementary to that in the main body of the notes.

After reading this text, the reader should be prepared to tackle more advanced treatments of Lie groups and their representation theory, such as mentioned in the references. In particular, this text should serve as preparation for study of the monograph [**38**].

Some basic notation

$\mathbb{R}$ is the set of real numbers.

$\mathbb{C}$ is the set of complex numbers.

$\mathbb{Z}$ is the set of integers.

$\mathbb{Z}^+$ is the set of integers ≥ 0.

$\mathbb{N}$ is the set of integers ≥ 1 (the "natural numbers").

$x \in \mathbb{R}$ means x is an element of $\mathbb{R}$, i.e., x is a real number.

(a, b) denotes the set of $x \in \mathbb{R}$ such that $a < x < b$.

$[a, b]$ denotes the set of $x \in \mathbb{R}$ such that $a \leq x \leq b$.

$\{x \in \mathbb{R} : a \leq x \leq b\}$ denotes the set of x in $\mathbb{R}$ such that $a \leq x \leq b$.

$[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$ and $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$.

$\overline{z} = x - iy$ if $z = x + iy \in \mathbb{C},\ x, y \in \mathbb{R}$.

$\overline{\Omega}$ denotes the closure of the set Ω.

$f : A \to B$ denotes that the function f takes points in the set A to points in B. One also says f maps A to B.

Chapter 1

First look at Lie groups

In this chapter we define the concept of a Lie group, as a smooth manifold with a group structure whose operations are smooth. For examples, we describe a number of matrix groups, such as the group of invertible linear transformations on $\mathbb{R}^n$ or $\mathbb{C}^n$, the group of orthogonal transformations on $\mathbb{R}^n$ or of unitary transformations on $\mathbb{C}^n$, and various other groups. Complementing $\mathbb{R}$ and $\mathbb{C}$, we introduce the algebra $\mathbb{H}$ of quaternions and describe some groups of $n \times n$ matrices of quaternions.

In §1.3, we define the matrix exponential Exp and establish several key properties, such as the fact that if A is a skew-symmetric $n \times n$ real matrix, then $\operatorname{Exp} A$ is an orthogonal matrix.

In the last section of this chapter, we define left and right invariant metric tensors on a Lie group G, and associated volume elements (Haar measures) and invariant integrals. If these two Haar measures coincide, we say G is unimodular. We note that if G is compact, then it is unimodular. Further characterization will be done in Chapter 3. The invariant integral will play an important role in representation theory.

1.1. Definition and first examples

A Lie group G is a group that is also a smooth manifold, such that the group operations $G \times G \to G$ and $G \to G$ given by $(g,h) \mapsto gh$ and $g \mapsto g^{-1}$ are smooth maps.

We consider some examples, starting with

$$\mathrm{Gl}(n,\mathbb{R}) = \{A \in \mathrm{M}(n,\mathbb{R}) : A^{-1} \text{ exists}\}, \tag{1.1.1}$$

where $\mathrm{M}(n,\mathbb{R})$ consists of $n \times n$ real matrices.

Proposition 1.1.1. *The set* $\mathrm{Gl}(n,\mathbb{R})$ *is open in* $\mathrm{M}(n,\mathbb{R})$.

Proof. One way to see this is to note that $\mathrm{Gl}(n,\mathbb{R}) = \{A \in \mathrm{M}(n,\mathbb{R}) : \det A \neq 0\}$, and $\det : \mathrm{M}(n,\mathbb{R}) \to \mathbb{R}$ is continuous. Here is another.

Given $A \in \mathrm{Gl}(n,\mathbb{R})$, we have $A + B = A(I + A^{-1}B)$, which is invertible provided $I + A^{-1}B$ is invertible. Now if $C \in \mathrm{M}(n,\mathbb{R})$ we have the operator norm

$$\|C\| = \sup\{\|Cv\| : v \in \mathbb{R}^n, \|v\| \le 1\}, \tag{1.1.2}$$

and we see that $\|C^k\| \le \|C\|^k$, and hence

$$\|C\| < 1 \Longrightarrow (I + C)^{-1} = \sum_{k \ge 0} (-C)^k, \tag{1.1.3}$$

with absolute convergence, so $\|A^{-1}B\| < 1$ implies $A + B$ is invertible. $\square$

The group $\mathrm{Gl}(n,\mathbb{R})$ inherits a manifold structure from the vector space $\mathrm{M}(n,\mathbb{R})$. Since $(A,B) \mapsto AB$ is bilinear, it is clearly smooth. Furthermore, $\kappa(A) = A^{-1}$ gives a smooth map on $\mathrm{Gl}(n,\mathbb{R})$, with

$$D\kappa(A)X = -A^{-1}XA^{-1}. \tag{1.1.4}$$

In fact, for $\|X\|$ small,

$$\begin{aligned} (A+X)^{-1} &= (A(I + A^{-1}X))^{-1} = (I + A^{-1}X)^{-1}A^{-1} \\ &= A^{-1} + \sum_{k \ge 1} (-1)^k (A^{-1}X)^k A^{-1}, \end{aligned} \tag{1.1.5}$$

which yields (1.1.4).

Similar considerations apply to

$$\mathrm{Gl}(n,\mathbb{C}) = \{A \in \mathrm{M}(n,\mathbb{C}) : A^{-1} \text{ exists}\}, \tag{1.1.6}$$

where $\mathrm{M}(n,\mathbb{C})$ consists of $n \times n$ complex matrices.

Many other basic examples of Lie groups arise as subgroups of $\mathrm{Gl}(n,\mathbb{R})$ and $\mathrm{Gl}(n,\mathbb{C})$. For example, we have

$$\mathrm{Sl}(n,\mathbb{F}) = \{A \in \mathrm{M}(n,\mathbb{F}) : \det A = 1\} \subset \mathrm{Gl}(n,\mathbb{F}), \quad \mathbb{F} = \mathbb{R} \text{ or } \mathbb{C}. \tag{1.1.7}$$

Other examples are

$$\begin{aligned} \mathrm{O}(n) &= \{A \in \mathrm{M}(n,\mathbb{R}) : A^*A = I\}, \\ \mathrm{U}(n) &= \{A \in \mathrm{M}(n,\mathbb{C}) : A^*A = I\}, \end{aligned} \tag{1.1.8}$$

where

$$A = (a_{jk}) \Longrightarrow A^* = (\overline{a}_{kj}). \tag{1.1.9}$$

Also we have

$$\begin{aligned} \mathrm{SO}(n) &= \{A \in \mathrm{O}(n) : \det A = 1\}, \\ \mathrm{SU}(n) &= \{A \in \mathrm{U}(n) : \det A = 1\}. \end{aligned} \tag{1.1.10}$$

The proof that (1.1.7)–(1.1.10) define Lie groups follows from the fact these groups are all smooth submanifolds of $\mathrm{M}(n,\mathbb{F})$. This fact in turn can be deduced from the following result, which is a consequence of the inverse function theorem.

Theorem 1.1.2. *(***Submersion mapping theorem.***) Let V and W be finite-dimensional vector spaces, and $F : V \to W$ a smooth map. Fix $p \in W$, and consider*

$$S = \{x \in V : F(x) = p\}. \tag{1.1.11}$$

Assume that, for each $x \in S$, $DF(x) : V \to W$ is surjective. Then S is a smooth submanifold of V. Furthermore, for each $x \in S$,

$$T_xS = \mathit{ker}\, DF(x). \tag{1.1.12}$$

For a proof of this result, see §A.1. The proof takes the following approach. Given $q \in S$, define

$$G_q : V \longrightarrow W \oplus \ker DF(q), \quad G_q(x) = (F(x), P_q(x-q)),$$

where $P_q : V \to \ker DF(q)$ is a projection. Then the inverse function theorem can be applied to G_q.

We show how Theorem 1.1.2 can be applied to show that the groups described in (1.1.7)–(1.1.10) are smooth submanifolds of $\mathrm{M}(n,\mathbb{F})$. We start with (1.1.7). Here we take

$$V = \mathrm{M}(n,\mathbb{F}), \quad W = \mathbb{F}, \quad F : V \to W, \quad F(A) = \det A. \tag{1.1.13}$$

Now given A invertible,

$$F(A+B) = \det(A+B) = (\det A)\det(I + A^{-1}B), \tag{1.1.14}$$

and inspection shows that, for $X \in \mathrm{M}(n,\mathbb{F})$,

$$\det(I+X) = 1 + \operatorname{Tr} X + O(\|X\|^2), \tag{1.1.15}$$

so

$$DF(A)B = (\det A)\,\mathrm{Tr}(A^{-1}B). \tag{1.1.16}$$

Now, given $A \in \mathrm{Sl}(n,\mathbb{F})$, or even $A \in \mathrm{Gl}(n,\mathbb{F})$, it is readily verified that

$$\tau_A : \mathrm{M}(n,\mathbb{F}) \longrightarrow \mathbb{F}, \quad \tau_A(B) = \mathrm{Tr}(A^{-1}B), \tag{1.1.17}$$

is nonzero, hence surjective, and Theorem 1.1.2 applies.

We turn to $\mathrm{O}(n)$, defined in (1.1.8). In this case,

$$\begin{gathered} V = \mathrm{M}(n,\mathbb{R}), \quad W = \{X \in \mathrm{M}(n,\mathbb{R}) : X = X^*\}, \\ F : V \longrightarrow W, \quad F(A) = A^*A. \end{gathered} \tag{1.1.18}$$

Now, given $A \in V$,

$$F(A+B) = A^*A + A^*B + B^*A + O(\|B\|^2), \tag{1.1.19}$$

so

$$DF(A)B = A^*B + B^*A = A^*B + (A^*B)^*. \tag{1.1.20}$$

We claim that

$$A \in \mathrm{O}(n) \Longrightarrow DF(A) : \mathrm{M}(n,\mathbb{R}) \to W \text{ is surjective.} \tag{1.1.21}$$

Indeed, given $X \in W$, i.e., $X \in \mathrm{M}(n,\mathbb{R}), X = X^*$, we have

$$B = \frac{1}{2}AX \Longrightarrow DF(A)B = X. \tag{1.1.22}$$

Again, Theorem 1.1.2 applies.

Similar arguments apply to $\mathrm{U}(n)$ in (1.1.8) and to the groups in (1.1.10). For $\mathrm{SU}(n)$ we take

$$\begin{gathered} V = M(n,\mathbb{C}), \quad W = \{X \in M(n,\mathbb{C}) : X = X^*\} \oplus \mathbb{R}, \\ F : V \longrightarrow W, \quad F(A) = (A^*A, \mathrm{Im}\det A). \end{gathered}$$

Note that $A \in \mathrm{U}(n)$ implies $|\det A| = 1$, so $\mathrm{Im}\det A = 0 \Leftrightarrow \det A = \pm 1$.

As a further comment on $\mathrm{O}(n)$, we note that, given $A \in \mathrm{M}(n,\mathbb{R})$, defining $A : \mathbb{R}^n \to \mathbb{R}^n$,

$$A \in \mathrm{O}(n) \Longleftrightarrow (Au, Av) = (u,v), \quad \forall\, u,v \in \mathbb{R}^n, \tag{1.1.23}$$

where (u,v) is the Euclidean inner product on $\mathbb{R}^n$:

$$(u,v) = \sum_j u_j v_j, \tag{1.1.24}$$

where $u = (u_1,\dots,u_n)$, $v = (v_1,\dots,v_n)$. Similarly, given $A \in \mathrm{M}(n,\mathbb{C})$, defining $A : \mathbb{C}^n \to \mathbb{C}^n$,

$$A \in \mathrm{U}(n) \Longleftrightarrow (Au, Av) = (u,v), \quad \forall\, u,v \in \mathbb{C}^n, \tag{1.1.25}$$

where (u, v) denotes the Hermitian inner product on $\mathbb{C}^n$:

$$(u, v) = \sum_j u_j \overline{v}_j. \tag{1.1.26}$$

Note that

$$\langle u, v \rangle = \mathrm{Re}(u, v) \tag{1.1.27}$$

defines the Euclidean inner product on $\mathbb{C}^n \approx \mathbb{R}^{2n}$, and we have

$$\mathrm{U}(n) \hookrightarrow \mathrm{O}(2n). \tag{1.1.28}$$

Analogues of $\mathrm{O}(n)$ and $\mathrm{U}(n)$, with $\mathbb{R}$ and $\mathbb{C}$ replaced by the ring $\mathbb{H}$ of quaternions, will be discussed in §1.2.

Having defined several matrix groups, we now define a family of Lie groups that are not *a priori* subgroups of $\mathrm{Gl}(N, \mathbb{F})$. Namely we define the Euclidean group $\mathrm{E}(n)$ as a group of isometries of $\mathbb{R}^n$. As a set, $\mathrm{E}(n) = \mathrm{O}(n) \times \mathbb{R}^n$, and the action of (A, v) on $\mathbb{R}^n$ is given by

$$(A, v)x = Ax + v, \quad A \in \mathrm{O}(n),\ v, x \in \mathbb{R}^n. \tag{1.1.29}$$

The group law is seen to be

$$(A, v) \cdot (B, w) = (AB, Aw + v). \tag{1.1.30}$$

Actually, $\mathrm{E}(n)$ is isomorphic to a matrix group, via

$$(A, v) \mapsto \begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix}, \tag{1.1.31}$$

as one verifies that

$$\begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} B & w \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} AB & Aw + v \\ 0 & 1 \end{pmatrix}. \tag{1.1.32}$$

There are Lie groups that are not isomorphic to matrix groups, but it is a fact (not established here) that every connected Lie group is *locally isomorphic* to a matrix group. This is a consequence of a result known as Ado's theorem.

Exercises

1. Making use of (1.1.12), show that

$$\begin{aligned} T_I O(n) &= \{A \in M(n,\mathbb{R}) : A^* = -A\}, \\ T_I U(n) &= \{A \in M(n,\mathbb{C}) : A^* = -A\}, \\ T_I SU(n) &= \{A \in M(n,\mathbb{C}) : A^* = -A,\ \operatorname{Tr} A = 0\}. \end{aligned}$$

We denote these spaces by $\mathfrak{o}(n), \mathfrak{u}(n)$, and $\mathfrak{su}(n)$, respectively.

2. Show that, for each of the groups G listed above,

$$g \in G,\ X \in T_I G \Longrightarrow gXg^{-1} \in T_I G.$$

We write $\operatorname{Ad}(g)X = gXg^{-1}$. Show that

$$g, h \in G \Longrightarrow \operatorname{Ad}(gh) = \operatorname{Ad}(g)\operatorname{Ad}(h).$$

3. In the setting of Exercise 2, show that the action of $\operatorname{Ad}(g)$ on $T_I G$ preserves the Hilbert-Schmidt norm, defined by $\|A\|_{HS}^2 = \operatorname{Tr}(A^*A)$.

4. Show that a basis of $\mathfrak{su}(2)$ is given by

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

In particular, $\mathfrak{su}(2)$ is a vector space of dimension 3, $\mathfrak{su}(2) \approx \mathbb{R}^3$.

5. Deduce that, for $G = SU(2)$,

$$\operatorname{Ad} : G \longrightarrow \mathcal{L}(T_I G)$$

gives rise to a group homomorphism

$$p : SU(2) \longrightarrow SO(3), \quad \ker p = \{\pm I\}.$$

1.2. Quaternions and the groups Sp(n)

The space $\mathbb{H}$ of quaternions is a four-dimensional real vector space, identified with $\mathbb{R}^4$, with basis elements $1, i, j, k$, the element 1 identified with the real number 1. Elements of $\mathbb{H}$ are represented as follows:

$$\xi = a + bi + cj + dk, \tag{1.2.1}$$

with $a, b, c, d \in \mathbb{R}$. We call a the real part of ξ ($a = \operatorname{Re}\xi$) and $bi + cj + dk$ the vector part. We also have a multiplication on $\mathbb{H}$, an $\mathbb{R}$-bilinear map $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$ coinciding with the standard product on the real part, and otherwise governed by the rules

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik, \tag{1.2.2}$$

and

$$i^2 = j^2 = k^2 = -1. \tag{1.2.3}$$

Otherwise stated, if we write

$$\xi = a + u, \quad a \in \mathbb{R}, \quad u \in \mathbb{R}^3, \tag{1.2.4}$$

and similarly write $\eta = b + v,\ b \in \mathbb{R},\ v \in \mathbb{R}^3$, the product is given by

$$\xi\eta = (a + u)(b + v) = (ab - u \cdot v) + av + bu + u \times v. \tag{1.2.5}$$

Here $u \cdot v$ is the dot product in $\mathbb{R}^3$ and $u \times v$ is the cross product of vectors in $\mathbb{R}^3$. The quantity $ab - u \cdot v$ is the real part of $\xi\eta$ and $av + bu + u \times v$ is the vector part. Multiplication on $\mathbb{H}$ has the following important property.

Proposition 1.2.1. *The product on* $\mathbb{H}$ *is associative, i.e.,* $(\xi\eta)\zeta = \xi(\eta\zeta)$, *for* $\xi, \eta, \zeta \in \mathbb{H}$.

We present an approach to the proof in the exercises. Another proof will be given in §10.1.

We also have a conjugation operation on $\mathbb{H}$:

$$\overline{\xi} = a - bi - cj - dk = a - u. \tag{1.2.6}$$

A calculation gives

$$\xi\overline{\eta} = (ab + u \cdot v) - av + bu - u \times v. \tag{1.2.7}$$

In particular,

$$\operatorname{Re}(\xi\overline{\eta}) = \operatorname{Re}(\overline{\eta}\xi) = (\xi, \eta), \tag{1.2.8}$$

the right side denoting the Euclidean inner product on $\mathbb{R}^4$. Setting $\eta = \xi$ in (1.2.7) gives

$$\xi\overline{\xi} = |\xi|^2, \tag{1.2.9}$$

the Euclidean square-norm of ξ. In particular, whenever $\xi \in \mathbb{H}$ is nonzero, it has a multiplicative inverse:

$$\xi^{-1} = |\xi|^{-2}\overline{\xi}. \tag{1.2.10}$$

A routine calculation gives

$$\overline{\xi\eta} = \overline{\eta}\,\overline{\xi}. \tag{1.2.11}$$

Hence (via associativity of the product on $\mathbb{H}$)

$$|\xi\eta|^2 = (\xi\eta)(\overline{\xi\eta}) = \xi\eta\overline{\eta}\overline{\xi} = |\eta|^2\xi\overline{\xi} = |\xi|^2|\eta|^2, \tag{1.2.12}$$

or

$$|\xi\eta| = |\xi|\,|\eta|. \tag{1.2.13}$$

Note that $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$ sits in $\mathbb{H}$ as a commutative subring, for which the properties (1.2.9) and (1.2.13) are familiar.

We consider the set of unit quaternions:

$$\mathrm{Sp}(1) = \{\xi \in \mathbb{H} : |\xi| = 1\}. \tag{1.2.14}$$

Using (1.2.10) and (1.2.13) it is clear that Sp(1) is a group under multiplication. It sits in $\mathbb{R}^4$ as the unit sphere S^3. We compare Sp(1) with the group SU(2), consisting of 2×2 complex matrices of the form

$$U = \begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix}, \quad \xi, \eta \in \mathbb{C}, \quad |\xi|^2 + |\eta|^2 = 1.x \tag{1.2.15}$$

The group SU(2) is also diffeomorphic to S^3. Furthermore we have:

Proposition 1.2.2. *The groups* SU(2) *and* Sp(1) *are isomorphic under the correspondence*

$$U \mapsto \xi + j\eta, \tag{1.2.16}$$

for U as in (1.2.15).

Proof. The correspondence (1.2.16) is clearly bijective. To see that it is a homomorphism of groups, we calculate:

$$\begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix} \begin{pmatrix} \xi' & -\overline{\eta}' \\ \eta' & \overline{\xi}' \end{pmatrix} = \begin{pmatrix} \xi\xi' - \overline{\eta}\eta' & -\xi\overline{\eta}' - \overline{\eta}\overline{\xi}' \\ \eta\xi' + \overline{\xi}\eta' & -\eta\overline{\eta}' + \xi\overline{\xi}' \end{pmatrix}, \tag{1.2.17}$$

given $\xi, \eta \in \mathbb{C}$. Noting that, for $a, b \in \mathbb{R}$, $j(a + bi) = (a - bi)j$, we have

$$\begin{aligned} (\xi + j\eta)(\xi' + j\eta') &= \xi\xi' + \xi j\eta' + j\eta\xi' + j\eta j\eta' \\ &= \xi\xi' - \overline{\eta}\eta' + j(\eta\xi' + \overline{\xi}\eta'). \end{aligned} \tag{1.2.18}$$

Comparison of (1.2.17) and (1.2.18) verifies that (1.2.16) yields a homomorphism of groups. $\square$

To proceed, we consider $n \times n$ matrices of quaternions:

$$A = (a_{jk}) \in \mathrm{M}(n, \mathbb{H}), \quad a_{jk} \in \mathbb{H}. \tag{1.2.19}$$

If $\mathbb{H}^n$ denotes the space of column vectors of length n, whose entries are quaternions, then $A \in \mathrm{M}(n, \mathbb{H})$ acts on $\mathbb{H}^n$ by the usual formula. If $\xi = (\xi_j)$, $\xi_j \in \mathbb{H}$, we have

$$(A\xi)_j = \sum_k a_{jk}\xi_k. \tag{1.2.20}$$

Note that

$$A : \mathbb{H}^n \longrightarrow \mathbb{H}^n \tag{1.2.21}$$

is $\mathbb{R}$-linear, and commutes with the *right action* of $\mathbb{H}$ on $\mathbb{H}^n$, defined by

$$(\xi b)_j = \xi_j b, \quad \xi \in \mathbb{H}^n, \ b \in \mathbb{H}. \tag{1.2.22}$$

Composition of such matrix operations on $\mathbb{H}^n$ is given by the usual matrix product. If $B = (b_{jk})$, then

$$(AB)_{jk} = \sum_\ell a_{j\ell} b_{\ell k}. \tag{1.2.23}$$

We define a conjugation on $\mathrm{M}(n, \mathbb{H})$; with A given by (1.2.19),

$$A^* = (\overline{a}_{kj}). \tag{1.2.24}$$

A calculation using (1.2.11) gives

$$(AB)^* = B^* A^*. \tag{1.2.25}$$

We are ready to define the groups $\mathrm{Sp}(n)$ for $n > 1$:

$$\mathrm{Sp}(n) = \{A \in \mathrm{M}(n, \mathbb{H}) : A^* A = I\}. \tag{1.2.26}$$

Note that A^* is a left inverse of the $\mathbb{R}$-linear map $A : \mathbb{H}^n \to \mathbb{H}^n$ if and only if it is a right inverse (by real linear algebra). In other words, given $A \in \mathrm{M}(n, \mathbb{H})$,

$$A^* A = I \Longleftrightarrow AA^* = I. \tag{1.2.27}$$

In particular,

$$A \in \mathrm{Sp}(n) \Longleftrightarrow A^* \in \mathrm{Sp}(n) \Longleftrightarrow A^{-1} \in \mathrm{Sp}(n). \tag{1.2.28}$$

Also, given $A, B \in \mathrm{Sp}(n)$,

$$(AB)^* AB = B^* A^* AB = B^* B = I. \tag{1.2.29}$$

Hence $\mathrm{Sp}(n)$, defined by (1.2.26), is a group. We claim that (1.2.26) defines a smooth, compact submanifold of $\mathrm{M}(n, \mathbb{H})$, so $\mathrm{Sp}(n)$ is a compact Lie group. We omit the check of smoothness, which goes along the lines of (1.1.18)–(1.1.22), but we will establish compactness, using a construction of separate interest.

We define a quaternionic inner product on $\mathbb{H}^n$ as follows. If $\xi = (\xi_j),\ \eta = (\eta_j) \in \mathbb{H}^n$, set

$$\langle \xi, \eta \rangle = \sum_j \overline{\eta}_j \xi_j. \tag{1.2.30}$$

From (1.2.8) we have

$$\mathrm{Re}\langle \xi, \eta \rangle = (\xi, \eta), \tag{1.2.31}$$

where the right side denotes the Euclidean inner product on $\mathbb{H}^n = \mathbb{R}^{4n}$. Now, if $A \in \mathrm{M}(n, \mathbb{H}),\ A = (a_{jk})$, then

$$\begin{aligned} \langle A\xi, \eta \rangle &= \sum_{j,k} \overline{\eta}_j a_{jk} \xi_k \\ &= \sum_{j,k} \overline{\overline{a}_{jk} \eta_j}\, \xi_k \\ &= \langle \xi, A^* \eta \rangle. \end{aligned} \tag{1.2.32}$$

Hence

$$\langle A\xi, A\eta \rangle = \langle \xi, A^* A \eta \rangle. \tag{1.2.33}$$

In particular, given $A \in \mathrm{M}(n, \mathbb{H})$, we have $A \in \mathrm{Sp}(n)$ if and only if $A : \mathbb{H}^n \to \mathbb{H}^n$ preserves the quaternionic inner product (1.2.30). Given (1.2.31), we have

$$\mathrm{Sp}(n) \hookrightarrow \mathrm{O}(4n). \tag{1.2.34}$$

From here it is easy to show that $\mathrm{Sp}(n)$ is closed in $\mathrm{O}(4n)$, and hence compact.

Remark. A refinement of (1.2.34) is given in the exercises below. Further results on quaternions are given in §10.1, where there is an additional refinement of (1.2.34).

Exercises

1. Define the $\mathbb{R}$-linear map $\tau : \mathbb{H} \to M(2,\mathbb{C})$ by

$$\tau(a+bi+cj+dk) = \begin{pmatrix} a+ib & -c-id \\ c-id & a-ib \end{pmatrix} = \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix}.$$

Note the resemblance to (1.2.16). In particular, $a+bi+cj+dk = \alpha + j\beta$. Show that, for $\xi, \eta \in \mathbb{H}$,

$$\tau(\xi\eta) = \tau(\xi)\tau(\eta).$$

Deduce that the product on $\mathbb{H}$ is associative, from the associativity of $M(2,\mathbb{C})$. Furthermore, show that

$$\tau(\overline{\xi}) = \tau(\xi)^*.$$

2. The proof of Proposition 1.2.2 used the associativity of the product on $\mathbb{H}$. Rework it to use Exercise 1 instead.

3. Using Exercise 1, define an injective ring homomorphism

$$\tau_n : M(n,\mathbb{H}) \longrightarrow M(2n,\mathbb{C}).$$

Deduce that

$$\tau_n : Sp(n) \longrightarrow U(2n),$$

and this is a group homomorphism. Note how this refines (1.2.34).

4. Show that

$$\xi \in \mathbb{H},\ |\xi| = 1 \Longrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} \xi & -1 \\ 1 & \overline{\xi} \end{pmatrix} \in Sp(2).$$

5. Show that

$$\xi \in \mathbb{H},\ |\xi| = 1 \Longrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -\overline{\xi} \\ \xi & 1 \end{pmatrix} \in Sp(2),$$

and more generally, if $c, s \in [-1,1],\ c^2 + s^2 = 1$,

$$\begin{pmatrix} c & -s\overline{\xi} \\ s\xi & c \end{pmatrix} \in Sp(2).$$

6. Show that

$$\alpha, \beta \in Sp(1) \Longrightarrow \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix} \in Sp(2).$$

7. Define $\gamma : Sp(1) \times Sp(1) \to \mathcal{L}(\mathbb{H})$ by

$$\gamma(\alpha,\beta)\xi = \alpha\xi\overline{\beta}, \quad \alpha,\beta \in Sp(1),\ \xi \in \mathbb{H}.$$

Show that $\gamma(\alpha, \beta)$ preserves the inner product (1.2.8), and yields a group homomorphism

$$\gamma : Sp(1) \times Sp(1) \longrightarrow SO(4).$$

Show that $\operatorname{Ker} \gamma = \{(1, 1), (-1, -1)\}$.

8. Note that for $\theta \in \mathbb{R}$, $e^{i\theta} = \cos\theta + i \sin\theta \in Sp(1)$. If also $\varphi \in \mathbb{R}$, analyze $\gamma(e^{i\theta}, e^{i\varphi}) \in SO(4)$ as follows:

$$e^{i\theta}(a + bi + cj + dk)e^{-i\varphi} = e^{i\theta}(a + bi)e^{-i\varphi} + e^{i\theta}(c + di)je^{-i\varphi}.$$

Show that $je^{-i\varphi} = e^{i\varphi}j$, and deduce that

$$\begin{aligned} &\gamma(e^{i\theta}, e^{i\varphi})(a + bi + cj + dj) \\ &= e^{i(\theta-\varphi)}(a + bi) + e^{i(\theta+\varphi)}(c + di)j. \end{aligned}$$

In matrix representation,

$$\gamma(e^{i\theta}, e^{i\varphi}) = \begin{pmatrix} R(\theta - \varphi) & \\ & R(\theta + \varphi) \end{pmatrix}, \quad R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

9. The isomorphism $SU(2) \approx Sp(1)$ given in Proposition 1.2.2 is

$$\begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix} \mapsto \xi + j\eta.$$

Show that composing $SU(2) \times SU(2) \approx Sp(1) \times Sp(1)$ with the map γ from Exercise 7 yields $\tilde{\gamma} : SU(2) \times SU(2) \to SO(4)$, satisfying

$$\tilde{\gamma}\left(\begin{pmatrix} u(\theta) & \\ & u(\varphi) \end{pmatrix}\right) = \begin{pmatrix} R(\theta - \varphi) & \\ & R(\theta + \varphi) \end{pmatrix},$$

where

$$u(\theta) = \begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{pmatrix}.$$

10. Show that, for $\xi, \eta \in \mathbb{H}$ and (ξ, η) as in (1.2.8),

$$(\overline{\xi}, \overline{\eta}) = (\xi, \eta).$$

11. A special case of (1.2.30)–(1.2.32) is that

$$\begin{aligned} a, \xi, \eta \in \mathbb{H} &\Rightarrow \langle a\xi, \eta \rangle = \langle \xi, \overline{a}\eta \rangle \\ &\Rightarrow (a\xi, \eta) = (\xi, \overline{a}\eta), \end{aligned}$$

hence

$$u \in \operatorname{Im} \mathbb{H} = \mathbb{R}^3 \Rightarrow (u\xi, \eta) = -(\xi, u\eta).$$

Show that also

$$u \in \mathbb{R}^3 \Rightarrow (\xi u, \eta) = -(\xi, \eta u).$$

Hint. By Exercise 10, $(\xi u, \eta) = -(u\overline{\xi}, \overline{\eta})$ and $(\xi, \eta u) = -(\overline{\xi}, u\overline{\eta})$.

1.3. The matrix exponential and other functions of matrices

If $A \in \mathrm{M}(n, \mathbb{C})$, we define

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k. \tag{1.3.1}$$

We also denote this by $\mathrm{Exp}(tA)$. Making use of the operator norm (1.1.2) and noting that $\|A^k\| \le \|A\|^k$, we see that (1.3.1) is absolutely convergent for all A and all t. The power series (1.3.1) can be differentiated term by term, and we obtain

$$\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A. \tag{1.3.2}$$

Using this we can establish the identity

$$e^{(s+t)A} = e^{sA} e^{tA}. \tag{1.3.3}$$

To get this, we can first compute

$$\frac{d}{dt}\bigl[e^{(s+t)A} e^{-tA}\bigr] = e^{(s+t)A} A e^{-tA} - e^{(s+t)A} A e^{-tA} = 0, \tag{1.3.4}$$

using the product rule; hence $e^{(s+t)A}e^{-tA}$ is independent of t. Evaluating at $t = 0$ gives

$$e^{(s+t)A} e^{-tA} = e^{sA}. \tag{1.3.5}$$

Setting $s = 0$ gives

$$e^{tA} e^{-tA} = I. \tag{1.3.6}$$

Thus e^{-tA} is the multiplicative inverse of e^{tA}. Using this, we can multiply both sides of (1.3.5) on the right by e^{tA} and obtain (1.3.3).

A similar argument, which we leave to the reader, gives

$$AB = BA \Longrightarrow e^{t(A+B)} = e^{tA} e^{tB}, \tag{1.3.7}$$

though such an identity fails when A and B do not commute.

We note a few easy identities:

$$e^{tX^{-1}AX} = X^{-1} e^{tA} X, \quad e^{tA^*} = \bigl(e^{tA}\bigr)^*, \tag{1.3.8}$$

given X invertible, $t \in \mathbb{R}$. If A is diagonal, e^{tA} is obtained by exponentiating the diagonal entries. Also one has

$$\det e^{tA} = e^{t \,\mathrm{Tr}\, A}. \tag{1.3.9}$$

If A is diagonal this is checked by the remarks above; it then follows for A diagonalizable, by (1.3.8). It can be shown that the set of diagonalizable matrices is dense in $\mathrm{M}(n, \mathbb{C})$, and then (1.3.9) holds for all A, by continuity. Alternatively, it is quite easy to show that there exists an *open* subset of

$M(n, \mathbb{C})$ consisting of diagonalizable matrices. Since both sides of (1.3.9) are holomorphic on $M(n, \mathbb{C})$, this suffices.

We remark on the behavior of the exponential map on the tangent space at the identity to the groups described in (1.1.7)–(1.1.10). Making use of the criterion (1.1.12), one can calculate the following:

$$
\begin{aligned}
T_I \operatorname{Sl}(n, \mathbb{F}) &= \{A \in \mathrm{M}(n, \mathbb{F}) : \operatorname{Tr} A = 0\}, \\
T_I \operatorname{O}(n) &= \{A \in \mathrm{M}(n, \mathbb{R}) : A^* = -A\} = T_I \operatorname{SO}(n), \\
T_I \operatorname{U}(n) &= \{A \in \mathrm{M}(n, \mathbb{C}) : A^* = -A\}, \\
T_I \operatorname{SU}(n) &= \{A \in \mathrm{M}(n, \mathbb{C}) : A^* = -A, \operatorname{Tr} A = 0\}.
\end{aligned}
\tag{1.3.10}
$$

For the first two, take $A = I$ in (1.1.16) and (1.1.20), respectively, yielding $DF(I)A = \operatorname{Tr} A$ and $DF(I)A = A + A^*$, respectively. Having (1.3.10) and making use of (1.3.8)–(1.3.9), one readily verifies the following.

Proposition 1.3.1. *For each Lie group listed above,*

$$
\operatorname{Exp} : T_I G \longrightarrow G. \tag{1.3.11}
$$

We will discuss how this result fits in a more general framework in §§3.1–3.2.

We next want to calculate the derivative of the map $\operatorname{Exp} : \mathrm{M}(n, \mathbb{R}) \to \operatorname{Gl}(n, \mathbb{R})$. Equivalently, if $A, B \in \mathrm{M}(n, \mathbb{R})$, we calculate

$$
\frac{d}{dt} e^{A+tB}\Big|_{t=0}. \tag{1.3.12}
$$

When A and B commute, this is easily calculated via (1.3.7). Otherwise, matters are more complicated. To calculate (1.3.12), it is useful to look at

$$
U(s,t) = e^{s(A+tB)}, \tag{1.3.13}
$$

which satisfies

$$
\frac{\partial U}{\partial s} = (A + tB)U(s,t), \quad U(0,t) = I. \tag{1.3.14}
$$

Then $U_t = \partial U / \partial t$ satisfies

$$
\frac{\partial}{\partial s} U_t(s,t) = (A + tB)U_t(s,t) + BU(s,t), \quad U_t(0,t) = 0, \tag{1.3.15}
$$

and in particular

$$
\frac{\partial}{\partial s} U_t(s,0) = AU_t(s,0) + BU(s,0), \quad U_t(0,0) = 0. \tag{1.3.16}
$$

This is an inhomogeneous linear ODE, whose solution is

$$
\begin{aligned}
U_t(s,0) &= \int_0^s e^{(s-\sigma)A} BU(\sigma, 0)\, d\sigma \\
&= \int_0^s e^{(s-\sigma)A} B e^{\sigma A}\, d\sigma.
\end{aligned}
\tag{1.3.17}
$$

We get (1.3.12) by setting $s = 1$:

$$\frac{d}{dt} e^{A+tB}\big|_{t=0} = \int_0^1 e^{(1-\sigma)A} B e^{\sigma A}\, d\sigma, \tag{1.3.18}$$

so

$$D \operatorname{Exp}(A)B = e^A \int_0^1 e^{-\sigma A} B e^{\sigma A}\, d\sigma. \tag{1.3.19}$$

The method (1.3.1) of defining the matrix exponential extends to other cases. Suppose $F(z)$ is a holomorphic function with a power series expansion

$$F(z) = \sum_{k=0}^{\infty} a_k z^k. \tag{1.3.20}$$

If (1.3.20) converges on the disk $D_R = \{z \in \mathbb{C} : |z| < R\}$, and if $A \in \mathrm{M}(n, \mathbb{C})$, $\|A\| < R$, then we can define

$$F(A) = \sum_{k=0}^{\infty} a_k A^k, \tag{1.3.21}$$

and this power series is absolutely convergent. Power series manipulations show that if also $G(z)$ is holomorphic on D_R, and we set $H(z) = F(z)G(z)$, then, for $\|A\| < R$,

$$F(A)G(A) = H(A). \tag{1.3.22}$$

We will see more examples of (1.3.21) in subsequent sections.

Here we look into one other example, namely, for $\|tA\| < 1$, set

$$\log(I + tA) = tA - \frac{t^2}{2}A^2 + \frac{t^3}{3}A^3 - \cdots = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} t^k A^k. \tag{1.3.23}$$

We aim to prove that

$$e^{\log(I+tA)} = I + tA. \tag{1.3.24}$$

To see this, note that for $\|tA\| < 1$,

$$\begin{aligned} X(t) = \log(I + tA) \Rightarrow X'(t) &= A(I - tA + t^2A^2 - \cdots) \\ &= A(I + tA)^{-1}, \end{aligned} \tag{1.3.25}$$

as follows from (1.3.23) by differentiating term by term. For such $X(t)$, we see that $X(t)$ and $X(s)$ always commute, so it follows from (1.3.19) (or otherwise) that

$$\frac{d}{dt} e^{X(t)} = X'(t) e^{X(t)}. \tag{1.3.26}$$

Consequently, if we set

$$V(t) = (I + tA)^{-1} e^{\log(I+tA)}, \tag{1.3.27}$$

we have $V(0) = I$ and

$$V'(t) = -A(I+tA)^{-2}e^{\log(I+tA)} + A(I+tA)^{-2}e^{\log(I+tA)} = 0, \tag{1.3.28}$$

so (2.24) is established.

It follows directly from (1.3.1) that

$$\mathrm{Exp}(0+B) = I + B + O(\|B\|^2),$$

and hence

$$D\,\mathrm{Exp}(0)B = B, \tag{1.3.29}$$

i.e., $D\,\mathrm{Exp}(0)$ is the identity operator on $\mathrm{M}(n,\mathbb{R})$. (This is of course also a special case of (1.3.19).) It follows from the inverse function theorem that there are neighborhoods $\mathcal{O}$ of $0 \in \mathrm{M}(n,\mathbb{R})$ and Ω of $I \in \mathrm{Gl}(n,\mathbb{R})$ such that

$$\mathrm{Exp} : \mathcal{O} \longrightarrow \Omega, \quad \text{diffeomorphically}, \tag{1.3.30}$$

hence there is a smooth inverse from Ω to $\mathcal{O}$. The results (1.3.23)–(1.3.24) provide an explicit formula for this inverse. Putting this together with Proposition 1.3.1 yields the following.

Proposition 1.3.2. *For each Lie group G listed in* (1.3.10)*, there exists a neighborhood $\mathcal{O}$ of 0 in T_IG and a neighborhood Ω of I in G such that* (1.3.30) *holds.*

Exponentiation of quaternions and quaternionic matrices

If $\xi \in \mathbb{H}$, then, parallel to (1.3.1), we have the exponential

$$e^{t\xi} = \sum_{k=0}^{\infty} \frac{t^k}{k!}\xi^k, \tag{1.3.31}$$

a convergent power series that can be differentiated term by term, to give

$$\frac{d}{dt}e^{t\xi} = \xi e^{t\xi}. \tag{1.3.32}$$

Parallel to (1.3.3), we have

$$e^{(s+t)\xi} = e^{s\xi}e^{t\xi}, \quad \forall\, s,t \in \mathbb{R},\ \xi \in \mathbb{H}, \tag{1.3.33}$$

and parallel to (1.3.7), we have

$$\xi\eta = \eta\xi \Longrightarrow e^{t(\xi+\eta)} = e^{t\xi}e^{t\eta}. \tag{1.3.34}$$

It is of interest to know we have the following explicit computation.

Proposition 1.3.3. *If $u \in \mathbb{R}^3 \subset \mathbb{H}$ and $|u| = 1$, then*

$$e^{tu} = \cos t + (\sin t)u. \tag{1.3.35}$$

Proof. This extends Euler's identity $e^{ti} = \cos t + i \sin t$, and has a similar proof. Since $|u| = 1 \Rightarrow u^{2\ell} = (-1)^\ell,\ u^{2\ell+1} = (-1)^\ell u$, we have

$$e^{tu} = \sum_{\ell=0}^{\infty} \frac{t^{2\ell}}{(2\ell)!} u^{2\ell} + \sum_{\ell=0}^{\infty} \frac{t^{2\ell+1}}{(2\ell+1)!} u^{2\ell+1}, \tag{1.3.36}$$

leading directly to (1.3.35). □

Bringing in (1.3.34), we see that, for $a \in \mathbb{R},\ u \in \mathbb{R}^3,\ u \neq 0$,

$$\xi = a + u \Rightarrow e^{t\xi} = e^{ta}\Big[(\cos t|u|) + (\sin t|u|)\frac{u}{|u|}\Big]. \tag{1.3.37}$$

Moving on, if $A \in M(n, \mathbb{H})$ then, parallel to (1.3.1) we have

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \in M(n, \mathbb{H}), \tag{1.3.38}$$

a convergent power series, yielding

$$\frac{d}{dt} e^{tA} = A e^{tA}, \tag{1.3.39}$$

and analogues of (1.3.3) and (1.3.7) hold. Also, for $A \in M(n, \mathbb{H}),\ t \in \mathbb{R}$,

$$\left(e^{tA}\right)^* = e^{tA^*}. \tag{1.3.40}$$

Furthermore, complementing (1.3.10), we have

$$T_I Sp(n) = \{A \in M(n, \mathbb{H}) : A^* = -A\}, \tag{1.3.41}$$

and

$$A \in T_I Sp(n),\ t \in \mathbb{R} \Longrightarrow e^{tA} \in Sp(n). \tag{1.3.42}$$

Exercises

1. Take

$$X = \begin{pmatrix} 0 & -\overline{\xi} \\ \xi & 0 \end{pmatrix} \in \mathfrak{sp}(2) = T_I Sp(2), \quad \xi \in \mathbb{H}, \ |\xi| = 1.$$

Show that

$$\begin{aligned} e^{tX} &= (\cos t)I + (\sin t)X \\ &= \begin{pmatrix} c & -s\overline{\xi} \\ s\xi & c \end{pmatrix}, \quad c = \cos t, \ s = \sin t. \end{aligned}$$

By (1.3.42), $e^{tX} \in Sp(2)$, for $t \in \mathbb{R}$. Relate this to Exercise 5 of §1.2.
Hint. $X^2 = -|\xi|^2 I = -I$.

2. Given

$$Y = \begin{pmatrix} u & \\ & v \end{pmatrix} \in \mathfrak{sp}(u), \quad u, v \in \operatorname{Im} \mathbb{H} = \mathbb{R}^3,$$

show that

$$e^{tY} = \begin{pmatrix} e^{tu} & \\ & e^{tv} \end{pmatrix} = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}, \quad \alpha, \beta \in Sp(1).$$

Relate this to Exercise 6 of §1.2.

3. Show that if $F(z)$ is holomorphic on D_R, with power series (1.3.20), $A \in M(n, \mathbb{C}), \|A\| < R$, and $F(A)$ is defined as in (1.3.21), then, for $t \in \mathbb{R}$,

$$\|tA\| < R \Longrightarrow \frac{d}{dt} F(tA) = AF'(tA).$$

Apply this to

$$F(z) = \log(1+z), \quad f(z) = (1+z)^{-1},$$

to re-derive (1.3.25) and (1.3.28).

4. Let $G(z)$ be holomorphic on D_S and assume $G : D_S \to D_R$, and that $G(0) = 0$. Thus, with F as in Exercise 3,

$$F(G(tA)) = \sum a_k G(tA)^k.$$

Show that if $\|tA\| < S$ and $\|G(tA)\| < R$, then

$$\begin{aligned} \frac{d}{dt} F(G(tA)) &= \sum k a_k A G'(tA) G(tA)^{k-1} \\ &= AG'(tA)F'(G(tA)). \end{aligned}$$

5. Apply Exercise 4 to $F(z) = \log(1+z)$, $G(z) = e^z - 1$, $R = 1$. Show that if $\|e^{tA} - I\| < 1$, then

$$\frac{d}{dt} \log(e^{tA}) = A,$$

and deduce that

$$\log(e^{tA}) = tA.$$

1.4. Integration on a Lie group

For our first construction, assume G is a compact subgroup of the unitary group $\mathrm{U}(n)$, sitting in $\mathrm{M}(n,\mathbb{C})$, the space of complex $n \times n$ matrices. The space $\mathrm{M}(n,\mathbb{C})$ has a Hermitian inner product,

$$(A,B) = \operatorname{Tr} AB^* = \operatorname{Tr} B^* A, \tag{1.4.1}$$

giving a real inner product $\langle A,B\rangle = \operatorname{Re}(A,B)$. This induces a Riemannian metric on G. Let us define, for $g \in G$,

$$L_g, R_g : \mathrm{M}(n,\mathbb{C}) \longrightarrow \mathrm{M}(n,\mathbb{C}), \quad L_g X = gX, \quad R_g X = Xg. \tag{1.4.2}$$

Clearly each such map is a linear isometry on $\mathrm{M}(n,\mathbb{C})$ (given that $g \in U(n)$), and we have isometries L_g and R_g on G.

A Riemannian metric tensor on a smooth manifold induces a volume element on M, as follows. In local coordinates $(x_1,\dots,x_N)$ on $U \subset M$, say the metric tensor has components $h_{jk}(x)$. Then, on U,

$$dV(x) = \sqrt{\det(h_{jk})}\, dx_1 \cdots dx_N. \tag{1.4.3}$$

See §A.2 for a demonstration that dV is well defined, independent of the choice of coordinates.

In such a way we get a volume element on a compact group $G \subset \mathrm{U}(n)$, and since L_g and R_g are isometries, they also preserve the volume element. We normalize this volume element to define normalized Haar measure on G:

$$\int_G f(g)\, dg = \frac{1}{V(G)} \int_G f\, dV. \tag{1.4.4}$$

We have left invariance

$$\int_G f(hg)\, dg = \int_G f(g)\, dg \tag{1.4.5}$$

and right invariance

$$\int_G f(gh)\, dg = \int_G f(g)\, dg, \tag{1.4.6}$$

for all $h \in G$, in such a situation.

We give a more general construction of Haar measure, working on any Lie group G. To start, we fix some Euclidean inner product on $T_eG \approx \mathfrak{g}$; call it $\langle\ ,\ \rangle_{\mathfrak{g}}$. Here e denotes the identity element of G. Defining L_g and R_g on G as in (1.4.2), we have

$$DL_{g^{-1}},\ DR_{g^{-1}} : T_gG \longrightarrow T_eG \approx \mathfrak{g}. \tag{1.4.7}$$

We define two metric tensors on G as follows. Given $U, V \in T_gG$, we define inner products

$$\begin{aligned} \langle U, V\rangle_\ell &= \langle DL_{g^{-1}}U, DL_{g^{-1}}V\rangle_{\mathfrak{g}}, \\ \langle U, V\rangle_r &= \langle DR_{g^{-1}}U, DR_{g^{-1}}V\rangle_{\mathfrak{g}}. \end{aligned} \tag{1.4.8}$$

A straightforward computation shows that, for each $g \in G$, $L_g : G \to G$ is an isometry for $\langle\ ,\ \rangle_\ell$ and $R_g : G \to G$ is an isometry for $\langle\ ,\ \rangle_r$. Now the procedure (1.4.3) yields two volume elements on G, which we denote dV_ℓ and dV_r. As noted above, isometries of Riemannian manifolds naturally preserve the induced volume elements, so we have, for all $h \in G$,

$$\int_G f(hg)\, dV_\ell(g) = \int_G f(g)\, dV_\ell(g), \quad \int_G f(gh)\, dV_r(g) = \int_G f(g)\, dV_r(g). \tag{1.4.9}$$

Thus dV_ℓ is left-invariant and dV_r is right-invariant. We call these Haar measures.

We discuss the extent to which dV_ℓ is unique. If dV'_ℓ is another left-invariant measure, given in local coordinates by a smooth multiple of Lebesgue measure, then $dV'_\ell = \varphi(g)\, dV_\ell$ for a smooth positive function φ, and from the left invariance of both measures one can deduce that $\varphi(hg) = \varphi(g)$ for all $g, h \in G$, so φ must be constant. A similar remark holds for dV_r.

We consider the effect of a right translation on dV_ℓ. For convenience set

$$I_\ell(f) = \int_G f(g)\, dV_\ell(g), \tag{1.4.10}$$

so right translation by h yields

$$I_\ell^h(f) = \int_G f(gh)\, dV_\ell(g). \tag{1.4.11}$$

It is easy to check that I_ℓ^h is left-invariant, so by the uniqueness described above we have

$$I_\ell^h(f) = \alpha(h) I_\ell(f), \tag{1.4.12}$$

for a map

$$\alpha : G \longrightarrow (0, \infty). \tag{1.4.13}$$

It is easy to show that α is smooth, and that

$$\alpha(h_1h_2) = \alpha(h_1)\alpha(h_2), \quad \forall\ h_j \in G, \tag{1.4.14}$$

i.e., α is a group homomorphism from G to the multiplicative group $(0, \infty)$.

We say G is unimodular if $\alpha \equiv 1$. In such a case, the left-invariant Haar measure is also right-invariant; we say Haar measure is bi-invariant on G, and that G is unimodular. The Haar measure constructed on a compact group $G \subset \mathrm{U}(n)$ at the begining of this section is bi-invariant. More generally, note that for any Lie group G the image of G under α is a subgroup of $(0,\infty)$; if G is compact this must be a compact subgroup, hence $\{1\}$, so every compact Lie group has a bi-invariant Haar measure. If G is compact, we normalize Haar measure as in (1.4.4), so

$$\int_G 1\,dg = 1. \tag{1.4.15}$$

Lots of noncompact Lie groups are also unimodular, but some are not unimodular. We will discuss this further in a later section.

We now give yet another construction of Haar measure, making use of differential forms. See §A.3 for material on this. Let G be any Lie group, say of dimension N. Pick any nonzero $\omega_e \in \Lambda^N T_e^* G$, where e denotes the identity element of G. Then there is a unique N-form ω_ℓ on G such that

$$\omega_\ell(e) = \omega_e, \quad L_g^*\omega_\ell = \omega_\ell, \quad \forall\ g \in G, \tag{1.4.16}$$

and a unique N-form ω_r on G such that

$$\omega_r(e) = \omega_e, \quad R_g^*\omega_r = \omega_r, \quad \forall\ g \in G. \tag{1.4.17}$$

In fact $\omega_e = L_g^*\omega_\ell(g)$ and $\omega_e = R_g^*\omega_r(g)$. If we use ω_ℓ (or ω_r) to define an orientation on G, then we have volume elements, which we denote dV_ℓ and dV_r. Again we have (1.4.9). Since $\Lambda^N T_e^* G$ is 1-dimensional, it is clear that both dV_ℓ and dV_r are unique, up to a constant positive multiple; this provides another demonstration of such uniqueness.

Note that L_g^* and R_h^* commute for each $g, h \in G$. Hence $R_g^*\omega_\ell$ is left-invariant and $L_g^*\omega_r$ is right-invariant for each $g, h \in G$. The uniqueness mentioned above implies

$$R_h^*\omega_\ell = \alpha(h)\omega_\ell, \tag{1.4.18}$$

for all $h \in G$, with α as in (1.4.12). From this point of view, (1.4.14) follows from the identity

$$R_{h_1 h_2}^* = R_{h_2}^* R_{h_1}^*. \tag{1.4.19}$$

We next comment on integrating $f(g^{-1})$. It is easily verified that for any left-invariant Haar measure dV_ℓ,

$$\int_G f(g^{-1})\,dV_\ell = I(f) \tag{1.4.20}$$

is right-invariant, i.e., equal to $\int_G f(g)\,dV_r(g)$ for some right-invariant Haar measure dV_r. See Exercise 4 below. If G is compact and (1.4.15) holds, then

$I(1) = 1$, and we have

$$\int\limits_G f(g^{-1})\,dg = \int\limits_G f(g)\,dg. \tag{1.4.21}$$

To illustrate some of the concepts discussed in this section, we will calculate explicitly Haar measure on $\mathrm{Gl}(n,\mathbb{R})$, in the form

$$dV_\ell(X) = \varphi(X)\,dX, \tag{1.4.22}$$

where

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}, \quad dX = dx_{11}\cdots dx_{nn}, \tag{1.4.23}$$

and $\varphi \in C^\infty(\mathrm{Gl}(n,\mathbb{R}))$. The condition φ must satisfy is described as follows: we have

$$L_g : \mathrm{Gl}(n,\mathbb{R}) \longrightarrow \mathrm{Gl}(n,\mathbb{R}), \quad L_g X = gX, \tag{1.4.24}$$

and the standard change of variable formula gives, for each $u \in C_0^\infty(\mathrm{Gl}(n,\mathbb{R}))$,

$$\begin{aligned} \int u(X)\varphi(X)\,dX &= \int u(L_gX)\varphi(L_gX)\,|\det DL_g(X)|\,dX \\ &= \int u(gX)\varphi(gX)\,|\det g|^n\,dX. \end{aligned} \tag{1.4.25}$$

The left invariance of dV_ℓ demands that this equal

$$\int u(gX)\varphi(X)\,dX. \tag{1.4.26}$$

Hence $\varphi(X)$ must satisfy the condition

$$\varphi(gX) = |\det g|^{-n}\varphi(X), \quad \forall\, g, X \in \mathrm{Gl}(n,\mathbb{R}). \tag{1.4.27}$$

This clearly holds if and only if φ is a constant multiple of

$$\varphi(g) = |\det g|^{-n}, \tag{1.4.28}$$

so we have dV_ℓ uniquely specified (up to a positive constant factor) as

$$dV_\ell(X) = |\det X|^{-n}\,dX. \tag{1.4.29}$$

Similar calculations show $dV_r(X)$ is given by the same formula, so $\mathrm{Gl}(n,\mathbb{R})$ is unimodular.

Exercises

1. $\mathbb{C}^* = \mathbb{C} \setminus 0$ is a multiplicative group. Show that, with $z = x + iy$,

$$dV_\ell(z) = dV_r(z) = |z|^{-2}\,dx\,dy.$$

2. Consider $Gl(n, \mathbb{C})$, an open subset of $M(n, \mathbb{C})$, with coordinates

$$Z = \begin{pmatrix} z_{11} & \cdots & z_{1n} \\ \vdots & & \vdots \\ z_{n1} & \cdots & z_{nn} \end{pmatrix}, \quad dZ = dx_{11} \cdots dx_{nn}\,dy_{11} \cdots dy_{nn},$$

with $z_{jk} = x_{jk} + iy_{jk}$. Note that we have an analogue of (1.4.25), where, for the factor

$$|\det DL_g(Z)|,$$

we regard $DL_g(Z) : M(n, \mathbb{C}) \to M(n, \mathbb{C})$ as an $\mathbb{R}$-linear map. Generally, if $A : \mathbb{C}^k \to \mathbb{C}^k$ is $\mathbb{C}$-linear, it induces an $\mathbb{R}$-linear map on $\mathbb{R}^{2k} = \mathbb{C}^k$, and the determinants are related by

$$\det_{\mathbb{R}} A = |\det_{\mathbb{C}} A|^2.$$

With this in mind, show that

$$dV_\ell(Z) = |\det Z|^{-2n}\,dZ,$$

with $\det Z = \det_{\mathbb{C}} Z$.

3. $\mathbb{H}^* = \mathbb{H} \setminus 0$ is a multiplicative group. Show that, in coordinates

$$\xi = \xi_1 + \xi_2 i + \xi_3 j + \xi_4 k,$$

we have

$$dV_\ell(\xi) = dV_r(\xi) = |\xi|^{-4}\,d\xi,$$

where $d\xi = d\xi_1 \cdots d\xi_4$.

4. Let dV_ℓ be left invariant, so $I_\ell(f) = \int_G f(x)\,dV_\ell(x)$ satisfies $I_\ell({}_g f) = I_\ell(f)$, ${}_g f(x) = f(gx)$. Set

$$I(f) = I_\ell(f^\vee), \quad f^\vee(x) = f(x^{-1}).$$

Show that I is right-invariant.
Hint. With $f_g(x) = f(xg)$, show that $(f_g)^\vee = {}_{g^{-1}}(f^\vee)$.

Chapter 2

Lie groups and representations

A major theme in this text is the study of representations of a Lie group G on a vector space V, with emphasis on the cases where G is compact and V is finite dimensional and endowed with an inner product, and the representation π is unitary. We also assume $\pi(g)$ is a continuous function of g. We show that this automatically implies $\pi(g)$ is smooth in g, when V is finite dimensional. A major tool for this result, as well as for most of the results of this chapter is the invariant integral.

Key results established here include the break-up of representations of a compact group G into irreducible components, the Weyl orthogonality relations for irreducible unitary representations of G, and the Peter-Weyl theorem, which says the matrix entries of such representations, suitably normalized, form an orthonormal basis of $L^2(G)$. We first prove this when G is a compact matrix group. Then, after introducing the convolution algebra $L^1(G)$, we provide a proof valid for an arbitrary compact Lie group, which incidentally implies that any such group is isomorphic to a matrix group.

2.1. Basic notions of representation theory

We define a representation of a Lie group on a finite-dimensional vector space V to be a continuous map

$$\pi : G \longrightarrow \operatorname{End}(V) \tag{2.1.1}$$

such that

$$\pi(e) = I, \quad \pi(gg') = \pi(g)\pi(g'), \quad \forall\, g, g' \in G. \tag{2.1.2}$$

Note that then $\pi(g^{-1}) = \pi(g)^{-1}$, so in fact $\pi : G \to \mathrm{Gl}(V)$. If V is a real vector space with a Euclidean inner product and

$$(\pi(g)v, \pi(g)w) = (v, w), \quad \forall\, g \in G,\ v, w \in V, \tag{2.1.3}$$

we say π is an orthogonal representation. If V is complex with a Hermitian inner product and (2.1.3) holds, we say π is a unitary representation. Representations of a *compact* Lie group are unitarizable, as follows.

Proposition 2.1.1. *If π is a representation of a compact Lie group on a finite-dimensional vector space V, then V has an inner product for which* (2.1.3) *holds.*

Proof. Pick some Hermitian inner product $((\ ,\ ))$ on V. Then define $(\ ,\ )$ on V by

$$(u, v) = \int_G ((\pi(g)u, \pi(g)v))\, dg. \tag{2.1.4}$$

We have, for all $h \in G$,

$$\begin{aligned}(\pi(h)u, \pi(h)v) &= \int_G ((\pi(g)\pi(h)u, \pi(g)\pi(h)v))\, dg \\ &= \int_G ((\pi(gh)u, \pi(gh)v))\, dg \\ &= (u, v),\end{aligned} \tag{2.1.5}$$

by right invariance of Haar measure on G. $\square$

We say a representation π of G on V is irreducible if V has no proper invariant linear subspace. Not all representations break up into irreducibles, but all unitary representations do.

Proposition 2.1.2. *If π is a unitary representation of G on a finite-dimensional space V, then V is a direct sum of subspaces on which π acts irreducibly.*

Proof. If $V_0 \subset V$ is a linear space invariant under the action of π and π is unitary, then $V_0^\perp$ is also invariant. If V_0 and/or $V_0^\perp$ have proper invariant subspaces, repeat this process. Since $\dim V < \infty$, it must terminate. $\square$

We now discuss an important result in representation theory known as Schur's lemma. This has two parts.

Lemma 2.1.3. *Suppose π and λ are finite-dimensional, irreducible unitary representations of G on V and W. Assume $A : V \to W$ satisfies*

$$A\pi(g) = \lambda(g)A, \quad \forall g \in G. \tag{2.1.6}$$

Then either $A = 0$ or A is an isomorphism. In the latter case, A must be a scalar multiple of a unitary map from V to W.

Proof. One sees that $\operatorname{Ker} A \subset V$ is invariant under $\pi(g)$ for all $g \in G$, so $\operatorname{Ker} A = 0$ or V. Also the range $\operatorname{Ran} A \subset W$ is invariant under $\lambda(g)$ for all $g \in G$, so $\operatorname{Ran} A = 0$ or W. The last statement of Lemma 2.1.3 follows from the next lemma. $\square$

Lemma 2.1.4. *Suppose π is a finite-dimensional, irreducible unitary representation of G on V. Assume $B : V \to V$ satisfies*

$$B\pi(g) = \pi(g)B, \quad \forall\, g \in G. \tag{2.1.7}$$

Then B is a scalar multiple of the identity.

Proof. Set $B = B_1 + iB_2$, $B_j^* = B_j$. It follows from (2.1.7) and unitarity that

$$B_j\pi(g) = \pi(g)B_j, \quad \forall\, g \in G. \tag{2.1.8}$$

Now each B_j is diagonalizable, and

$$B_jv = av \Rightarrow B_j\pi(g)v = \pi(g)B_jv = a\pi(g)v, \quad \forall\, g \in G, \tag{2.1.9}$$

so π leaves each eigenspace of B_j invariant. Irreducibility implies each B_j is scalar, so the lemma is proven. $\square$

Finally, we set $B = A^*A$ to prove the last assertion in Lemma 2.1.3. In fact, if λ and π are unitary, (2.1.6) implies also $A^*\lambda(g) = \pi(g)A^*$, so $A^*A\pi(g) = A^*\lambda(g)A = \pi(g)A^*A$, for all g, hence $A^*A = aI$ for some $a \in \mathbb{C}$. In fact, $a > 0$ since $A^*A \geq 0$.

Given two finite-dimensional representations π and λ of G on V and W, we say π and λ are equivalent ($\pi \approx \lambda$) if and only if there is an isomorphism $A : V \to W$ such that $A^{-1}\lambda(g)A = \pi(g)$ for all $g \in G$. If these representations are unitary, we say they are unitarily equivalent provided such a unitary A exists. It follows that when π and λ are irreducible and unitary, they are equivalent if and only if they are unitarily equivalent. In fact, this holds regardless of whether π and λ are irreducible.

The following result is a first version of what will be a very important extension in the next section.

Proposition 2.1.5. *Let G be a compact Lie group, π a unitary representation of G on V, a finite-dimensional vector space with an inner product. Set*

$$Pv = \int_G \pi(g)v\,dg. \tag{2.1.10}$$

Then P is the orthogonal projection of V on the space where π acts trivially.

The proof consists of four easy pieces:

$$\pi(g)Pv = Pv, \quad \forall\, g \in G, \tag{2.1.11}$$

$$P^* = \int \pi(g^{-1})\,dg = P, \tag{2.1.12}$$

$$P^2 = \iint \pi(g)\pi(h)\,dg\,dh = \iint \pi(gh)\,dg\,dh = P, \tag{2.1.13}$$

$$\pi(g)v = v\ \forall\, g \Longrightarrow Pv = v. \tag{2.1.14}$$

Each step follows from the bi-invariance of Haar measure on G when it is compact.

We next record some ways of producing new representations from old. Let π be a representation of G on V and λ a representation of G on W. First, there is the direct sum of two representations. We define $\pi \oplus \lambda$, a representation of G on $V \oplus W$, as

$$(\pi \oplus \lambda)(g)(v, w) = (\pi(g)v, \lambda(g)w), \quad v \in V,\ w \in W, \tag{2.1.15}$$

with $(v, w) \in V \oplus W$.

Next, we define the representation λ/π of G on $\mathrm{Hom}(V, W)$, by

$$\lambda/\pi(g)A = \lambda(g)A\pi(g)^{-1}, \quad A \in \mathrm{Hom}(V, W). \tag{2.1.16}$$

If V and W are finite-dimensional inner product spaces, then $\mathrm{Hom}(V, W)$ gets the Hermitan inner product

$$(A, B) = \mathrm{Tr}\, AB^*, \tag{2.1.17}$$

and if π and λ are unitary, so is λ/π, since

$$\begin{aligned}
&(\lambda(g)A\pi(g)^{-1}, \lambda(g)B\pi(g)^{-1}) \\
&\quad= \mathrm{Tr}\,\lambda(g)A\pi(g)^{-1}\pi(g)B^*\lambda(g)^{-1} \\
&\quad= \mathrm{Tr}\, AB^*.
\end{aligned} \tag{2.1.18}$$

If we take orthonormal bases of V and W, and write matrix entries $\pi_{jk}(g)$, $A_{\ell m}$, etc., we have

$$\big(\lambda/\pi(g)A\big)_{jk} = \sum_{\ell,m} \lambda_{j\ell}(g)\, A_{\ell m}\, \overline{\pi_{km}(g)}. \tag{2.1.19}$$

In case $W = \mathbb{C}$ and $\lambda(g) \equiv 1$, we have $\mathrm{Hom}(V, W) = \mathrm{Hom}(V, \mathbb{C}) = V'$, the dual of V. In such a case, we use the notation $\overline{\pi}$, so for $v \in V, \omega \in V'$,

$$\langle v, \overline{\pi}(g)\omega\rangle = \langle \pi(g^{-1})v, \omega\rangle. \tag{2.1.20}$$

Equivalently,

$$\overline{\pi}(g) = \pi(g^{-1})^t : V' \longrightarrow V'. \tag{2.1.21}$$

If V is an inner product space, π is a unitary representation of G, and $(\pi_{jk}(g))$ is the matrix representation of $\pi(g)$ with respect to a given orthonormal basis, then

$$\big(\overline{\pi_{jk}(g)}\big) \tag{2.1.22}$$

is the matrix representation of $\overline{\pi}$ with respect to the dual basis of V'.

Another notation for (2.1.16) is

$$\lambda/\pi = \lambda \otimes \overline{\pi}, \tag{2.1.23}$$

the tensor product representation, acting on

$$W \otimes V' \approx \mathrm{Hom}(V, W). \tag{2.1.24}$$

We also have $\lambda \otimes \pi$, acting on $W \otimes V \approx \mathrm{Hom}(V', W)$, given by

$$\lambda \otimes \pi(g)A = \lambda/\overline{\pi}(g)A = \lambda(g)A\overline{\pi}(g)^{-1}, \quad A \in \mathrm{Hom}(V', W). \tag{2.1.25}$$

Exercises

1. Deduce from (2.1.19) that

$$\begin{aligned}\operatorname{Tr} \lambda/\pi(g) &= \sum_{\ell,m} \lambda_{\ell\ell}(g)\overline{\pi_{mm}(g)} \\ &= \operatorname{Tr} \lambda(g) \cdot \overline{\operatorname{Tr} \pi(g)}.\end{aligned}$$

Hence

$$\operatorname{Tr} \lambda \otimes \pi(g) = \operatorname{Tr} \lambda(g) \cdot \operatorname{Tr} \pi(g).$$

2. For $k \in \mathbb{Z}^+,\ n \in \mathbb{N}$, let $\mathcal{P}_k(\mathbb{C}^n)$ denote the space of polynomials on $\mathbb{C}^n$, homogeneous of degree k. Use the inner product

$$(u, v) = \int\limits_{|z|\le 1} u(z)\overline{v(z)}\, dV(z),$$

with $dV(z) = dx_1 \cdots dx_n\, dy_1 \cdots dy_n$. Define

$$\pi_k : SU(n) \longrightarrow \mathcal{L}(\mathcal{P}_k(\mathbb{C}^n))$$

by

$$\pi_k(g)u(z) = u(g^{-1}z).$$

Show that π_k is a unitary representation of $SU(n)$ on $\mathcal{P}_k(\mathbb{C}^n)$.

3. If also $\ell \in \mathbb{Z}$, show that

$$\pi_{k,\ell}(g)u(z) = (\det g)^\ell u(g^{-1}z)$$

defines a unitary representation of $U(n)$ on $\mathcal{P}_k(\mathbb{C}^n)$.

4. Do Exercise 2 with $\mathcal{P}_k(\mathbb{C}^n)$ replaced by $\mathcal{P}_k(\mathbb{R}^n)$ and $SU(n)$ replaced by $SO(n)$. Similarly, modify Exercise 3 to produce unitary representations of $O(n)$.

2.2. Weyl orthogonality

Let G be a compact Lie group. Assume π is an irreducible unitary representation of G on V and λ an irreducible unitary representation of G on W. Define P acting on $\mathrm{Hom}(V, W)$ as follows. If $A : V \to W$, set

$$P(A) = \int_G \lambda(g) A \pi(g)^{-1} \, dg. \tag{2.2.1}$$

It is readily verified that

$$\lambda(g) P(A) \pi(g)^{-1} = P(A), \quad \forall\, g \in G. \tag{2.2.2}$$

In other words, $P(A)$ intertwines π and λ. Now Schur's lemma, established in §2.1, gives the following:

$$\begin{aligned} \pi \text{ not} \approx \lambda &\Longrightarrow P(A) = 0, \quad \forall A, \\ \pi = \lambda &\Longrightarrow P(A) = c_\pi(A) I, \end{aligned} \tag{2.2.3}$$

where $c_\pi(A)$ is scalar and I the identity operator on $V = W$.

In the latter case, taking the trace yields $d_\pi \, c_\pi(A) = \mathrm{Tr}\, A$ (where $d_\pi = \dim\ V$), hence $c_\pi(A) = d_\pi^{-1}\ \mathrm{Tr}\, A$, so

$$\int_G \pi(g) A \pi(g)^* \, dg = d_\pi^{-1} (\mathrm{Tr}\, A) I. \tag{2.2.4}$$

If matrix entries are denoted $\pi(g)_{jk},\ A_{jk}$, etc., we have (cf. (2.1.19))

$$\begin{aligned} \sum_{k,\ell} \int_G \pi(g)_{jk} A_{k\ell} \overline{\pi(g)}_{m\ell} \, dg &= d_\pi^{-1}\, \delta_{jm}\ \mathrm{Tr}\, A \\ &= d_\pi^{-1}\, \delta_{jm} \sum_{k,\ell} \delta_{k\ell} A_{k\ell}, \end{aligned} \tag{2.2.5}$$

hence

$$\int_G \pi(g)_{jk} \overline{\pi(g)}_{m\ell} \, dg = d_\pi^{-1}\, \delta_{jm}\, \delta_{k\ell}. \tag{2.2.6}$$

On the other hand, if π is not $\approx \lambda$, the first case of (2.2.3) applies. In this case, $P(A)_{jm}$ is equal to

$$\sum_{k,\ell} \int_G \lambda(g)_{jk} A_{k\ell} \overline{\pi(g)}_{m\ell} \, dg = 0, \quad \forall\, A \in \mathrm{Hom}(V, W), \tag{2.2.7}$$

and this yields

$$\int_G \lambda(g)_{jk} \overline{\pi(g)}_{m\ell} \, dg = 0, \quad \text{if } \pi \text{ not } \approx \lambda, \tag{2.2.8}$$

for each $j, k \in \{1, \ldots, d_\lambda\}$ and $\ell, m \in \{1, \ldots, d_\pi\}$. Together, (2.2.6) and (2.2.8) make up the Weyl orhogonality relations. The following is a convenient restatement.

Proposition 2.2.1. *Let $\{\pi^\alpha : \alpha \in \mathcal{I}\}$ be a mutually inequivalent family of ireducible unitary representations of a compact Lie group G, on spaces V_α of dimension d_α. Then*

$$\{d_\alpha^{1/2} \pi_{jk}^\alpha : \alpha \in \mathcal{I},\ 1 \le j, k \le d_\alpha\} \tag{2.2.9}$$

is an orthonormal set in $L^2(G)$, where $\pi_{jk}^\alpha(g)$ denotes the matrix representation of $\pi^\alpha(g)$ with respect to an orthonormal basis of V_α.

Exercises

1. Let π and λ be unitary representations of a compact Lie group G on V and W, and define $Q_{\lambda\pi} : \mathcal{L}(V, W) \to \mathcal{L}(V, W)$ by

$$Q_{\lambda\pi} A = \int_G \lambda(g) A \pi(g)^{-1} \, dg.$$

Show that $Q_{\lambda\pi}$ is the orthogonal projection of $\mathcal{L}(V, W)$ onto

$$\mathcal{I}(\lambda, \pi) = \{A \in \mathcal{L}(V, W) : \lambda(g)A = A\pi(g), \ \forall \, g \in G\}.$$

Relate this to Proposition (2.1.5), applied to the representation ν of G on $\mathcal{L}(V, W)$ given by $\nu(g)A = \lambda(g)A\pi(g)^{-1}$, or, as defined in §2.1, $\nu = \lambda/\pi$.

2. In case λ and π are also irreducible, relate the conclusion of Exercise 1 to the results (2.2.3)–(2.2.4).

2.3. The Peter-Weyl theorem, I

Let G be a compact Lie group and let $\pi^\alpha,\ \alpha \in \mathcal{I}$, be a maximal set of mutually inequivalent irreducible unitary representations of G, on spaces V_α, of dimension d_α. Pick an orthonormal basis of V_α and say the corresponding matrix entries of π^α are $\pi^\alpha_{jk},\ 1 \le j,k \le d_\alpha$. In §2.2 it was shown that $\{d_\alpha^{1/2}\,\pi^\alpha_{jk}\}$ forms an orthonormal set in $L^2(G)$. The Peter-Weyl theorem asserts the completeness of this orthonormal set.

Theorem 2.3.1. *The set* $\{d_\alpha^{1/2}\,\pi^\alpha_{jk} : \alpha \in \mathcal{I}, 1 \le j,k \le d_\alpha\}$ *is an orthonormal basis of* $L^2(G)$.

What it remains to prove is that the linear span $\mathcal{E}$ of $\{\pi^\alpha_{jk}\}$ is dense in $L^2(G)$. We will give a proof of this here, under the additional hypothesis that G is isomorphic to a subgroup of $\mathrm{U}(N)$, for some N. We will show that $\mathcal{E}$ is dense in the space $C(G)$ of continuous functions on G, using the Stone-Weierstrass theorem. Since $C(G)$ is dense in $L^2(G)$ and has a stronger topology, this suffices. Clearly $\mathcal{E}$ is a linear space and $1 \in \mathcal{E}$. To apply the Stone-Weierstrass theorem, we need to have the following:

$$\mathcal{E} \ \text{ separates points of } \ G, \tag{2.3.1}$$

$$u \in \mathcal{E} \Longrightarrow \overline{u} \in \mathcal{E}, \tag{2.3.2}$$

$$u, v \in \mathcal{E} \Longrightarrow uv \in \mathcal{E}. \tag{2.3.3}$$

Of these conditions, (2.3.1) follows directly from the hypothesis $G \subset \mathrm{U}(N)$. As for (2.3.2), if π^α has matrix representation (π^α_{jk}), then $(\overline{\pi}^\alpha_{jk})$ is also the matrix of an irreducible unitary representation of G. Finally, we note that the tensor product representation $\pi^\alpha \otimes \pi^\beta$ on $V_\alpha \otimes V_\beta$ (defined as in (2.1.23)–(2.1.25)) can be decomposed into irreducibles, by Proposition 2.1.2, and this gives (2.3.3).

That these arguments can be applied to all compact G can be stated as follows.

Proposition 2.3.2. *If G is a compact Lie group, then there is an injective representation*

$$\rho : G \longrightarrow \mathrm{U}(N). \tag{2.3.4}$$

We say G has a faithful unitary representation.

Actually, we will prove this result in §2.8, as a *corollary* to the Peter-Weyl theorem, which will be proven for all compact Lie groups in that section.

From the Peter-Weyl theorem it follows that, if $u \in L^2(G)$, then

$$u = \sum_{\alpha \in \mathcal{I}} d_\alpha^{1/2} \sum_{j,k} \hat{u}_{jk}(\alpha)\, \pi_{jk}^\alpha(g), \tag{2.3.5}$$

where

$$\hat{u}_{jk}(\alpha) = d_\alpha^{1/2} \int\limits_G u(g)\, \overline{\pi_{jk}^\alpha(g)}\, dg, \tag{2.3.6}$$

the convergence in (2.3.5) holding in L^2-norm. See the basic development of Hilbert space theory in Appendix C, §C.2.

Let us also set

$$\mathcal{P}_\alpha u = d_\alpha^{1/2} \sum_{j,k} \hat{u}_{jk}(\alpha) \pi_{jk}^\alpha(g), \tag{2.3.7}$$

the orthogonal projection of u onto the space

$$\mathcal{V}_\alpha = \operatorname{span}\{\pi_{jk}^\alpha : 1 \le j, k \le d_\alpha\}. \tag{2.3.8}$$

We have, for $u \in L^2(G)$,

$$u = \sum_{\alpha \in \mathcal{I}} \mathcal{P}_\alpha u, \tag{2.3.9}$$

convergence in L^2-norm.

Another way to write (2.3.7) is as

$$\mathcal{P}_\alpha u(g) = d_\alpha^{1/2} \operatorname{Tr}\big(\hat{u}(\alpha)^t \pi^\alpha(g)\big), \tag{2.3.10}$$

where

$$\hat{u}(\alpha) = d_\alpha^{1/2} \int\limits_G u(g) \overline{\pi}^\alpha(g)\, dg. \tag{2.3.11}$$

Here, as in (2.1.21), we can define $\overline{\pi}^\alpha$ as the representation of G on V_α' given by

$$\overline{\pi}^\alpha(g) = \pi^\alpha(g^{-1})^t : V_\alpha' \longrightarrow V_\alpha', \tag{2.3.12}$$

so $\hat{u}(\alpha) : V_\alpha' \to V_\alpha'$. If $(\pi_{jk}^\alpha(g))$ is the matrix representation of $\pi^\alpha(g)$ with respect to an orthonormal basis of V_α, then $(\overline{\pi_{jk}^\alpha(g)})$ is the matrix representation of $\overline{\pi}^\alpha(g)$ with respect to the dual basis of V_α'. The Hermitian inner product $(\ ,\ )$ on V_α gives rise to a conjugate linear isomorphism

$$C : V_\alpha \longrightarrow V_\alpha', \quad (u, v) = \langle u, Cv \rangle, \tag{2.3.13}$$

and a straightforward calculation gives

$$\overline{\pi}^\alpha(g) = C \pi^\alpha(g) C^{-1}. \tag{2.3.14}$$

We also note that for a unitary representation π^α we have (as asserted shortly below (2.3.3))

$$\pi^\alpha \text{ irreducible} \Longrightarrow \overline{\pi}^\alpha \text{ irreducible.} \tag{2.3.15}$$

Indeed, if $E \subset V'_\alpha$ is a $\mathbb{C}$-linear subspace invariant under $\overline{\pi}^\alpha(g)$ for all g, then $C^{-1}E \subset V_\alpha$ is a $\mathbb{C}$-linear subspace invariant under $\pi^\alpha(g)$.

Exercises

We look at an explicit version of Theorem 2.3.1 in case $G = \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, with one-dimensional unitary representations

$$e_k : \mathbb{T} \longrightarrow S^1, \quad e_k(\theta) = e^{ik\theta}, \quad k \in \mathbb{Z}.$$

We have the inner product on $L^2(\mathbb{T})$,

$$(u, v) = \frac{1}{2\pi} \int_0^{2\pi} u(\theta)\overline{v(\theta)}\, d\theta.$$

1. Show that $\{e_k : k \in \mathbb{Z}\}$ is an orthonormal set, i.e.,

$$(e_k, e_\ell) = \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-\ell)}\, d\theta = \delta_{k\ell}.$$

2. Let $\mathcal{L} = \mathrm{Span}\{e_k : k \in \mathbb{Z}\}$. Note that

$$1 = e_0, \quad \overline{e}_k = e_{-k}, \quad e_k e_\ell = e_{k+\ell}.$$

Deduce that $\mathcal{L}$ is an algebra of continuous functions on $\mathbb{T}$ satisfying (2.3.1)–(2.3.3), and hence, by the Stone-Weierstrass theorem,

$$\mathcal{L} \text{ is dense in } C(\mathbb{T}).$$

3. Deduce that $\mathcal{L}$ is an orthonormal basis of $L^2(\mathbb{T})$ and, by the general theory treated in Appendix C,

$$u \in L^2(\mathbb{T}), \quad S_N u = \sum_{|k|\le N} \hat{u}(k) e_k, \quad \hat{u}(k) = \frac{1}{2\pi} \int_0^{2\pi} u(\theta) e^{-ik\theta}\, d\theta$$

$$\Longrightarrow S_n \to u \text{ in } L^2\text{-norm, as } N \to \infty.$$

Note. The series

$$u = \sum_{k=-\infty}^{\infty} \hat{u}(k) e^{ik\theta}$$

is called the Fourier series of u.

4. Extend the constriction above to $G = \mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$, with

$$e_k(\theta) = e^{ik\cdot\theta}, \quad k \in \mathbb{Z}^n, \quad k \cdot \theta = k_1\theta_1 + \cdots + k_n\theta_n.$$

2.4. Characters and central functions

Let G be a Lie group. A function u on G is said to be central provided that, for each $h \in G$, $u(h^{-1}gh) = u(g)$ (for a.e. $g \in G$ if $u \in L^1_{\mathrm{loc}}(G)$). Examples of central functions on G include

$$\operatorname{Tr} \rho(g) = \chi_\rho(g), \tag{2.4.1}$$

where ρ is a representation of G on a finite-dimensional vector space. We call χ_ρ the *character* of ρ.

Suppose now that G is compact. Let π^α, $\alpha \in \mathcal{I}$ be a maximal set of mutually inequivalent unitary representations of G, on vector spaces V_α. We set $\chi_\alpha = \operatorname{Tr} \pi^\alpha$. Note that, if $\alpha \neq \beta$,

$$\int\limits_G \chi_\alpha(g)\overline{\chi_\beta(g)}\, dg = \sum_{j,k} \int\limits_G \pi^\alpha_{jj}(g)\overline{\pi^\beta_{kk}(g)}\, dg = 0, \tag{2.4.2}$$

as a consequence of (2.2.8). On the other hand, using (2.2.6) we have

$$\begin{aligned} \int\limits_G \chi_\alpha(g)\overline{\chi_\alpha(g)}\, dg &= \sum_{j,k} \int\limits_G \pi^\alpha_{jj}(g)\overline{\pi^\alpha_{kk}(g)}\, dg \\ &= \sum_{j,k} d_\alpha^{-1}\delta_{jk} \\ &= 1. \end{aligned} \tag{2.4.3}$$

Hence $\{\chi_\alpha : \alpha \in \mathcal{I}\}$ is an orthonormal set in $L^2(G)$. We have more, from the Peter-Weyl theorem, proved for a certain class of compact G in §2.3, and to be proved for general compact Lie groups in §2.8.

Proposition 2.4.1. *The set $\{\chi_\alpha : \alpha \in \mathcal{I}\}$ is an orthonormal basis of*

$$L^2_{\mathcal{C}}(G) = \{u \in L^2(G) : u \text{ is central}\}. \tag{2.4.4}$$

Proof. Take $u \in L^2_{\mathcal{C}}(G)$. By (2.3.9)–(2.3.10) we can write

$$u = \sum_{\alpha \in \mathcal{I}} \mathcal{P}_\alpha u, \quad \mathcal{P}_\alpha u = d_\alpha^{1/2} \operatorname{Tr}\big(\hat{u}(\alpha)^t\, \pi^\alpha(g)\big) \in \mathcal{V}_\alpha, \tag{2.4.5}$$

with convergence in L^2-norm. We claim that each term $\mathcal{P}_\alpha u$ is a multiple of χ_α. Note that, for each $h \in G$,

$$u_h(g) = u(h^{-1}gh) \Longrightarrow \hat{u}_h(\alpha) = \overline{\pi}^\alpha(h)\hat{u}(\alpha)\overline{\pi}^\alpha(h)^{-1}. \tag{2.4.6}$$

Hence

$$u \ \text{ central} \Longrightarrow \hat{u}(\alpha)\overline{\pi}^\alpha(h) = \overline{\pi}^\alpha(h)\hat{u}(\alpha), \quad \forall\, h \in G. \tag{2.4.7}$$

Thus, by Schur's lemma, $\hat{u}(\alpha)$ is a scalar multiple of the identity. Taking traces in (2.3.11) gives $\operatorname{Tr}\hat{u}(\alpha) = d_\alpha^{1/2}\int_G u(g)\overline{\chi_\alpha(g)}\,dg$, which gives

$$\begin{aligned} u \text{ central} \Longrightarrow \hat{u}(\alpha) &= c^\alpha(u)I, \quad c^\alpha(u) = d_\alpha^{-1/2}(u,\chi_\alpha)_{L^2(G)} \\ \Longrightarrow \mathcal{P}_\alpha u(g) &= d_\alpha^{1/2}c^\alpha(u)\operatorname{Tr}\pi^\alpha(g) = (u,\chi_\alpha)_{L^2(G)}\,\chi_\alpha(g), \end{aligned} \tag{2.4.8}$$

finishing the proof. □

We now establish a generalization of Proposition 2.1.5.

Proposition 2.4.2. *Let G be a compact Lie group and ρ a unitary representation of G on a finite-dimensional vector space V. Set*

$$P_\alpha = d_\alpha \int_G \overline{\chi_\alpha(g)}\,\rho(g)\,dg. \tag{2.4.9}$$

Then P_α is the orthogonal projection of V onto the space where G acts like copies of π^α.

Proof. As shown in §2.1, one has an orthogonal direct sum decomposition

$$V = V_1 \oplus \cdots \oplus V_K, \tag{2.4.10}$$

with V_j invariant under ρ and $\rho_j = \rho|_{V_j}$ irreducible; say $\rho_j \approx \pi^{\beta_j}$. The content of the proposition is that

$$\begin{aligned} u \in V_j,\ \beta_j \neq \alpha &\Longrightarrow P_\alpha u = 0, \\ u \in V_j,\ \beta_j = \alpha &\Longrightarrow P_\alpha u = u. \end{aligned} \tag{2.4.11}$$

The first part of (2.4.11) follows from the identity

$$\int_G \overline{\chi_\alpha(g)}\,\pi^\beta_{k\ell}(g)\,dg = 0 \ \Leftarrow \beta \neq \alpha, \tag{2.4.12}$$

a consequence of (2.2.8). The second part of (2.4.11) follows from the identity

$$d_\alpha \int_G \overline{\chi_\alpha(g)}\,\pi^\alpha_{k\ell}(g)\,dg = \delta_{k\ell}, \tag{2.4.13}$$

a consequence of (2.2.6). □

The number of factors V_j in (2.4.10) for which $\rho_j \approx \pi^\alpha$ is called the multiplicity of the irreducible representation π^α in ρ and denoted $\mu(\pi^\alpha,\rho)$. This is seen to be the dimension of the image of P_α divided by d_α, i.e., by (2.4.9),

$$\mu(\pi^\alpha,\rho) = d_\alpha^{-1}\operatorname{Tr}P_\alpha = \int_G \chi_\rho(g)\overline{\chi_\alpha(g)}\,dg. \tag{2.4.14}$$

It is apparent that two finite-dimensional unitary representations of a compact Lie group G are equivalent if and only if they break up into the same irreducible components, with the same multiplicities. Thus we have the following.

Proposition 2.4.3. *If ρ and λ are finite-dimensional unitary representations of a compact Lie group G, then*

$$\rho \approx \lambda \Longleftrightarrow \chi_\rho = \chi_\lambda. \tag{2.4.15}$$

Conjugacy classes.

If G is a Lie group, we say g_1 and $g_2 \in G$ are conjugate (in G) if there exists $h \in G$ such that $g_1 = h^{-1} g_2 h$. We write $g_1 \sim g_2$, This is an equivalence relation, and the set $G/\sim$ of equivalence classes is called the set of conjugacy classes of G. Central functions are precisely those that are constant on each conjugacy class, so it is of interest to understand what conjugacy classes look like.

Suppose $G = U(n)$. Given $A \in U(n)$, $\mathbb{C}^n$ has an orthonormal basis $\{u_k\}$ of eigenvectors of A,

$$Au_k = \lambda_k u_k, \quad |\lambda_k| = 1. \tag{2.4.16}$$

It is clear that an element $B \in U(n)$ is unitarily conjugate to A if and only if A and B have the same eigenvalues, with the same multiplicity, or equivalently

$$A \sim B \text{ in } U(n) \Longleftrightarrow \det(\lambda I - A) = \det(\lambda I - B), \tag{2.4.17}$$

as polynomials in λ.

In case $G = SU(n)$, $A, B \in SU(n)$, we have the criterion (2.4.17) for conjugacy in $U(n)$. However, it is elementary that, given $A, B \in SU(n)$,

$$A \sim B \text{ in } SU(n) \Longleftrightarrow A \sim B \text{ in } U(n). \tag{2.4.18}$$

Indeed, if $X \in U(n)$ and $A = X^{-1}BX$, we can set $X = \omega Y$, $Y \in SU(n)$, $\omega \in \mathbb{C}$, $\omega^n = \det X$, and note that $A = Y^{-1}BY$.

Now consider $G = O(n)$. Given $A \in O(n)$, $\operatorname{Spec} A$ consists of points in $\{z \in \mathbb{C} : |z| = 1\}$, and non-real eigenvalues occur in complex conjugate pairs. For example, if $\lambda_1 \in \operatorname{Spec} A$, $\lambda_1 \notin \mathbb{R}$, $Av_1 = \lambda_1 v_1$, say $v_1 = x_1 + iy_1$, $x_1, y_1 \in \mathbb{R}^n$. Note that

$$A\overline{v}_1 = \overline{\lambda}_1 \overline{v}_1, \text{ so } \overline{v}_1 \perp v_1, \text{ hence } |x_1| = |y_1|, \ x_1 \perp y_1. \tag{2.4.19}$$

Writing $\lambda_1 = c_1 + is_1$, $c_1, s_1 \in \mathbb{R}$, $c_1^2 + s_1^2 = 1$, we have

$$A(x_1 + iy_1) = (c_1 + is_1)(x_1 + iy_1), \tag{2.4.20}$$

hence

$$Ax_1 = c_1 x_1 - s_1 y_1, \quad Ay_1 = s_1 x_1 + c_1 y_1. \tag{2.4.21}$$

The space $V_1 = \operatorname{Span}\{x_1, y_1\}$ is invariant under A, hence $A : V_1^\perp \to V_1^\perp$, a real inner product space of dimension $n-2$. Inductively, we obtain the following.

Proposition 2.4.4. *Let $A \in O(n)$. Then $\mathbb{R}^n$ has an orthonormal basis of the following form. If $\{\lambda_j, \overline{\lambda}_j : 1 \le j \le k\}$ are the complex conjugate pairs of non-real eigenvalues of A, counted by multiplicity, then there is an orthonormal set $S = \{x_j, y_j \in \mathbb{R}^n : 1 \le j \le k\}$ such that, with $\lambda_j = c_j + is_j,\ c_j, s_j \in \mathbb{R}$,*

$$Ax_j = c_j x_j - s_j y_j, \quad Ay_j = s_j x_j + c_j y_j. \tag{2.4.22}$$

If $m = n - 2k > 0$, there is an orthonormal set $\{x_j : k+1 \le j \le k+m\}$ which, together with S, forms an orthonormal basis of $\mathbb{R}^n$, such that

$$Ax_j = \lambda_j x_j, \quad \text{for } \ k+1 \le j \le k+m, \quad \lambda_j \in \{\pm 1\}. \tag{2.4.23}$$

Hence we have the following analogue of (2.4.17).

Corollary 2.4.5. *Given $A, B \in O(n)$, $A \sim B$ in $O(n)$ if and only if these matrices have the same eigenvalues, counted according to multiplicity. Equivalently,*

$$A \sim B \ \text{in } O(n) \Longleftrightarrow \det(\lambda I - A) = \det(\lambda I - B). \tag{2.4.24}$$

For $G = SO(n)$, there is a partial analogue of (2.4.18). Namely, given $A, B \in SO(n)$,

$$A \sim B \text{ in } SO(n) \Longleftrightarrow A \sim B \text{ in } O(n), \text{ provided } n \text{ is odd.} \tag{2.4.25}$$

Indeed, n odd $\Rightarrow$ $\det(-I) = -1$, so if $X \in O(n)$, $A = X^{-1}BX$, and $\det X = -1$, set $Y = -X \in SO(n)$ to get $A = Y^{-1}BY$.

For n even, matters are different. For example, $SO(2)$ is commutative, so

$$\text{for } A, B \in SO(2), \ A \sim B \text{ in } SO(2) \Longleftrightarrow A = B. \tag{2.4.26}$$

On the other hand, for $c^2 + s^2 = 1,\ s \ne 0$,

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix} \sim \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \ \text{in} \ O(2), \quad \text{not in } SO(2). \tag{2.4.27}$$

Exercises

1. Consider $A, B \in SU(2)$,

$$A = \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}, \quad z, w \in \mathbb{C},\ |z|^2 + |w|^2 = 1.$$

Show that

$$A \sim B \text{ in } SU(2) \Longleftrightarrow \operatorname{Tr} A = \operatorname{Tr} B.$$

Note that $\operatorname{Tr} A = 2 \operatorname{Re} z$.

2. Given $\xi, \eta \in Sp(1)$, i.e., $\xi, \eta \in \mathbb{H}$, $|\xi| = |\eta| = 1$, show that

$$\operatorname{Re}(\eta\xi\overline{\eta}) = \operatorname{Re}\xi.$$

Going further, if also $\zeta \in Sp(1)$, show that

$$\xi \sim \zeta \Longleftrightarrow \operatorname{Re}\xi = \operatorname{Re}\zeta.$$

3. Given $A, B \in SO(2)$, show that

$$A \sim B \text{ in } O(2) \Longleftrightarrow \operatorname{Tr} A = \operatorname{Tr} B.$$

4. Given $A \in SO(3)$, show that there exists $R \in SO(2)$ such that

$$A \sim \begin{pmatrix} R & \\ & 1 \end{pmatrix} \text{ in } SO(3).$$

5. Given $A \in SO(4)$, show there exist $R_j \in SO(2)$ such that

$$A \sim \begin{pmatrix} R_1 & \\ & R_2 \end{pmatrix} \text{ in } O(4).$$

See if you can replace conjugation in $O(4)$ by conjugation in $SO(4)$.

6. Set $O_-(n) = O(n) \setminus SO(n)$. Show that

$$A \in O_-(2) \Longrightarrow A \sim \begin{pmatrix} 1 & \\ & -1 \end{pmatrix},$$

$$A \in O_-(3) \Longrightarrow A \sim \begin{pmatrix} R & \\ & -1 \end{pmatrix}, \quad R \in SO(2),$$

$$A \in O_-(4) \Longrightarrow A \sim \begin{pmatrix} R & & \\ & 1 & \\ & & -1 \end{pmatrix}, \quad R \in SO(2).$$

7. Let π be an irreducible unitary representation of a compact Lie group G on V, of dimension d_π. Take $B \in \mathcal{L}(V)$. Set

$$\chi_{\pi,B}(g) = \operatorname{Tr}(B\pi(g)).$$

Let B and $\pi(g)$ have matrix representations (b_{jk}) and $\pi_{jk}(g))$ with respect to some orthonormal basis of V. Show that

$$\chi_{\pi,B}(g) = \sum_{j,k} b_{jk}\pi_{kj}(g).$$

Deduce that if also $C \in \mathcal{L}(V)$,

$$(\chi_{\pi,B}, \chi_{\pi,C})_{L^2(G)} = \frac{1}{d_\pi}\operatorname{Tr}(BC^*).$$

2.5. Representations of O(2)

The group $O(2)$ has two connected components, $O_+(2) = SO(2)$ and $O_-(2)$, components on which the determinant is 1 and -1, respectively. We have

$$SO(2) = \{R(\theta) : \theta \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}\}, \quad O_-(2) = \{R(\theta)T : \theta \in \mathbb{T}\}, \tag{2.5.1}$$

where

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.5.2}$$

Note that $O(2)$ is generated by $SO(2)$ and T, with relations

$$R(0) = I, \quad R(\theta + \varphi) = R(\theta)R(\varphi), \quad R(\theta)T = TR(-\theta), \quad T^2 = I. \tag{2.5.3}$$

The 2×2 representation R of $SO(2)$, regarded as a representation on $\mathbb{C}^2$, is not irreducible. We have

$$R(\theta) \sim \begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{pmatrix}. \tag{2.5.4}$$

The conjugation is done via the unitary matrix

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix},$$

which commutes with T.

Guided by these formulas, we define the following representations π_n of $O(2)$ on $\mathbb{C}^2$, for $n \in \mathbb{Z}$. To start, we set

$$\pi_n(R(\theta)) = \begin{pmatrix} e^{in\theta} & \\ & e^{-in\theta} \end{pmatrix}, \quad \pi_n(T) = T. \tag{2.5.5}$$

A calculation gives

$$\pi_n(R(\theta))T = \begin{pmatrix} & e^{in\theta} \\ e^{-in\theta} & \end{pmatrix} = T\pi_n(R(-\theta)), \tag{2.5.6}$$

paralleling the relations in (2.5.3). Hence (2.5.5) uniquely defines a group homomorphism $\pi_n : O(2) \to U(2) \subset \mathcal{L}(\mathbb{C}^2)$, and we have

$$\pi_n(R(\theta)T) = \begin{pmatrix} & e^{in\theta} \\ e^{-in\theta} & \end{pmatrix} = \pi_n(TR(-\theta)). \tag{2.5.7}$$

We see that, for $g \in O(2)$,

$$T^{-1}\pi_n(g)T = \pi_{-n}(g), \quad \text{so} \quad \pi_n \approx \pi_{-n}. \tag{2.5.8}$$

The representations π_n of $O(2)$ are irreducible for $n \neq 0$, but π_0 is not; indeed

$$\pi_0(R(\theta)) \equiv I, \quad \pi_0(T) = T. \tag{2.5.9}$$

so the eigenspaces of T in $\mathbb{C}^2$ are invariant under π_0. We have

$$\pi_0 \approx \alpha_0 \oplus \beta_0, \tag{2.5.10}$$

where α_0 and β_0 are the following one-dimensional representations of $O(2)$:

$$\alpha_0(g) = 1, \quad \beta_0(g) = \det g, \quad \forall\, g \in O(2). \tag{2.5.11}$$

Noting that the matrix entries of $\{\sqrt{2}\pi_n : n \in \mathbb{N}\}$, α_0, and β_0 yield an orthonormal basis of $O(2)$, we have the following.

Proposition 2.5.1. *A complete set of irreducible unitary representations of $O(2)$ is given by*

$$\{\pi_n : n \in \mathbb{N}\}, \quad \alpha_0, \quad \textit{and} \ \ \beta_0. \tag{2.5.12}$$

It is of interest to compute characters:

$$\begin{aligned} &\operatorname{Tr} \pi_n(R(\theta)) = 2\cos n\theta, \quad \operatorname{Tr} \pi_n(R(\theta)T) = 0, \quad n \neq 0, \\ &\operatorname{Tr} \alpha_0(g) = 1, \quad \operatorname{Tr} \beta_0(g) = \det g, \quad \forall\, g \in O(2). \end{aligned} \tag{2.5.13}$$

Note that each irreducible character of $O(2)$ is constant on $O_-(2)$, consistent with the analysis of conjugacy classes indicated in Exercise 6 of §2.4, implying that all of $O_-(2)$ is one conjugacy class in $O(2)$.

2.6. Comments on representations of finite groups

Throughout this section G will be a finite group (i.e., a compact Lie group of dimension zero). We denote its order by $o(G)$. Then the integral is given by

$$\int_G f(g)\,dg = \frac{1}{o(G)} \sum_{g\in G} f(g). \tag{2.6.1}$$

In this case $L^2(G)$ is a finite-dimensional vector space, of dimension $o(G)$, and the regular representation of G on $L^2(G)$, given by

$$L(g)u(x) = u(g^{-1}x), \quad g, x \in G,\ u \in L^2(G), \tag{2.6.2}$$

is a faithful unitary representation of G.

If $\pi^\alpha,\ \alpha \in \mathcal{I}$ is a maximal set of mutually inequivalent irreducible unitary representations of G, onto vector spaces V_α, of dimension d_α, then it is a special case of the results of §2.3 that $\{\pi^\alpha_{jk} : \alpha \in \mathcal{I}, 1 \le j, k \le d_\alpha\}$ forms an orthonormal basis of $L^2(G)$. In particular, with $\mathcal{V}_\alpha$ as in (2.3.8), we have

$$L^2(G) = \bigoplus_{\alpha\in\mathcal{I}} \mathcal{V}_\alpha, \quad \dim \mathcal{V}_\alpha = d_\alpha^2, \tag{2.6.3}$$

and hence

$$\sum_{\alpha\in\mathcal{I}} d_\alpha^2 = o(G). \tag{2.6.4}$$

Note that, for finite G, $L^2_{\mathcal{C}}(G)$ is equal to the set of all central functions on G. Hence

$$\dim L^2_{\mathcal{C}}(G) = o(\mathcal{C}), \tag{2.6.5}$$

where $\mathcal{C}$ denotes the set of conjugacy classes in G and $o(\mathcal{C})$ its cardinality. Since $\{\chi_\alpha : \alpha \in \mathcal{I}\}$ is an orthonormal basis of $L^2_{\mathcal{C}}(G)$, we deduce that

$$o(\mathcal{I}) = o(\mathcal{C}), \tag{2.6.6}$$

i.e., the number of distinct irreducible unitary representations of G is equal to the number of conjugacy classes of G.

We illustrate some of these results on a selection of finite groups, starting with a couple of the smallest symmetric groups. We denote by S_n the group of permutations of $\{1, \dots, n\}$; clearly $n! = o(S_n)$. Each group has a trivial representation, which we denote 1, acting on $\mathbb{C}$ by $1(g) = 1$, for all $g \in G$. Each group S_n has another one-dimensional representation,

$$\mathrm{sgn} : S_n \longrightarrow \{\pm 1\}. \tag{2.6.7}$$

One way to define $\operatorname{sgn}(\sigma)$ is the following. Consider

$$D_n(x) = \prod_{1\le j<k\le n} (x_j - x_k). \tag{2.6.8}$$

Then, for $\sigma \in S_n$,

$$\prod_{1\le j<k\le n} (x_{\sigma(j)} - x_{\sigma(k)}) = \operatorname{sgn}(\sigma)\, D_n(x). \tag{2.6.9}$$

It is easy to verify that $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$ for $\sigma, \tau \in S_n$.

We next define a representation ρ_n of S_n on $\mathbb{C}^n$ by

$$\rho_n(\sigma)e_j = e_{\sigma(j)}, \tag{2.6.10}$$

where $e_1, \dots, e_n$ is the standard basis of $\mathbb{C}^n$. This representation is not irreducible, since

$$\rho_n(\sigma)(e_1 + \cdots + e_n) = e_1 + \cdots + e_n, \quad \forall\, \sigma \in S_n. \tag{2.6.11}$$

The orthogonal complement of this vector is also invariant, so S_n acts on

$$V_{n-1} = \{u \in \mathbb{C}^n : u_1 + \cdots + u_n = 0\}. \tag{2.6.12}$$

Let us denote the action of S_n on V_{n-1} by π_S^n.

Lemma 2.6.1. *The representation π_S^n of S_n on V_{n-1} is irreducible.*

Proof. We note the result is trivial for S_2 acting on V_1. Now take $n \ge 3$ and consider a nonzero $v \in V_{n-1}$. We aim to show the span W of $\{\pi_S^n(\sigma)v : \sigma \in S_n\}$ is all of V_{n-1}. If we can show

$$e_1 - e_2 \in W, \tag{2.6.13}$$

this is easily accomplished. Note that W must contain a vector of the form

$$(v_1, v_2, \dots, v_n), \quad v_1 \ne v_2. \tag{2.6.14}$$

Then

$$(v_1, v_2, v_3, \dots, v_n) - (v_2, v_1, v_3, \dots, v_n) = (v_1 - v_2, v_2 - v_1, 0, \dots, 0) \in W \tag{2.6.15}$$

is nonzero, and we have (9.13). The proof is done. ☐

Note that ρ_n acts on $\mathbb{R}^n$ and this complexifies to the action on $\mathbb{C}^n$ given above. Similarly π_S^n acts on $V_{n-1}^R = \{u \in \mathbb{R}^n : u_1 + \cdots + u_n = 0\}$ and complexifies to to action on V_{n-1}. Acting on $\mathbb{R}^n$, ρ_n acts as the group of symmetries of the simplex spanned by $e_1, \dots, e_n$, lying in the surface $\{u : u_1 + \cdots + u_n = 1\}$. The projection onto V_{n-1}^R sends $\{e_1, \dots, e_n\}$ to the vertices of a simplex centered at the origin, and π_S^n acts as the group of symmetries of this simplex.

For example, via π^n_S, S_3 acts as the group of symmetries of an equilateral triangle in $\mathbb{R}^2$ and S_4 acts as the group of symmetries of a regular tetrahedron in $\mathbb{R}^3$.

We claim that, when $n = 3$, the set

$$1,\ \text{sgn},\ \pi^n_S \tag{2.6.16}$$

exhausts the set of irreducible representations of S_3. In fact, in view of (2.6.4), the dimension check

$$o(S_3) = 6 = 1^2 + 1^2 + 2^2 \tag{2.6.17}$$

verifies this.

The group S_4 has the irreducible representations (2.6.16) and a couple more. One is given by

$$\pi^4_Q(\sigma) = \text{sgn}(\sigma)\,\pi^4_S(\sigma), \tag{2.6.18}$$

acting on V_3. Since the representations π^4_S and π^4_Q are three-dimensional representations, we have

$$\det \pi^4_Q(\sigma) = \text{sgn}(\sigma)\ \det \pi^4_S(\sigma), \tag{2.6.19}$$

so they cannot be equivalent. (By contrast, $\text{sgn}\cdot\pi^3_S$ is equivalent to π^3_S.)

So far the representations of S_4 we have contribute $1 + 1 + 3^2 + 3^2 = 20$ to $o(S_4) = 24$. In addition, there is a two-dimensional representation of S_4, coming from a surjective homomorphism

$$\beta : S_4 \longrightarrow S_3. \tag{2.6.20}$$

To construct β, we need to have S_4 act on a 3-point set. To this end, consider the following situation. The regular tetrahedron $\mathcal{T}$ has 4 vertices, 4 faces, and 6 edges. The edges come in 3 sets of opposite pairs. The action of S_4 on $\mathcal{T}$ preserves this pairing, and gives the action of S_4 on a 3-point set, yielding (2.6.20). Then the representation

$$\pi^3_S \circ \beta \tag{2.6.21}$$

is a 2-dimensional irreducible representation of S_4, completing the list.

We make some more comments on the representations π^4_S and π^4_Q. Note that, for all $\sigma \in S_4$,

$$\det \pi^4_S(\sigma) = \text{sgn}(\sigma), \quad \det \pi^4_Q(\sigma) = 1. \tag{2.6.22}$$

Hence π^4_Q acts as a group of *rotations* on $V^R_3 \approx \mathbb{R}^3$. In fact, we claim π^4_Q acts as the group of rotational symmetries of a cube $\mathcal{Q} \subset \mathbb{R}^3$, centered at the origin. To see this, let G_Q denote the group of such symmetries of $\mathcal{Q}$ and refer to Figure 2.6.1, which shows a tetrahedron $\mathcal{T}$, with vertices A, B, C, D, sitting in a cube, with vertices A, B, C, D and also $A' = -A, B' = -B, C' = -C, D' = -D$. Each $g \in G_Q$ either takes $\mathcal{T}$ to $\mathcal{T}$ or takes $\mathcal{T}$ to $-\mathcal{T}$.

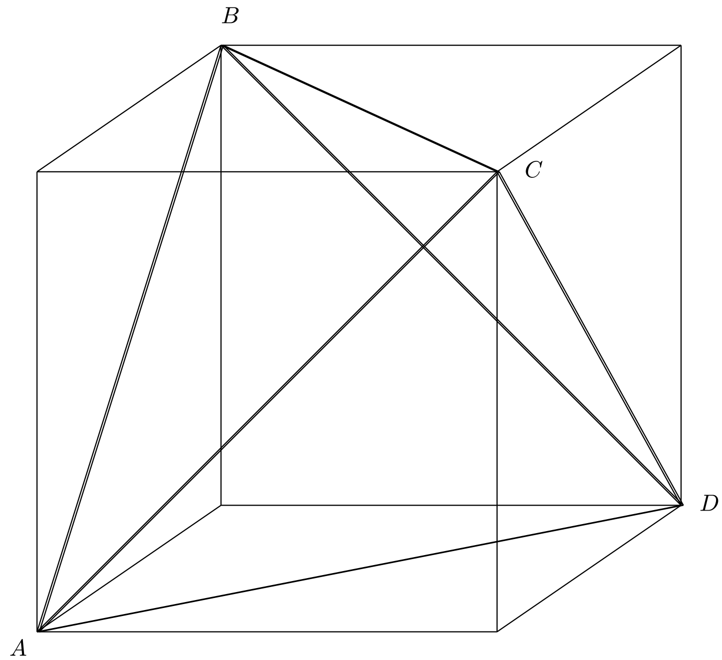

Figure 2.6.1. Geometrical setting for the actions of π_S^4 and π_Q^4

This dichotomy defines a homomorphism $\gamma : G_Q \to \{\pm 1\}$, and we see that $g \mapsto \gamma(g)g$ gives the group of symmetries of $\mathcal{T}$. This is equivalent to

$$\pi_S^4(\sigma) = \operatorname{sgn}(\sigma)\, \pi_Q^4(\sigma), \tag{2.6.23}$$

which is another way of putting (2.6.18).

We note another perspective on (2.6.20). Namely π_Q^4 acts on the 3-point set consisting of opposite pairs of faces of the cube $\mathcal{Q}$.

Dihedral groups. We next consider, for $k \geq 3$, the group D^k of symmetries of a regular k-gon in the plane. We have

$$D^k = D_+^k \cup D_-^k, \tag{2.6.24}$$

where

$$\begin{aligned} D_+^k &= \{R(\theta_j) : j \in \mathbb{Z}/(k)\}, \quad \theta_j = \frac{2\pi j}{k}, \\ D_-^k &= \{R(\theta_j)T : j \in \mathbb{Z}/(k)\}, \end{aligned} \tag{2.6.25}$$

and we set

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.6.26}$$

We see that D^k is a subgroup of $O(2)$, and $D^k_\pm$ consists of elements in $O_\pm(2)$. We define unitary representations π_n of D^k on $\mathbb{C}^2$ by specializing representations of $O(2)$ constructed in §2.5. Consequently

$$\pi_n(R(\theta_j)) = \begin{pmatrix} e^{in\theta_j} & \\ & e^{-in\theta_j} \end{pmatrix}, \quad \pi_n(T) = T, \quad n \in \mathbb{Z}/(k). \tag{2.6.27}$$

We have

$$\pi_n(R(\theta_j)T) = \begin{pmatrix} & e^{in\theta_j} \\ e^{-in\theta_j} & \end{pmatrix}. \tag{2.6.28}$$

As in §2.5,

$$\pi_n \approx \pi_{-n}, \tag{2.6.29}$$

and π_0 is reducible:

$$\pi_0 \approx \alpha_0 \oplus \beta_0, \quad \alpha_0(g) = 1, \quad \beta_0(g) = \det g, \quad g \in D^k. \tag{2.6.30}$$

For $n \in \mathbb{Z}/(k)$, we see that

$$\begin{aligned} \pi_n \text{ is reducible} \Longleftrightarrow & \begin{pmatrix} e^{in\theta_1} & \\ & e^{-in\theta_1} \end{pmatrix} \text{ commutes with } T \\ \Longleftrightarrow & \; e^{in\theta_1} \in \{\pm 1\}. \end{aligned} \tag{2.6.31}$$

Since $e^{in\theta_1} = e^{2\pi i n/k}$, we have

Lemma 2.6.2. *For $n \neq 0$ in $\mathbb{Z}/(k)$, the representation π_n of D^k is reducible if and only if*

$$k \textit{ is even and } n = \frac{k}{2}. \tag{2.6.32}$$

Note that if (2.6.32) holds, then

$$\pi_{k/2}(R(\theta_1)) = -I, \tag{2.6.33}$$

hence

$$\begin{aligned} \pi_{k/2}(R(\theta_j)) &= (-1)^j I, \quad \text{and} \\ \pi_{k/2}(R(\theta_j)T) &= (-1)^j T. \end{aligned} \tag{2.6.34}$$

We hence have

$$\pi_{k/2} \approx \alpha_{k/2} \oplus \beta_{k/2}, \tag{2.6.35}$$

with

$$\begin{aligned} \alpha_{k/2}(R(\theta_j)) &= (-1)^j = \alpha_{k/2}(R(\theta_j)T), \\ \beta_{k/2}(R(\theta_j)) &= (-1)^j = -\beta_{k/2}(R(\theta_j)T). \end{aligned} \tag{2.6.36}$$

To draw connections to (2.6.4), we note that when k is even, we have described

$$\frac{k-2}{2} \text{ 2D representations, and 4 1D representations,} \tag{2.6.37}$$

so the left side of (2.6.4) gives

$$4\,\frac{k-2}{2} + 4 = 2k = o(D^k), \tag{2.6.38}$$

and when k is odd, we have described

$$\frac{k-1}{2} \text{ 2D representations, and 2 1D representations,} \tag{2.6.39}$$

so the left side of (2.6.4) gives

$$4\,\frac{k-1}{2} + 2 = 2k = o(D^k). \tag{2.6.40}$$

In both cases we verify (2.6.4).

2.7. The convolution product and group algebras

Let G be a Lie group. Given integrable functions $u, v : G \to \mathbb{C}$ (for example, continuous functions with compact support) we define the convolution $u*v : G \to \mathbb{C}$ by

$$u * v(x) = \int_G u(g)v(g^{-1}x)\, dg. \tag{2.7.1}$$

We use left-invariant Haar measure. One easily sees that $u * v$ is continuous with compact support if u and v are. It is a consequence of Fubini's theorem that $u, v \in L^1(G) \Rightarrow u * v \in L^1(G)$. This convolution product is easily seen to have the associative property:

$$u * (v * w) = (u * v) * w. \tag{2.7.2}$$

Let π be a unitary representation of G. For $u \in L^1(G)$, we set

$$\pi(u) = \int_G u(g)\, \pi(g)\, dg. \tag{2.7.3}$$

The following relates (2.7.3) to the convolution product.

Proposition 2.7.1. *We have*

$$\pi(u * v) = \pi(u)\pi(v). \tag{2.7.4}$$

Proof. The definitions give

$$\begin{aligned}
\pi(u * v) &= \int (u * v)(g)\, \pi(g)\, dg \\
&= \iint u(h)v(h^{-1}g)\pi(g)\, dg\, dh \\
&= \iint u(h)\pi(h)\, v(h^{-1}g)\pi(h^{-1}g)\, dg\, dh \\
&= \int u(h)\pi(h)\, dh\ \pi(v) \\
&= \pi(u)\pi(v),
\end{aligned} \tag{2.7.5}$$

where left invariance of Haar measure is used in the fourth identity. □

In the rest of this section we restrict attention to the case where G is compact; in particular its Haar measure is bi-invariant. The following result bears on the meaning of "central."

Proposition 2.7.2. *If u is central, then for all $v \in L^1(G)$, $u * v = v * u$.*

Proof. Recall that to say u is central is to say $u(g^{-1}xg) = u(x)$. We have

$$\begin{aligned} v * u(x) &= \int v(g)u(g^{-1}x)\,dg \\ &= \int v(g)u(xg^{-1})\,dg \quad \text{(if } u \text{ is central)} \\ &= \int v(h^{-1}x)u(h)\,dh \\ &= u * v(x). \end{aligned} \tag{2.7.6}$$

□

When u is central, $\pi(u)$ has a special behavior, as we now derive. To see this, let us set

$$C_g u(x) = u(g^{-1}xg), \tag{2.7.7}$$

and note that

$$\begin{aligned} \pi(C_h u) &= \int \pi(g)u(h^{-1}gh)\,dg \\ &= \int \pi(hxh^{-1})u(x)\,dx \\ &= \pi(h)\pi(u)\pi(h)^{-1}, \end{aligned} \tag{2.7.8}$$

so

$$u \ \text{ central} \Longrightarrow \pi(u)\pi(h) = \pi(h)\pi(u), \quad \forall\, h \in G. \tag{2.7.9}$$

In particular, if u is central and π^α is irreducible (on V_α, of dimension d_α), then, by Schur's lemma, $\pi^\alpha(u)$ must be scalar. Taking traces yields

$$\begin{aligned} u \ \text{ central} &\Longrightarrow \pi^\alpha(u) = \sigma^\alpha(u)I, \\ \sigma^\alpha(u) &= \frac{1}{d_\alpha}\int \chi_\alpha(g)u(g)\,dg, \end{aligned} \tag{2.7.10}$$

where $\chi_\alpha(g) = \operatorname{Tr}\pi^\alpha(g)$. Compare (2.4.8), noting that $\hat{u}(\alpha) = d_\alpha^{1/2}\overline{\pi}^\alpha(u)$.

Generalizing (2.7.7)–(2.7.8), we note that

$$G_{g,h}u(x) = u(g^{-1}xh) \Rightarrow \pi(C_{g,h}u) = \pi(g)\pi(u)\pi(h)^{-1}. \tag{2.7.11}$$

The following is a useful formula for the projection $\mathcal{P}_\alpha$ defined in (2.3.7).

Proposition 2.7.3. *For $u \in L^2(G)$,*

$$\mathcal{P}_\alpha u = d_\alpha\,\chi_\alpha * u = d_\alpha\,u * \chi_\alpha. \tag{2.7.12}$$

Proof. By Proposition 2.7.2, the last two functions in (2.7.12) are equal. By (2.3.10)–(2.3.11) we have

$$\begin{aligned} \mathcal{P}_\alpha u(g) &= d_\alpha^{1/2}\ \mathrm{Tr}\big(\hat{u}(\alpha)^t \pi^\alpha(g)\big) \\ &= d_\alpha\ \mathrm{Tr}\big(\overline{\pi}^\alpha(u)^t \pi^\alpha(g)\big). \end{aligned} \tag{2.7.13}$$

Meanwhile

$$\begin{aligned} u * \chi_\alpha(g) &= \mathrm{Tr} \int_G u(h)\pi^\alpha(h^{-1}g)\, dh \\ &= \mathrm{Tr}\Big[\int_G u(h)\pi^\alpha(h^{-1})\, dh\ \pi^\alpha(g)\Big]. \end{aligned} \tag{2.7.14}$$

Finally,

$$\overline{\pi}^\alpha(h) = \pi^\alpha(h^{-1})^t \Rightarrow \overline{\pi}^\alpha(u)^t = \int_G u(h)\pi^\alpha(h^{-1})\, dh. \tag{2.7.15}$$

Then (2.7.13)–(2.7.15) yield (2.7.12). □

We define the following involution on functions on G:

$$u^*(g) = \overline{u(g^{-1})}, \tag{2.7.16}$$

and note that

$$(u * v)^* = v^* * u^*, \tag{2.7.17}$$

and if π is a unitary representation of G,

$$\pi(u^*) = \pi(u)^*, \tag{2.7.18}$$

as is readily checked.

Given $f \in L^1(G)$, we can define the operator

$$K_f : L^2(G) \longrightarrow L^2(G), \quad K_f u(x) = f * u(x) = \int f(g)u(g^{-1}x)\, dg. \tag{2.7.19}$$

The estimate

$$\|K_f u\|_{L^2} \le \|f\|_{L^1} \|u\|_{L^2} \tag{2.7.20}$$

follows from the triangle inequality for the L^2 norm. Also, if (u, v) denotes the L^2-inner product, we have

$$\begin{aligned} (K_f u, v) &= \iint f(g)u(g^{-1}x)\overline{v(x)}\, dg\, dx \\ &= \iint f(xy^{-1})u(y)\overline{v(x)}\, dy\, dx \\ &= (u, K_{f^*} v), \end{aligned} \tag{2.7.21}$$

or

$$K_f^* = K_{f^*}. \tag{2.7.22}$$

Also it follows from (2.7.2) that

$$K_{f*g} = K_f K_g. \tag{2.7.23}$$

We can draw some parallels between K_f and $\pi(f)$ as follows. Consider the left- and right-regular representations of G on $L^2(G)$:

$$L(g)u(x) = u(g^{-1}x), \quad R(g)u(x) = u(xg). \tag{2.7.24}$$

These unitary representations are infinite dimensional, but many of the previously studied concepts apply. We have, for $f \in L^1(G),\ u \in L^2(G)$,

$$L(f)u(x) = \int f(g)u(g^{-1}x)\,dg = f * u(x) = K_f u(x), \tag{2.7.25}$$

and

$$R(f)u(x) = \int f(g)u(xg)\,dg = \int f(hx^{-1})u(h)\,dh = u * \check{f}(x), \tag{2.7.26}$$

where

$$\check{f}(g) = f(g^{-1}). \tag{2.7.27}$$

Then (2.7.23) becomes $L(f * g) = L(f)L(g)$, a result parallel to (2.7.4). Similarly one has $R(f * g) = R(f)R(g)$.

Here is another useful result.

Proposition 2.7.4. *For all $f \in L^1(G)$, we have*

$$R(g)K_f = K_f R(g), \quad \forall\, g \in G. \tag{2.7.28}$$

In addition,

$$f \quad \textit{central} \Longrightarrow L(g)K_f = K_f L(g), \quad \forall\, g \in G. \tag{2.7.29}$$

The proof involves more calculations like those done above. We leave it as an exercise.

We make note on the continuity of the representations $L(g)$ and $R(g)$. They are strongly continuous on $L^2(G)$, in the sense that

$$\forall\, u \in L^2(G),\ \ L(g)u \text{ and } R(g)u \ \text{ are continuous from } G \text{ to } L^2(G). \tag{2.7.30}$$

This continuity is obvious if $u \in C(G)$ and it follows for general u via the denseness of $C(G)$ in $L^2(G)$ and the fact that $\|L(g)u\|_{L^2} = \|u\|_{L^2} = \|R(g)u\|_{L^2}$ for all g.

We make some comments on the convolution algebra of a *finite* group G, with integral given by (2.6.1). In such a case, the convolution algebra $L^1(G)$ is also denoted $\ell^1(G)$. Another common notation for $u \in \ell^1(G)$ is

$$u = \sum_{g \in G} u(g)g. \tag{2.7.31}$$

Then convolution is given by

$$\begin{aligned} u * v &= \frac{1}{o(G)} \sum_{g,h \in G} u(g)v(h)\, gh \\ &= \frac{1}{o(G)} \sum_{g,x \in G} u(g)v(g^{-1}x)\, x, \end{aligned} \tag{2.7.32}$$

which is consistent with (2.7.1).

Exercises

1. Verify the associativity property (2.7.2), i.e.,
$$u * (v * w) = (u * v) * w,$$
for $u, v, w \in L^1(G)$.

2. Show from (2.3.10)–(2.3.11) that, for $u \in L^2(G),\ g \in G$,
$$\hat{u}(\alpha) = d_\alpha^{1/2} \overline{\pi}^\alpha(u),$$
and hence
$$\begin{aligned} \mathcal{P}_\alpha u(g) &= d_\alpha \operatorname{Tr}\big(\overline{\pi}^\alpha(u)^t \pi^\alpha(g)\big) \\ &= d_\alpha \operatorname{Tr}\big(\pi^\alpha(g)^t \overline{\pi}^\alpha(u)\big). \end{aligned}$$

3. Verify the identities in (2.7.17) and (2.7.18), when G is compact.

4. Extending (2.7.19), show that
$$f \in L^1(G) \Longrightarrow K_f : L^p(G) \to L^p(G), \quad 1 \le p < \infty,$$
and
$$\|K_f u\|_{L^p} \le \|f\|_{L^1} \|u\|_{L^p}.$$

5. Extending (2.7.25), show that
$$f \in L^1(G),\ u \in L^p(G) \Longrightarrow L(f)u = K_f u = f * u.$$

6. In the setting of Exercise 3, show that the identity
$$L(f * g) = L(f)L(g)$$
implies the associative law (2.7.2).

7. Assume G is compact. Show that
$$f \in L^2(G) \Longrightarrow K_f : L^2(G) \to C(G), \quad \sup |K_f u| \le \|f\|_{L^2} \|u\|_{L^2},$$
and, more generally, for $1 < p < \infty$,
$$f \in L^{p'}(G) \Longrightarrow K_f : L^p(G) \to C(G), \quad \sup |K_f u| \le \|f\|_{L^{p'}} \|u\|_{L^p}.$$

2.8. The Peter-Weyl theorem, II

Let G be a compact Lie group. We give G a bi-invariant Riemannian metric. Indeed, if any left-invariant Riemannian metric tensor is put on G, as discussed in §1.4, we can integrate over G its pull-back under the action of right translations to get a bi-invariant metric tensor. It is clear that the pull-back of such a metric tensor under $g \mapsto g^{-1}$ is also bi-invariant. In fact, the two agree, but rather than argue this let us just average the two, obtaining a bi-invariant metric tensor that is also invariant under $g \mapsto g^{-1}$.

If $d(x, y)$ denotes the resulting distance function between x and y in G, we note that, for any continuous function $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\psi(g) = \varphi(d(g, e)) \quad \text{is a central function on} \quad G. \tag{2.8.1}$$

Also, with ψ^* defined as in (2.7.12), we have $\psi^* = \psi$. Hence, for any unitary representation π of G, $\pi(\psi)$ is self-adjoint.

Let us assume φ is ≥ 0, Lipschitz, and satisfies $\varphi(s) = 1$ for $|s| \leq 1/2$, $\varphi(s) = 0$ for $|s| \geq 1$, and let us set

$$\psi_\nu(g) = \varphi(\nu d(g, e)), \quad \nu \geq 1. \tag{2.8.2}$$

Then $\psi_\nu \in \mathrm{Lip}(G)$, and it is supported on $B_{1/\nu}(e)$, where $B_r(h) = \{g \in G : d(g, h) \leq r\}$. Set

$$\Psi_\nu(g) = A_\nu^{-1} \psi_\nu(g), \quad A_\nu = \int \psi_\nu(g)\, dg, \tag{2.8.3}$$

so $\Psi_\nu \in \mathrm{Lip}(G)$ is also supported in $B_{1/\nu}(e)$ and $\int \Psi_\nu(g)\, dg = 1$. Now set

$$\Phi_\nu(g) = \Psi_\nu * \Psi_\nu(g). \tag{2.8.4}$$

Then $\Phi_\nu \in \mathrm{Lip}(G)$ is supported in $B_{2/\nu}(e)$ and $\int \Phi_\nu(g)\, dg = 1$. Now define the convolution operators

$$C_\nu u = \Psi_\nu * u, \quad K_\nu u = \Phi_\nu * u. \tag{2.8.5}$$

Proposition 2.8.1. *The operators C_ν and K_ν are approximate identities. That is, as $\nu \to \infty$,*

$$u \in C(G) \Longrightarrow C_\nu u \to u \quad \text{and} \quad K_\nu u \to u \quad \text{uniformly.} \tag{2.8.6}$$

Also

$$u \in L^2(G) \Longrightarrow C_\nu u \to u \quad \text{and} \quad K_\nu u \to u \quad \text{in} \quad L^2\text{-norm.} \tag{2.8.7}$$

Proof. We note that, for every $g \in G$, $C_\nu u(g)$ is a weighted average of u over the set $B_{1/\nu}(g)$ and $K_\nu u(g)$ is a weighted average of u over the set $B_{2/\nu}(g)$.

Thus if u has the modulus of continuity ω, i.e., $|u(g) - u(h)| \le \omega(d(g,h))$, then

$$\|C_\nu u - u\|_{\sup} \le \omega(1/\nu), \quad \|K_\nu u - u\|_{\sup} \le \omega(2/\nu). \tag{2.8.8}$$

This gives (2.8.6). The result (2.8.7) for C_ν follows from

$$\|C_\nu u\|_{L^2} \le \|u\|_{L^2}, \quad \|u\|_{L^2} \le \|u\|_{\sup}, \quad C(G) \text{ dense in } L^2(G). \tag{2.8.9}$$

In fact, given $u \in L^2(G)$ and $\varepsilon > 0$, pick $v \in C(G)$ such that $\|u - v\|_{L^2} < \varepsilon$. Pick N such that $\nu \ge N \Rightarrow \|C_\nu u - u\|_{\sup} < \varepsilon$. Then, for $\nu \ge N$,

$$\begin{aligned} \|C_\nu u - u\|_{L^2} &= \|C_\nu u - C_\nu v + C_\nu v - v + v - u\|_{L^2} \\ &\le \|C_\nu(u - v)\|_{L^2} + \|C_\nu v - v\|_{\sup} + \|v - u\|_{L^2} \\ &< 3\varepsilon, \end{aligned} \tag{2.8.10}$$

giving (2.8.7) for $C_\nu u$. The proof for $K_\nu u$ is similar. □

For other properties of C_ν and K_ν on $L^2(G)$, we note from (2.7.22)–(2.7.23) that

$$C_\nu = C_\nu^*, \quad K_\nu = C_\nu^2; \tag{2.8.11}$$

hence

$$(K_\nu u, u) = \|C_\nu u\|_{L^2}^2 \ge 0, \quad \forall\, u \in L^2(G), \tag{2.8.12}$$

i.e., K_ν is a positive semi-definite self-adjoint operator on $L^2(G)$.

Proposition 2.8.2. *For each ν, C_ν and K_ν are compact operators on $L^2(G)$.*

There are several ways to prove this. The integral kernel of C_ν is Lipschitz on $G \times G$, hence square-integrable, so C_ν is a Hilbert-Schmidt operator, hence compact. Also, for each ν,

$$C_\nu : L^2(G) \longrightarrow \mathrm{Lip}(G). \tag{2.8.13}$$

Now $\mathrm{Lip}(G) \hookrightarrow C(G)$ is compact, by Ascoli's theorem, while $C(G) \hookrightarrow L^2(G)$ is continuous.

Now if K is a compact self-adjoint operator on $L^2(G)$, then the eigenspaces E_λ corresponding to nonzero eigenvalues are all finite dimensional, and these spaces together with $\operatorname{Ker} K$ span $L^2(G)$. The following result will help us prove the Peter-Weyl theorem for general compact Lie groups. As usual, let $\{\pi^\alpha : \alpha \in \mathcal{I}\}$ be a maximal set of inequivalent irreducible unitary representations of G, acting on spaces V_α.

Proposition 2.8.3. *Let $K : L^2(G) \to L^2(G)$ be compact and self-adjoint, and assume*

$$KR(g) = R(g)K, \quad \forall\, g \in G, \tag{2.8.14}$$

where $R(g)$ denotes the right-regular representation of G on $L^2(G)$, given by (2.7.24). *Let E_λ be the eigenspace of K for some nonzero eigenvalue λ. Then each $u \in E_\lambda$ is a finite linear combination of matrix entries π^α_{jk}.*

Proof. By (2.8.14), $R(g) : E_\lambda \to E_\lambda$. We know this representation decomposes into irreducibles; $E_\lambda = V_1 \oplus \cdots \oplus V_K$. Say $R(g)|_{V_\ell} \approx \pi^\alpha$, so there is a unitary map $U : V_\ell \to V_\alpha$ intertwining these representations. Say an orthonormal basis $\{u_j\}$ of V_ℓ corresponds to an orthonormal basis $\{e_j\}$ of V_α, with respect to which π^α has matrix entries π^α_{jk}. Then, for all $g \in G$, all k,

$$R(g)u_k = U^{-1}\pi^\alpha(g)e_k = U^{-1}\sum_j \pi^\alpha_{jk}(g)e_j, \tag{2.8.15}$$

or

$$u_k(xg) = \sum_j u_j(x)\,\pi^\alpha_{jk}(g), \quad \forall\; x, g \in G. \tag{2.8.16}$$

Taking $x = e$ gives

$$u_k(g) = \sum_j u_j(e)\,\pi^\alpha_{jk}(g), \tag{2.8.17}$$

proving the proposition. □

We are now ready for our second proof of the Peter-Weyl theorem, one that works for all compact Lie groups G.

Proposition 2.8.4. *Let G be a compact Lie group, $\{\pi^\alpha : \alpha \in \mathcal{I}\}$ a maximal set of irreducible unitary representations of G, on vector spaces V_α, of dimension d_α. Then $\{d_\alpha^{1/2}\,\pi^\alpha_{jk} : \alpha \in \mathcal{I}, 1 \le j, k \le d_\alpha\}$ is an orthonormal basis of $L^2(G)$.*

Proof. It suffices to prove the span of π^α_{jk} is dense in $L^2(G)$. Suppose we have a positive, self-adjoint, compact operator $K : L^2(G) \to L^2(G)$, satisfying (2.8.14), and suppose $\operatorname{Ker} K = 0$. Then the span of the eigenspaces of K is dense in $L^2(G)$, and the result then follows from Proposition 2.8.3. The task that remains is to construct such an injective operator K.

To this end, set

$$K = \sum_{\nu \ge 0} 2^{-\nu} K_\nu, \tag{2.8.18}$$

with K_ν as in (2.8.4)–(2.8.5). Then K is a norm limit of compact operators, hence compact, and also clearly positive, self-adjoint. By Proposition 2.7.4 each K_ν has the property (2.8.14), hence so does K.

Finally, we show K is injective. Suppose $u \in \operatorname{Ker} K$. Thus $0 = (Ku, u) = \sum 2^{-\nu}(K_\nu u, u)$. Since each K_ν is ≥ 0, we must have $K_\nu u = 0$ for each ν. But Proposition 2.8.1 implies $K_\nu u \to u$ in L^2-norm, so $u = 0$, and we are done. ☐

Let us again denote by $\mathcal{V}_\alpha$ the linear span of $\{\pi^\alpha_{jk} : 1 \le j, k \le d_\alpha\}$, and $\mathcal{P}_\alpha$ the orthogonal projection of $L^2(G)$ on $\mathcal{V}_\alpha$. We want to show that the span of $\{\mathcal{V}_\alpha\}$ is dense in $C(G)$, for any compact G (not yet knowing the properties used in the demonstration in §2.3 under the hypothesis $G \subset U(N)$). The following will be useful for this. Here, K_u denotes a convolution operator, as in (2.7.19).

Proposition 2.8.5. *For any $u \in L^1(G)$,*

$$K_u : \mathcal{V}_\alpha \longrightarrow \mathcal{V}_\alpha. \tag{2.8.19}$$

Furthermore, for any $v \in L^2(G)$,

$$\mathcal{P}_\alpha(u * v) = u * (\mathcal{P}_\alpha v). \tag{2.8.20}$$

Proof. Recall from Proposition 2.7.3 that

$$\mathcal{P}_\alpha v = d_\alpha\, \chi_\alpha * v. \tag{2.8.21}$$

Hence, by Proposition 2.7.2,

$$\begin{aligned} \mathcal{P}_\alpha(u * v) &= d_\alpha\, \chi_\alpha * u * v \\ &= d_\alpha\, u * \chi_\alpha * v, \end{aligned} \tag{2.8.22}$$

and we have (2.8.19)–(2.8.20). ☐

To proceed, we know that $\mathcal{I}$ is countable (i.e., $L^2(G)$ is separable), so make an ordering so $\mathcal{I} \approx \mathbb{Z}^+$, and set

$$\Pi_N\, f = \sum_{|\alpha| \le N} \mathcal{P}_\alpha f. \tag{2.8.23}$$

Then the content of Proposition 2.8.4 is that, as $N \to \infty$,

$$f \in L^2(G) \Longrightarrow \Pi_N f \to f \ \text{ in } \ L^2\text{-norm}. \tag{2.8.24}$$

Here is a result on uniform convergence.

Proposition 2.8.6. *Assume $f \in C(G)$ has the form*

$$f = u * v, \quad u, v \in L^2(G). \tag{2.8.25}$$

Then, as $N \to \infty$,

$$\Pi_N f \to f \ \text{ uniformly on } \ G. \tag{2.8.26}$$

Proof. It follows from (2.8.20) that

$$\Pi_N f = u * (\Pi_N v). \tag{2.8.27}$$

Now convolution yields a continuous bilinear map

$$L^2(G) \times L^2(G) \longrightarrow C(G), \tag{2.8.28}$$

so using $\Pi_N v \to v$ in L^2-norm in (2.8.27) yields (2.8.26). □

From Proposition 2.8.1 it follows that the set of functions of the form (2.8.25) is dense in $C(G)$, so we have:

Corollary 2.8.7. *The linear span $\mathcal{L}$ of $\{\pi^\alpha_{jk} : \alpha \in \mathcal{I}, 1 \le j, k \le d_\alpha\}$ is dense in $C(G)$.*

We use this to prove:

Proposition 2.8.8. *Every compact Lie group has a faithful finite-dimensional unitary representation.*

To see this, let

$$K_\alpha = \{g \in G : \pi^\alpha(g) = I\}. \tag{2.8.29}$$

We want to show that there is a finite set $\mathcal{S} \subset \mathcal{I}$ such that $\cap_{\alpha\in\mathcal{S}} K_\alpha = \{e\}$. Then $\oplus_{\alpha\in\mathcal{S}}\pi^\alpha$ provides such a representation.

By Corollary 2.8.7, for any $g \in G$, $g \neq e$, there exists $\alpha \in \mathcal{I}$ such that $\pi^\alpha(g) \neq I$. Otherwise, $\pi^\alpha(gx) = \pi^\alpha(x)$ for all $x \in G$, $\alpha \in \mathcal{I}$, hence $u(gx) = u(x)$ for all $x \in G$, $u \in \mathcal{L}$, which forces $g = e$. We use this as follows. Take any open neighborhood $\mathcal{O}$ of e in G. Then $G \setminus \mathcal{O}$ is compact. By the reasoning above, for each $g \in G \setminus \mathcal{O}$ there exists $\alpha \in \mathcal{I}$ and a neighborhood U_g of g such that $\pi^\alpha(h) \neq I$ for all $h \in U_g$. Since any open cover of $G \setminus \mathcal{O}$ has a finite subcover, we have the following.

For any open neighborhood $\mathcal{O}$ of e, there is a finite set $\mathcal{S} \subset \mathcal{I}$ such that

$$\bigcap_{\alpha\in\mathcal{S}} K_\alpha \subset \mathcal{O}. \tag{2.8.30}$$

Thus the proof of Proposition 2.8.8 is completed by the following assertion.

Proposition 2.8.9. *If G is a Lie group, then there is an open $\mathcal{O} \ni e$ such that if K is a subgroup of G and $K \subset \mathcal{O}$, then $K = \{e\}$.*

Proof. Let $\mathrm{Sq} : G \to G$ be defined as $\mathrm{Sq}(g) = g^2$. In §3.1 we will prove

$$D\,\mathrm{Sq}(e) = 2I. \tag{2.8.31}$$

In other words, if we take a coordinate system on a neighborhood U of e, in which e corresponds to $0 \in \mathbb{R}^n$, then

$$\mathrm{Sq}(x) = 2x + R(x), \quad |R(x)| \le C|x|^2. \tag{2.8.32}$$

We use the Euclidean norm $|x|$. It follows that, in this coordinate system,

$$|x| < \frac{1}{2C} \Longrightarrow |\operatorname{Sq}(x)| > \frac{3}{2}|x|. \tag{2.8.33}$$

Thus if $\mathcal{O} = \{x : |x| < 1/4C\}$, we see that the orbit of any $g \neq e$ (given here by $e = 0$) under Sq cannot remain in $\mathcal{O}$. $\square$

We can also use approximate identities to study the smoothness of representations of a (not necessarily compact) Lie group G. Let us consider the case of a representation π of G on a Banach space V. We assume π is *strongly continuous*, i.e., for each $u \in V$, $\pi(g)u$ is a continuous function of g with values in V. As we have seen, the regular representations L and R of G on $L^2(G)$ have this property (when G has, respectively, left-invariant or right-invariant Haar measure). It is a consequence of the uniform boundedness principle that the operator norm $\|\pi(g)\|$ is bounded on compact subsets of G. If f is compactly supported and integrable on G, we can define $\pi(f)$ as before:

$$\pi(f)u = \int_G f(g)\pi(g)u\,dg, \quad \pi(f) : V \to V. \tag{2.8.34}$$

Here we will use left-invariant Haar measure. Note that, for any $h \in G$,

$$\pi(h)\pi(f)u = \int f(h^{-1}g)\pi(g)u\,dg. \tag{2.8.35}$$

From this it is easy to see that, for all $u \in V$,

$$f \in C_0^\infty(G) \Rightarrow \pi(h)\pi(f)u \ \text{ is a smooth } V\text{-valued function of } h, \tag{2.8.36}$$

Generally we say $v \in V$ is a smooth vector for the representation π if $\pi(g)v$ is a smooth function of g with values in V.

Now we can construct a sequence $f_\nu \in C_0^\infty(G)$, each integrating to 1, supported on progressively smaller neighborhoods of the identity element e, and (as in the proof of Proposition 2.8.1) we have

$$\pi(f_\nu)u \longrightarrow u \ \text{ in } \ V, \quad \forall\, u \in V. \tag{2.8.37}$$

We hence have:

Proposition 2.8.10. *If π is a strongly continuous representation of a Lie group G on a Banach space V, then the space $\mathcal{V}_0$ of smooth vectors is a dense linear subspace of V. In particular, if V is finite dimensional, then all vectors in V are smooth.*

Remark. Complementing Propositions 2.8.8 and 2.8.10, we can show that if π is a faithful finite dimensional representation of G on V, then $\pi : G \to \mathcal{L}(V)$ is a smooth embedding. See the exercises of Chapter 3, §3.3 for details.

We close this section with the following corollary to the Peter-Weyl theorem, which will prove useful later.

Proposition 2.8.11. *If G_1 and G_2 are two compact Lie groups, then the irreducible unitary representations of $G = G_1 \times G_2$ are, up to unitary equivalence, precisely those of the form*

$$\pi(g) = \pi_1(g_1) \otimes \pi_2(g_2), \tag{2.8.38}$$

where $g = (g_1, g_2) \in G$, and π_j is a general irreducible unitary representation of G_j.

Proof. Given π_j irreducible unitary representations of G_j, the unitarity of (2.8.38) is clear, and irreducibility can be established as follows. We have $\chi_\pi(g_1, g_2) = \chi_{\pi_1}(g_1)\chi_{\pi_2}(g_2)$ (cf. Exercise 1 of §2.1), and hence

$$\iint\limits_{G_1 \times G_2} |\chi_\pi(g_1, g_2)|^2 \, dg_1 \, dg_2 = 1. \tag{2.8.39}$$

It remains to prove the completeness of the set of such representations. For this, it suffices to show that the matrix entries of such representations have dense linear span in $L^2(G_1 \times G_2)$. This follows from the general elementary fact that products $\varphi_j(g_1)\psi_k(g_2)$ of orthonormal bases $\{\varphi_j\}$ of $L^2(G_1)$ and $\{\psi_k\}$ of $L^2(G_2)$ form an orthonormal basis of $L^2(G_1 \times G_2)$. $\square$

Exercises

1. From the arguments proving Proposition 2.8.5, show that also

$$\mathcal{P}_\alpha(u * v) = (\mathcal{P}_\alpha u) * v = (\mathcal{P}_\alpha u) * (\mathcal{P}_\alpha v).$$

2. Deduce from Exercise 1 that

$$\begin{aligned} \|\mathcal{P}_\alpha(u * v)\|_{L^\infty} &\le \|\mathcal{P}_\alpha u\|_{L^2} \|\mathcal{P}_\alpha v\|_{L^2} \\ &\le \frac{t}{2}\|\mathcal{P}_\alpha u\|_{L^2}^2 + \frac{1}{2t}\|\mathcal{P}_\alpha v\|_{L^2}^2, \quad \forall\, t > 0. \end{aligned}$$

Hence, for $u, v \in L^2(G)$,

$$\begin{aligned} \sum_\alpha \|\mathcal{P}_\alpha(u * v)\|_{L^\infty} &\le \inf_{t>0} \frac{t}{2}\|u\|_{L^2}^2 + \frac{1}{2t}\|v\|_{L^2}^2 \\ &= \|u\|_{L^2}\|v\|_{L^2}. \end{aligned}$$

Note how this strengthens Proposition 2.8.6.

2.9. Denseness of Span $\{\pi^\alpha_{jk}\}$ in $C^m(G)$

Our goal here is to establish the following improvement of Corollary 2.8.7.

Proposition 2.9.1. *Let G be a compact Lie and $\{\pi^\alpha : \alpha \in \mathcal{I}\}$ a complete set of irreducible unitary representations of G, on spaces V_α, of dimension d_α. Then*

$$\mathcal{L} = \mathrm{Span}\{\pi^\alpha_{jk} : \alpha \in \mathcal{I},\ 1 \le j, k \le d_\alpha\} \tag{2.9.1}$$

is dense in $C^m(G)$, for each $m \in \mathbb{N}$.

We obtain this by tweaking arguments from §2.8. To begin, we define ψ_ν and Ψ_ν as in (2.8.2)–(2.8.3), requiring that ψ be C^∞, as well as $\psi(s) = 1$ for $|s| \le 1/2$ and 0 for $|s| \ge 1$. Then, for ν sufficiently large, say for $\nu \ge M$,

$$\Psi_\nu \in C^\infty_0(B_{2/\nu}(e)). \tag{2.9.2}$$

As in (2.8.5), we set

$$C_\nu u = \Psi_\nu * u. \tag{2.9.3}$$

We have

$$C_\nu : L^2(G) \longrightarrow C^m(G), \tag{2.9.4}$$

for each $m \in \mathbb{N}, \nu \ge M$. We can apply Proposition 2.8.5 to deduce that

$$\Pi_N C_\nu f = C_\nu \Pi_N f, \tag{2.9.5}$$

where

$$\Pi_N f = \sum_{|\alpha| \le N} \mathcal{P}_\alpha f. \tag{2.9.6}$$

We hence have, for each $m \in \mathbb{N}, \nu \ge M$,

$$\begin{aligned} f \in L^2(G) &\Longrightarrow \Pi_N f \to f \text{ in } L^2,\ \text{as } N \to \infty,\\ &\Longrightarrow C_\nu \Pi_N f \to C_\nu f \text{ in } C^m,\ \text{as } N \to \infty,\\ &\Longrightarrow \Pi_N C_\nu f \to C_\nu f \text{ in } C^m,\ \text{as } N \to \infty. \end{aligned} \tag{2.9.7}$$

The next ingredient in the proof of Proposition 2.9.1 is the following.

Lemma 2.9.2. *Given $m \in \mathbb{N},\ f \in C^m(G)$,*

$$C_\nu f \longrightarrow f \text{ in } C^m(G),\ \text{as } \nu \to \infty. \tag{2.9.8}$$

Proof. Best carried out with tools developed in Chapter 3. ☐

Accepting this, we proceed as follows. Take $f \in C^m(G)$ and pick $\varepsilon > 0$. Then, fix $\nu \in \mathbb{N}$ such that

$$\|C_\nu f - f\|_{C^m} < \varepsilon. \tag{2.9.9}$$

With such ν fixed, we have

$$\|\Pi_N(C_\nu f) - C_\nu f\|_{C^m} \longrightarrow 0, \quad \text{as } N \to \infty,$$

hence there exists M_1 such that

$$\|\Pi_N(C_\nu f) - C_\nu f\|_{C^m} < \varepsilon, \quad \text{for} \quad N \geq M_1, \tag{2.9.10}$$

so

$$\|\Pi_N(C_\nu f) - f\|_{C^m} < 2\varepsilon, \quad \text{for} \quad N \geq M_1. \tag{2.9.11}$$

Since $\Pi_N(C_\nu f) \in \mathcal{L}$, we have Proposition 2.9.1.

Chapter 3

Lie algebras

A very important tool in the study of Lie groups is the concept of a Lie algebra. Generally, if G is a Lie group, its Lie algebra $\mathfrak{g}$ can be defined as the space of left-invariant vector fields on G. If X and Y are two such vector fields, so is the commutator $[X, Y]$, and this (also called the Lie bracket) defines the Lie algebra structure. We introduce the basic notions in §3.1. We show that if $X \in \mathfrak{g}$, the flow $\mathcal{F}_X^t$ it generates has the property that

$$\mathcal{F}_X^t(g) = g\gamma_X(t),$$

where $\gamma_X(t) = \mathcal{F}_X^t(e)$ is a one-parameter subgroup, satisfying $\gamma_X'(0) = X(e)$. We also introduce the exponential map

$$\operatorname{Exp} : \mathfrak{g} \longrightarrow G, \quad \operatorname{Exp}(tX) = \gamma_X(t).$$

If G is a matrix group, one can also look at the tangent space to G at the identity element, a space of matrices closed under the matrix commutator. In this case, we show in §3.2 that the two notions of the Lie algebra of G are naturally isomorphic. Under this isomorphism, the exponential map introduced in §3.1 is taken to the matrix exponential.

In §3.3 we study how a representation π of G on a vector space V gives rise to a "derived" Lie algebra representation $d\pi$ of $\mathfrak{g}$ on V, satisfying

$$e^{t\, d\pi(X)} = \pi(\operatorname{Exp} tX).$$

The natural "adjoint" representation Ad of G on $\mathfrak{g}$, and its derived representation ad of $\mathfrak{g}$ on itself, studied in §3.4, provide important tools for the study of the internal structure of G. Key identities here are

$$\operatorname{Ad}(\operatorname{Exp} X) = e^{\operatorname{ad} X}, \quad \operatorname{ad}(X)Y = [X, Y].$$

In §3.5, we derive a formula for the product on G in terms of the action of ad on $\mathfrak{g}$, known as the Campbell-Hausdorff formula. This takes the form

$$(\operatorname{Exp} X)(\operatorname{Exp} Y) = \operatorname{Exp} \mathcal{C}(X, Y),$$

for X and Y in a suitable neighborhood of 0 in $\mathfrak{g}$, where

$$\mathcal{C}(X, Y) = X + \int_0^1 \Psi(e^{\operatorname{ad} X} e^{t \operatorname{ad} Y}) Y \, dt,$$

with

$$\Psi(\zeta) = \frac{\zeta \log \zeta}{\zeta - 1}.$$

One consequence of this result is that G, *a priori* assumed to be a C^∞ manifold, automatically carries a uniquely defined real analytic structure. Another consequence, established in §3.6, is that each Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{h}$, between Lie algebras of Lie groups G and H, gives rise to a Lie group homomorphism $\rho : G \to H$, provided G is simply connected.

This chapter concludes with an introduction of the universal enveloping algebra of $\mathfrak{g}$ in §3.7 and a result about its structure known as the Poincaré-Birkhoff-Witt theorem in §3.8.

3.1. Lie algebras of general Lie groups

Let G be a Lie group, T_eG the tangent space to the identity element. For each $X_0 \in T_eG$, there is a unique left-invariant vector field X on G such that $X(e) = X_0$. Here to say X is left-invariant is to say

$$DL_g(h)\,X(h) = X(gh), \tag{3.1.1}$$

where

$$L_g : G \to G, \quad L_g x = gx, \quad DL_g : T_hG \to T_{gh}G. \tag{3.1.2}$$

In fact, such X is uniquely specified by

$$X(g) = DL_g(e)\,X_0. \tag{3.1.3}$$

To justify this, we claim that one can deduce from the chain rule that if $X_0 \in T_eG$ and $X(g)$ is defined by (3.1.3), then (3.1.1) holds. In detail, if (3.1.3) holds for each g, then

$$\begin{aligned} DL_g(h)X(h) &= DL_g(h)DL_h(e)X_0 \\ &= D(L_g \circ L_h)(e)X_0 \\ &= DL_{gh}(e)X_0, \end{aligned}$$

giving (3.1.1).

We denote by $\mathfrak{g}$ the set of left-invariant vector fields on G, so $\mathfrak{g} \approx T_eG$ as a linear space.

A vector field $X \in \mathfrak{g}$ generates a flow $\mathcal{F}_X^t$ on G; cf. §A.4. The general theory of ODE gives us a local flow, but in fact calculations below will yield a global flow when $X \in \mathfrak{g}$. The defining property of $\mathcal{F}_X^t g$ is

$$\mathcal{F}_X^0 g = g, \quad \frac{d}{dt}\mathcal{F}_X^t g = X(\mathcal{F}_X^t g). \tag{3.1.4}$$

The following property helps reveal the nature of this flow.

Proposition 3.1.1. *Given $X \in \mathfrak{g}$, $g, h \in G$,*

$$g\,\mathcal{F}_X^t h = \mathcal{F}_X^t(gh). \tag{3.1.5}$$

Proof. Denote the left side of (3.1.5) by $x(t)$ and the right side by $y(t)$. Then $x(0) = y(0) = gh$. The result (3.1.4) easily gives $y'(t) = X(y)$. Meanwhile,

$$x'(t) = DL_g(\mathcal{F}_X^t h)\,X(\mathcal{F}_X^t h) = X(x),$$

the first identity by the chain rule and the second by (3.1.1). Uniqueness for ODE then yields $x(t) \equiv y(t)$. $\square$

Let us set

$$\gamma_X(t) = \mathcal{F}_X^t e, \tag{3.1.6}$$

for $X \in \mathfrak{g}$. Then taking $h = e$ in (3.1.5) gives

$$\mathcal{F}_X^t g = g\,\gamma_X(t). \tag{3.1.7}$$

The following is a key group property.

Proposition 3.1.2. *For $X \in \mathfrak{g}$,*

$$\gamma_X(s+t) = \gamma_X(s)\gamma_X(t). \tag{3.1.8}$$

Proof. This follows from (3.1.7) plus the fact that $\mathcal{F}_X^{s+t} = \mathcal{F}_X^t \circ \mathcal{F}_X^s$ (a general property of flows). In detail,

$$\begin{aligned}\gamma_X(s+t) &= \mathcal{F}_X^t(\mathcal{F}_X^s e) = \mathcal{F}_X^t \gamma_X(s) \\ &= \gamma_X(s)\gamma_X(t),\end{aligned}$$

the second identity by (3.1.6) and the third by (3.1.7). □

We say that $\gamma_X(t)$ is a one-parameter subgroup of G. Note that (3.1.8) implies $\gamma_X(t)$ is well defined for all $t \in \mathbb{R}$, hence, by (3.1.7), $\mathcal{F}_X^t$ is well defined for all $t \in \mathbb{R}$, when $X \in \mathfrak{g}$. We can characterize $\gamma_X(t)$ as follows.

Proposition 3.1.3. *The curve $\gamma_X(t)$ is the unique smooth one-parameter subgroup of G satisfying $\gamma_X'(0) = X(e)$.*

Proof. In fact, if $\gamma(t)$ is another such one-parameter subgroup and we set $\mathcal{F}^t g = g\gamma(t)$, we see that

$$\begin{aligned}\frac{d}{dt}\mathcal{F}^t g &= \frac{d}{ds} g\gamma(t+s)\big|_{s=0} \\ &= \frac{d}{ds} g\gamma(t)\gamma(s)\big|_{s=0} \\ &= DL_{g\gamma(t)}(e)\,X(e) \\ &= X(\mathcal{F}^t g),\end{aligned} \tag{3.1.9}$$

so by uniqueness for ODE, $\mathcal{F}^t \equiv \mathcal{F}_X^t$. □

We pause to prove (2.8.31), i.e.,

$$\mathrm{Sq}(x) = x\cdot x \Longrightarrow D\,\mathrm{Sq}(e)\,X = 2X, \quad \forall\ X \in T_eG. \tag{3.1.10}$$

To see this, since we know Sq is smooth, it suffices to note that

$$\mathrm{Sq}(\gamma_X(t)) = \gamma_X(2t) \Longrightarrow \frac{d}{dt}\,\mathrm{Sq}(\gamma_X(t))\big|_{t=0} = 2X. \tag{3.1.11}$$

Thus the last detail in the proof of Proposition 2.8.9 is taken care of.

We now define the *exponential map*:

$$\operatorname{Exp} : \mathfrak{g} \to G, \quad \operatorname{Exp} X = \gamma_X(1) = \mathcal{F}_X^1 e. \tag{3.1.12}$$

Results covered in §A.7 imply Exp is C^∞. In view of the uniqueness result from Proposition 3.1.3, we have $\gamma_{sX}(t) = \gamma_X(st)$, and hence

$$\operatorname{Exp} tX = \gamma_X(t). \tag{3.1.13}$$

Also the unique characterization of $\gamma_X(t)$ given above implies the following. If $G = \mathrm{Gl}(n, \mathbb{F})$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$), or if G is a matrix group, such as $\mathrm{O}(n)$ or $\mathrm{U}(n)$, then, with $X \in T_e G \approx \mathfrak{g}$,

$$\operatorname{Exp} tX = e^{tX}, \tag{3.1.14}$$

the right side denoting the matrix exponential. See §3.2 for more.

Note that (3.1.13) implies

$$D\operatorname{Exp}(0) : T_e G \to T_e G \ \text{ is the identity map.} \tag{3.1.15}$$

Hence, by the inverse function theorem, we have the following.

Proposition 3.1.4. Exp *is a diffeomorphism from some open neighborhood* $\mathcal{O}$ *of* 0 *in* $\mathfrak{g}$ *onto a neighborhood* U *of* e *in* G.

This provides what is known as an exponential coordinate system.

A vector field on G yields a first-order differential operator on smooth functions on G, via

$$Xf(x) = \frac{d}{dt} f(\mathcal{F}_X^t x)\big|_{t=0}. \tag{3.1.16}$$

See §A.4 for more details. If $X \in \mathfrak{g}$, then, by (3.1.7), we can write this as

$$Xf(x) = \frac{d}{dt} f(x\gamma_X(t))\big|_{t=0}. \tag{3.1.17}$$

It then follows that, when X is a vector field on G, then X is left-invariant (i.e., $X \in \mathfrak{g}$) if and only if

$$XL(g)f = L(g)Xf, \quad \forall\, g \in G,\ f \in C^\infty(G), \tag{3.1.18}$$

where, as usual,

$$L(g)f(x) = f(g^{-1}x). \tag{3.1.19}$$

In fact, the map $f(x) \mapsto f(x\gamma_X(t))$ commutes with $L(g)$, so any vector field X of the form (3.1.17), i.e., any $X \in \mathfrak{g}$, commutes with $L(g)$. For the converse, note that if X is a vector field that commutes with $L(g)$ for all $g \in G$, then, for each $f \in C^\infty(G)$,

$$\begin{aligned} Df(g)X(g) = Xf(g) &= L(g^{-1})Xf(e) \\ &= XL(g^{-1})f(e) = X(f \circ L_g)(e) \\ &= Df(g)\, DL_g(e)X(e), \end{aligned}$$

so $X(g) = DL_g(e)X(e)$, hence $X \in \mathfrak{g}$.

If X and Y have the property (3.1.18), then so does the commutator (or Lie bracket)

$$[X, Y] = XY - YX, \tag{3.1.20}$$

i.e., $X, Y \in \mathfrak{g} \Rightarrow [X, Y] \in \mathfrak{g}$. This structure makes $\mathfrak{g}$ a Lie algebra.

In general, a Lie algebra is a vector space $\mathfrak{g}$ on which there is a bilinear map

$$\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \quad (X, Y) \mapsto [X, Y], \tag{3.1.21}$$

satisfying two identities. One is

$$[X, Y] = -[Y, X]. \tag{3.1.22}$$

The other, known as the Jacobi identity, can be expressed as follows. Given $X \in \mathfrak{g}$, define the linear map $\operatorname{ad} X : \mathfrak{g} \to \mathfrak{g}$ by

$$\operatorname{ad} X(Y) = [X, Y]. \tag{3.1.23}$$

Then the Jacobi identity is

$$\operatorname{ad}[X, Y] = [\operatorname{ad} X, \operatorname{ad} Y], \tag{3.1.24}$$

where

$$[\operatorname{ad} X, \operatorname{ad} Y] = (\operatorname{ad} X)(\operatorname{ad} Y) - (\operatorname{ad} Y)(\operatorname{ad} X). \tag{3.1.25}$$

Plugging in the definition (3.1.23), one can write out the Jacobi identity as

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0. \tag{3.1.26}$$

It is routine to show that the commutator (3.1.20) satisfies (3.1.22) and (3.1.26).

The fact that each $X \in \mathfrak{g}$ generates a one-parameter subgroup of G has the following generalization, to a fundamental result of S. Lie. Suppose G is a Lie group with Lie algebra $\mathfrak{g}$, and suppose $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. That is, $\mathfrak{h}$ is a linear subspace of $\mathfrak{g}$ and $X_j \in \mathfrak{h} \Rightarrow [X_1, X_2] \in \mathfrak{h}$. By Frobenius' theorem (discussed in §A.6), through each point $p \in G$ there is a smooth manifold M_p, of dimension $k = \dim \mathfrak{h}$, which is an integral manifold for $\mathfrak{h}$ (i.e., $\mathfrak{h}$ spans the tangent space of M_p at each $q \in M_p$). We can take M_p to be the maximal such (connected) manifold, and then it is unique. Let H be the maximal integral manifold of $\mathfrak{h}$ containing the identity element.

Proposition 3.1.5. *If $\mathfrak{h}$ is a subalgebra of the Lie algebra $\mathfrak{g}$ of G, then the integral manifold H of $\mathfrak{h}$ through e is a subgroup of G.*

Proof. Take $h_0 \in H$ and consider $H_0 = h_0^{-1}H$. Clearly $e \in H_0$. By left-invariance, H_0 is also an integral manifold of $\mathfrak{h}$, so $H_0 \subset H$. This shows that $h_0, h_1 \in H \Rightarrow h_0^{-1}h_1 \in H$, so H is a group. $\square$

The next result gives one sense in which the Lie algebra $\mathfrak{g}$ of a Lie group G generates G, at least when G is connected.

Proposition 3.1.6. *Let G be a connected Lie group, $g \in G$. Then there exist $X_1, \dots, X_K \in \mathfrak{g}$ such that*

$$g = (\operatorname{Exp} X_1) \cdots (\operatorname{Exp} X_K). \tag{3.1.27}$$

Proof. Put a left-invariant Riemannian metric on G. By (3.1.15) and the inverse function theorem, there exists $\delta > 0$ such that, for $h \in G$,

$$\operatorname{dist}(h, e) < \delta \Longrightarrow h = \operatorname{Exp} Y(h), \quad \text{for some} \quad Y(h) \in \mathfrak{g}. \tag{3.1.28}$$

Given any $g \in G$, pick a smooth path $\sigma(t)$ from e to g, and find $g_0, \dots, g_K$ on the path such that

$$g_0 = e, \dots, g_K = g, \quad \operatorname{dist}(g_{j-1}, g_j) < \delta. \tag{3.1.29}$$

Then

$$g = (g_0^{-1} g_1)(g_1^{-1} g_2) \cdots (g_{K-2}^{-1} g_{K-1})(g_{K-1}^{-1} g_K), \tag{3.1.30}$$

and $\operatorname{dist}(g_{j-1}^{-1} g_j, e) = \operatorname{dist}(g_j, g_{j-1}) < \delta$, so each $g_{j-1}^{-1} g_j = \operatorname{Exp}(X_j)$ for some $X_j \in \mathfrak{g}$. $\square$

Exercises

In Exercises 1–3, we identify $\mathfrak{g}$ with T_eG, via the isomorphism described in the first paragraph of this section.

1. Define
$$M : G \times G \longrightarrow G, \quad M(g,h) = gh.$$
We have
$$DM(e,e) : \mathfrak{g} \oplus \mathfrak{g} \longrightarrow \mathfrak{g}, \quad \mathfrak{g} \approx T_eG.$$
Show that
$$DM(e,e)\begin{pmatrix} X \\ Y \end{pmatrix} = X + Y.$$

2. We can write $\mathrm{Sq}(g) = g^2$ as $\mathrm{Sq}(g) = M \circ E(g)$, with
$$E : G \longrightarrow G \times G, \quad E(g) = (g,g),$$
so $DE(e) : \mathfrak{g} \to \mathfrak{g} \oplus \mathfrak{g}$. Show that
$$DE(e)X = \begin{pmatrix} X \\ X \end{pmatrix},$$
and deduce that
$$D\,\mathrm{Sq}(e)X = 2X,$$
thus obtaining (3.1.20) without using the exponential map.

3. Assume $\gamma_0, \gamma_1 : (a,b) \to G$ are smooth curves, $a < 0 < b$, and $\gamma_j(0) = e$. Show that
$$\gamma_j'(0) = X_j \in T_eG \Longrightarrow \frac{d}{ds}\gamma_0(s)\gamma_1(s)\big|_{s=0} = X_0 + X_1.$$
Hint. $\gamma_0(s)\gamma_1(s) = M(\gamma_0(s), \gamma_1(s))$. Use the chain rule.
Remark. If $G \subset Gl(n,\mathbb{R})$ is a matrix group, we have
$$\frac{d}{ds}\gamma_0(s)\gamma_1(s) = \gamma_0'(s)\gamma_1(s) + \gamma_0(s)\gamma_1'(s), \quad \text{in } M(n,\mathbb{R}),$$
but we cannot use such a formula if G is not known to be a matrix group.

4. Recall that the convolution is given by
$$u * v(x) = \int_G u(g)v(g^{-1}x)\,dg, \quad g,x \in G.$$
Show that if $v \in C^1(G)$,
$$X \in \mathfrak{g} \Longrightarrow X(u * v) = u * Xv.$$

3.2. Lie algebras of matrix groups

Here we present another approach to the Lie algebra of a Lie subgroup

$$G \subset \mathrm{Gl}(n, \mathbb{R}). \tag{3.2.1}$$

As a linear space, the Lie algebra of G can be identified as

$$\mathfrak{g} = T_I G. \tag{3.2.2}$$

We can apply to an element $A \in \mathfrak{g}$ the matrix exponential, e^{tA}, introduced in §1.3,

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k.$$

We need to connect this with the notion of the exponential map developed in §3.1. This connection begins with the following fundamental property of the matrix exponential.

Proposition 3.2.1. *For $A \in \mathrm{M}(n, \mathbb{R})$, we have*

$$A \in \mathfrak{g} \Longleftrightarrow e^{tA} \in G, \quad \forall\, t \in \mathbb{R}. \tag{3.2.3}$$

Proof. The "$\Leftarrow$" part is clear from the identity $(d/dt)e^{tA}|_{t=0} = A$. As for the "$\Rightarrow$" part, it has been noted in Proposition 1.3.1 that this follows by inspection for G of the form (1.1.7)–(1.1.10), in which case $\mathfrak{g}$ is given by (1.3.10).

For a general matrix Lie group $G \subset Gl(n, \mathbb{R})$, we bring in results of §3.1. Given $A \in T_I G$, let X denote the unique left-invariant vector field on G such that $X(I) = A$. Then $\gamma_X(t) = \mathrm{Exp}(tX)$ is a one-parameter subgroup of G. Hence $\gamma_X(t)$ and $\gamma_A(t) = e^{tA}$ are both one-parameter subgroups of $Gl(n, \mathbb{R})$ satisfying $\gamma_X'(0) = \gamma_A'(0) = A$. The uniqueness result, Proposition 3.1.3 implies $\gamma_X(t) \equiv \gamma_A(t)$, so $\gamma_A(t) \in G$ for all t. $\square$

Using (3.2.3), we establish the following:

Proposition 3.2.2. *With $[A, B] = AB - BA$, we have*

$$A, B \in \mathfrak{g} \Longrightarrow [A, B] \in \mathfrak{g}. \tag{3.2.4}$$

Proof. Given $g \in G,\ A \in \mathfrak{g}$,

$$g e^{tA} g^{-1} = e^{tgAg^{-1}}, \quad \forall\, t, \tag{3.2.5}$$

and the left side of (3.2.5) belongs to G, so by (3.2.3) we have

$$gAg^{-1} \in \mathfrak{g}, \quad \forall\, g \in G,\ A \in \mathfrak{g}. \tag{3.2.6}$$

Now, for general $B \in \mathfrak{g}$, (3.2.3) yields $g = e^{tB} \in G$, so we have

$$e^{tB} A e^{-tB} \in \mathfrak{g}, \quad \forall\, A, B \in \mathfrak{g}. \tag{3.2.7}$$

Applying d/dt at $t = 0$ gives (3.2.4). $\square$

Remark. Again, for G of the form (1.1.7)–(1.1.10), the result (3.2.4) follows readily from the explicit description of $T_I G$ in (1.3.10).

The commutator $[A,B] = AB - BA$ gives $\mathfrak{g}$ the structure of a *Lie algebra*. We aim to establish further relations between the Lie algebra structure of $\mathfrak{g}$ and the group structure of G.

To begin, let us take $A, B \in \mathfrak{g}$ and record the calculation

$$\begin{aligned} e^{tA}e^{sB} &= \Bigl(I + tA + \frac{t^2}{2}A^2 + O(t^3)\Bigr)\Bigl(I + sB + \frac{s^2}{2}B^2 + O(s^3)\Bigr) \\ &= I + tA + sB + stAB + \frac{t^2}{2}A^2 + \frac{s^2}{2}B^2 + O(|(s,t)|^3), \end{aligned} \tag{3.2.8}$$

and similarly

$$e^{sB}e^{tA} = I + tA + sB + stBA + \frac{t^2}{2}A^2 + \frac{s^2}{2}B^2 + O(|(s,t)|^3), \tag{3.2.9}$$

Hence

$$e^{tA}e^{sB} = e^{sB}e^{tA} + st[A,B] + O(|(s,t)|^3). \tag{3.2.10}$$

We apply these calculations to show how the Lie algebra structure is preserved under representations of G. Thus, assume we have a smooth homomorphism

$$\pi : G \longrightarrow \mathrm{Gl}(m,\mathbb{R}). \tag{3.2.11}$$

(It is shown in Proposition 2.8.10 that every continuous homomorphism of G into $\mathrm{Gl}(m,\mathbb{R})$ is actually smooth.) Let us set

$$\sigma = D\pi(I) : \mathfrak{g} \longrightarrow \mathrm{M}(m,\mathbb{R}), \quad \text{so} \quad \sigma(A) = \frac{d}{ds}\pi(e^{sA})\big|_{s=0}, \tag{3.2.12}$$

for $A \in \mathfrak{g}$. Note that for such A,

$$\begin{aligned} \frac{d}{dt}\pi(e^{tA}) &= \frac{d}{ds}\pi(e^{(s+t)A})\big|_{s=0} \\ &= \frac{d}{ds}\pi(e^{sA})\pi(e^{tA})\big|_{s=0} \\ &= \sigma(A)\pi(e^{tA}), \end{aligned} \tag{3.2.13}$$

and since $\gamma(t) = \pi(e^{tA})$ satisfies $\gamma(0) = I$, this gives

$$\pi(e^{tA}) = e^{t\sigma(A)}. \tag{3.2.14}$$

We are ready to prove:

Proposition 3.2.3. *For π, σ as in (3.2.11)–(3.2.12), $A, B \in \mathfrak{g}$, we have*

$$\sigma([A,B]) = [\sigma(A), \sigma(B)] = \sigma(A)\sigma(B) - \sigma(B)\sigma(A). \tag{3.2.15}$$

Proof. Setting $s = t$ in (3.2.10), we have

$$\begin{aligned} e^{tA}e^{tB}e^{-tA}e^{-tB} &= I + t^2[A,B] + O(t^3) \\ &= e^{t^2[A,B]} + O(t^3). \end{aligned} \tag{3.2.16}$$

Applying π, we have

$$\pi(e^{tA}e^{tB}e^{-tA}e^{-tB}) = \pi(e^{t^2[A,B]}) + O(t^3), \tag{3.2.17}$$

which, by (3.2.14), is equal to

$$\begin{aligned} e^{t\sigma(A)}e^{t\sigma(B)}e^{-t\sigma(A)}e^{-t\sigma(B)} &= I + t^2[\sigma(A),\sigma(B)] + O(t^3) \\ &= e^{t^2[\sigma(A),\sigma(B)]} + O(t^3), \end{aligned} \tag{3.2.18}$$

the last two identities holding by (3.2.16), with A, B replaced by $\sigma(A), \sigma(B)$. From (3.2.17)–(3.2.18), we have

$$\pi(e^{s[A,B]}) = e^{s[\sigma(A),\sigma(B)]} + O(s^{3/2}), \tag{3.2.19}$$

while (3.2.14) yields

$$\pi(e^{s[A,B]}) = e^{s\sigma([A,B])}. \tag{3.2.20}$$

Applying d/ds at $s = 0$ yields (3.2.15). ☐

We next associate to each $A \in T_I G = \mathfrak{g}$ a certain vector field on G. To start, take $A \in \mathrm{M}(n,\mathbb{R})$, the Lie algebra of $\mathrm{Gl}(n,\mathbb{R})$. We define a vector field X_A on $\mathrm{Gl}(n,\mathbb{R})$ by

$$X_A(g) = gA, \tag{3.2.21}$$

for $g \in \mathrm{Gl}(n,\mathbb{R})$. This vector field is left-invariant. That is to say, if for each $h \in \mathrm{Gl}(n,\mathbb{R})$, we define $L_h : \mathrm{Gl}(n,\mathbb{R}) \to \mathrm{Gl}(n,\mathbb{R})$ by

$$L_h g = hg, \tag{3.2.22}$$

then we have

$$X_A(hg) = DL_h(g)X_A(g). \tag{3.2.23}$$

We now have the following simple result:

Proposition 3.2.4. *If $A \in \mathfrak{g} = T_I G$, then X_A is tangent to G.*

Proof. Given $g \in G$, we have $L_g : G \to G$, and hence

$$DL_g(I) : T_I G \longrightarrow T_g G, \tag{3.2.24}$$

hence $A \in \mathfrak{g} \Rightarrow X_A(g) \in T_g G$. ☐

Given $A \in \mathrm{M}(n,\mathbb{R})$, the flow $\mathcal{F}_A^t$ on $\mathrm{Gl}(n,\mathbb{R})$ generated by X_A is given by

$$\mathcal{F}_A^t g = g e^{tA}, \tag{3.2.25}$$

as is readily checked:

$$\begin{aligned} \frac{d}{dt}\mathcal{F}_A^t g\big|_{t=0} &= X_A(g) \quad \text{(by definition)} \\ &= gA, \end{aligned} \tag{3.2.26}$$

which coincides with $(d/dt)ge^{tA}|_{t=0}$. Note that $X = X_A$ is the vector field arising in the last part of the proof of Proposition 3.2.1.

Generally, a smooth vector field X defines a differential operator (also denoted X) on smooth functions by $Xu(x) = (d/dt)u(\mathcal{F}^t x)|_{t=0}$, where $\mathcal{F}^t$ is the flow generated by X. In particular, for $A \in \mathrm{M}(n,\mathbb{R})$,

$$\begin{aligned} X_A u(g) &= \frac{d}{dt} u(ge^{tA})\big|_{t=0} \\ &= Du(g)\cdot gA, \end{aligned} \tag{3.2.27}$$

where the "dot product" gives the action of $Du(g) \in \mathcal{L}(\mathrm{M}(n,\mathbb{R}),\mathbb{R})$ on $gA \in \mathrm{M}(n,\mathbb{R})$. Recall the Lie bracket of vector fields is given by

$$[X_A, X_B] = X_A X_B - X_B X_A. \tag{3.2.28}$$

The following result provides an equivalence between the Lie algebra structure on $\mathfrak{g}$ as we have defined it here and the Lie algebra structure as it is defined in §3.1.

Proposition 3.2.5. *Given $A, B \in \mathrm{M}(n,\mathbb{R})$, we have*

$$[X_A, X_B] = X_{[A,B]}. \tag{3.2.29}$$

Proof. To begin, we have, for u smooth on $Gl(n,\mathbb{R})$,

$$X_A X_B u(g) = \frac{\partial^2}{\partial s \partial t} u(ge^{tA}e^{sB})\big|_{s,t=0}, \tag{3.2.30}$$

and hence

$$(X_A X_B - X_B X_A)u(g) = \frac{\partial^2}{\partial s\partial t}\Big[u(ge^{tA}e^{sB}) - u(ge^{sB}e^{tA})\Big]\Big|_{s,t=0}. \tag{3.2.31}$$

Recalling (3.2.10), we see that

$$\begin{aligned} u(ge^{tA}e^{sB}) &= u(ge^{sB}e^{tA} + stg[A,B] + O(|(s,t)|^3)) \\ &= u(ge^{sB}e^{tA}) + stDu(ge^{sB}e^{tA})\cdot g[A,B] + O(|(s,t)|^3). \end{aligned} \tag{3.2.32}$$

Applying $(\partial^2/\partial s\partial t)|_{s,t=0}$, we obtain

$$\begin{aligned} [X_A, X_B]u(g) &= Du(g)\cdot g[A,B] \\ &= X_{[A,B]}u(g), \end{aligned} \tag{3.2.33}$$

the last identity holding by (3.2.27). This proves (3.2.29). □

This leads to an analogue of Proposition 3.1.5.

Proposition 3.2.6. *Let* $\mathfrak{h}$ *be a linear subspace of* $M(n,\mathbb{R})$*, and assume*

$$A, B \in \mathfrak{h} \Longrightarrow [A, B] \in \mathfrak{h}. \tag{3.2.34}$$

Then the set of left invariant vector fields on $Gl(n,\mathbb{R})$*,*

$$\{X_A : A \in \mathfrak{h}\} = \mathcal{H} \tag{3.2.35}$$

is closed under the Lie bracket of vector fields:

$$X_A, X_B \in \mathcal{H} \Longrightarrow [X_A, X_B] = X_{[A,B]} \in \mathcal{H}. \tag{3.2.36}$$

Hence the integral manifold H *of* $\mathcal{H}$ *through* I *is a subgroup of* $Gl(n,\mathbb{R})$*. We have*

$$T_I H = \mathfrak{h}. \tag{3.2.37}$$

Exercises

1. Define the conjugation map

$C_g : M(n,\mathbb{R}) \longrightarrow M(n,\mathbb{R}), \quad C_g(X) = gXg^{-1},\ g \in Gl(n,\mathbb{R}),\ X \in M(n,\mathbb{R}).$

Show that, for each $X \in M(n,\mathbb{R})$,

$$DC_g(X)A = gAg^{-1}, \quad \forall\, A \in M(n,\mathbb{R}).$$

Hint. $C_g(X)$ is linear in X.

2. If $G \subset Gl(n,\mathbb{R})$ is a smooth matrix group, define

$$C_g : G \longrightarrow G, \quad C_g(x) = gxg^{-1}.$$

We have $C_g(I) = I$, hence

$$DC_g(I) : T_I G \longrightarrow T_I G, \quad \forall\, g \in G.$$

Show that this leads to another proof of (3.2.6), one that does not rely on Proposition 3.2.1. (We still need Proposition 3.2.1 for (3.2.7).)

3. For $A \in \mathfrak{g} = T_I G$, set

$$\mathrm{Ad}(g)A = DC_g(I)A, \quad \mathrm{Ad}(g) : \mathfrak{g} \to \mathfrak{g}.$$

Hence

$$\mathrm{Ad}(g)A = gAg^{-1}, \quad A \in \mathfrak{g},\ g \in G.$$

Show that Ad is a representation of G on $\mathfrak{g}$.

Recall the correspondence between a representation π of G and the associated representation σ of $\mathfrak{g}$, defined by (3.2.12). Show that

$$\pi(g) = \mathrm{Ad}(g) \Longrightarrow \sigma(A)B = [A, B],$$

i.e., $\sigma(A)B = \mathrm{ad}(A)B$, where

$$\mathrm{ad}(A)B = [A, B].$$

Deduce from (3.2.14) that

$$\mathrm{Ad}(e^{tA}) = e^{t\,\mathrm{ad}\,A}, \quad \text{for} \ \ A \in \mathfrak{g}.$$

4. Complement the $s = t$ case of (3.2.8)–(3.2.10) with

$$e^{t(A+B)} = I + t(A+B) + \frac{t^2}{2}(A^2 + AB + BA + B^2) + O(t^3),$$

hence

$$e^{tA}e^{tB} = e^{t(A+B)} + \frac{t^2}{2}(AB - BA) + O(t^3).$$

5. Show that, for $X, Y \in M(n, \mathbb{R})$,
$$e^{tX+t^2Y} = e^{tX} + t^2Y + O(t^3).$$

6. Use Exercises 4 and 5 to show that
$$e^{tA}e^{tB} = e^{t(A+B)+(t^2/2)[A,B]} + O(t^3).$$
See §3.5 for a more precise result.

7. Verify using Proposition 3.2.1 and the inverse function theorem that if G is a matrix Lie group, $\mathfrak{g} = T_I G$,
$$\operatorname{Exp} : \mathfrak{g} \longrightarrow G, \quad \operatorname{Exp}(X) = e^X$$
yields a diffeomorphism of a neighborhood of 0 in $\mathfrak{g}$ onto a neighborhood of I in G.

8. Show that $\operatorname{Exp} : \mathfrak{g} \to G$ is onto for the following cases:
$$G = SU(n),\ U(n),\ SO(n).$$
Hint. Diagonalize elements of $SU(n)$ and $U(n)$. If $A \in SO(n)$, show that A is similar to a block diagonal matrix containing 2×2 blocks
$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = e^{\theta J},$$
if n is even, and similar to such a form plus one diagonal entry of 1 if n is odd.

9. In fact, Exp is onto whenever G is a compact, connected Lie group. See if you can prove this. Maybe peek at Appendix E.3.

10. Show that
$$A \in \mathfrak{sl}(2, \mathbb{R}) \Longrightarrow e^A \notin \left\{ -\begin{pmatrix} \lambda & \\ & 1/\lambda \end{pmatrix} : \lambda > 1 \right\}.$$
Hint. Say $\operatorname{Spec} A = \{\alpha, -\alpha\}$. $\alpha \in \mathbb{R} \Rightarrow \operatorname{Spec} e^A \subset \mathbb{R}^+$. $\alpha \notin \mathbb{R} \Rightarrow -\alpha = \overline{\alpha} \Rightarrow \alpha = ia,\ a \in \mathbb{R} \Rightarrow \operatorname{Spec} e^A \in \{\zeta \in \mathbb{C} : |\zeta| = 1\}$.

11. Show that
$$A \in \mathfrak{sl}(2, \mathbb{R}) \Longrightarrow e^A \neq -\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$
Hint. Say $\operatorname{Spec} A = \{-\alpha, \alpha\}$. $\alpha \neq 0 \Rightarrow A$ diagonalizable $\Rightarrow e^A$ diagonalizable (in $Gl(2, \mathbb{C})$). But $\operatorname{Spec} A = \{0\}$ implies A is nilpotent, so $A^2 = 0$, hence
$$e^A = I + A, \quad \text{but} \quad I + A \neq -I + N.$$

Note. $-I = e^{\pi J}$.

3.3. Lie algebra representations

Let G be a Lie group and π a (strongly continuous) representation of G on a finite-dimensional vector space V. As we have seen in Proposition 2.8.10, when $\dim V < \infty$ all vectors $v \in V$ are smooth. We define the map $d\pi : \mathfrak{g} \to \mathrm{End}(V)$ as follows:

$$d\pi(X)v = \frac{d}{dt}\,\pi(\mathrm{Exp}\,tX)v\Big|_{t=0}, \quad X \in \mathfrak{g},\ v \in V. \tag{3.3.1}$$

Recall that X is a left-invariant vector field on G. The following lemma will be helpful to understand $d\pi$. Given $f \in C_0^\infty(G)$, we set

$$\pi(f) = \int_G f(g)\pi(g)\,dg, \tag{3.3.2}$$

as before, except now we use *right-invariant* Haar measure on G.

Lemma 3.3.1. *Given $f \in C_0^\infty(G)$, $X \in \mathfrak{g}$, we have*

$$\pi(f)d\pi(X) = -\pi(Xf). \tag{3.3.3}$$

Proof. Plugging in the definitions yields

$$\begin{aligned}
\pi(f)d\pi(X)v &= \frac{d}{dt}\int f(g)\pi(g)\pi(\mathrm{Exp}\,tX)v\,dg\Big|_{t=0} \\
&= \frac{d}{dt}\int f(g)\pi(g\ \mathrm{Exp}\,tX)v\,dg\Big|_{t=0} \\
&= \frac{d}{dt}\int f\big(g\ \mathrm{Exp}(-tX)\big)\pi(g)v\,dg\Big|_{t=0} \\
&= -\int (Xf)(g)\pi(g)v\,dg \\
&= -\pi(Xf)v.
\end{aligned} \tag{3.3.4}$$

Here the third identity uses the right invariance of Haar measure and the fourth identity uses (3.1.17). ☐

We can deduce the following important consequence. Compare the result for matrix groups in Proposition 3.2.3, though the proof is perhaps more parallel to that of Proposition 3.2.5.

Proposition 3.3.2. *Given $X, Y \in \mathfrak{g}$,*

$$d\pi([X,Y]) = [d\pi(X), d\pi(Y)]. \tag{3.3.5}$$

Proof. For any $f \in C_0^\infty(G),\ v \in V$, we have

$$\begin{aligned}\pi(f)\big(d\pi(X)d\pi(Y) - d\pi(Y)d\pi(X)\big)v \\ &= \pi(YXf)v - \pi(XYf)v \\ &= -\pi([X,Y]f)v \\ &= \pi(f)d\pi([X,Y])v.\end{aligned} \tag{3.3.6}$$

Letting $f = f_\nu$ be an approximate identity gives the result. $\square$

Due to (3.3.5), we say $d\pi$ is a Lie algebra representation. We call it the derived representation associated to the representation π. The following is an important connection between Lie algebra and Lie group representations. Compare (3.2.14).

Proposition 3.3.3. *For all* $X \in \mathfrak{g}$,

$$\pi(\operatorname{Exp} tX) = e^{t\, d\pi(X)}. \tag{3.3.7}$$

Proof. Let $A = d\pi(X) \in \operatorname{End}(V)$ and let $\gamma(t)$ denote the left side of (3.3.7). We want to show that $\gamma(t) \equiv e^{tA}$. It is clear that $\gamma : \mathbb{R} \to \mathrm{Gl}(V)$ is a smooth one-parameter group, and (3.3.1) gives $\gamma'(0) = d\pi(X) = A$. The group property gives

$$\gamma'(t) = \frac{d}{ds}\gamma(s+t)\big|_{s=0} = A\gamma(t) = \gamma(t)A, \tag{3.3.8}$$

and hence

$$\frac{d}{dt}\,\gamma(t)e^{-tA} = \gamma(t)Ae^{-tA} - \gamma(t)Ae^{-tA} = 0, \tag{3.3.9}$$

so $\gamma(t)e^{-tA} \equiv I$.

Alternatively, the uniqueness of the one-parameter subgroup γ of $\mathrm{Gl}(V)$ satisfying $\gamma'(0) = A$ (cf. (3.1.9)) gives $\gamma(t) = e^{tA}$. $\square$

We next relate irreducibility of π and of $d\pi$.

Proposition 3.3.4. *Assume G is connected. Then π is an irreducible representation of G if and only if $d\pi$ is an irreducible representation of $\mathfrak{g}$.*

Proof. Let $V_0 \subset V$ be a linear subspace of V. First suppose V_0 is invariant under $\pi(g)$ for all $g \in G$. Then, for any $X \in \mathfrak{g}$,

$$v \in V_0 \Rightarrow d\pi(X)v = \frac{d}{dt}\pi(\operatorname{Exp} tX)v\big|_{t=0} \in V_0, \tag{3.3.10}$$

so V_0 is invariant under $d\pi(X)$ for all $X \in \mathfrak{g}$.

Next suppose V_0 is invariant under $d\pi(X)$ for all $X \in \mathfrak{g}$. Then, for any $X \in \mathfrak{g}$,

$$v \in V_0 \Rightarrow \pi(\operatorname{Exp} tX)v = e^{t\,d\pi(X)}v = \sum_{k\geq 0} \frac{t^k}{k!}\, d\pi(X)^k\, v \in V_0. \tag{3.3.11}$$

Now if G is connected, any $g \in G$ can be written in the form (3.1.27), so

$$v \in V_0 \Rightarrow \pi(g)v = \pi(\operatorname{Exp} X_1)\cdots\pi(\operatorname{Exp} X_K)v \in V_0. \tag{3.3.12}$$

$\square$

Suppose V has a Hermitian inner product and the representation π of G on V is unitary. Then, for $X \in \mathfrak{g}$,

$$e^{-t\,d\pi(X)} = \pi(\gamma_X(t))^{-1} = \pi(\gamma_X(t))^* = (e^{t\,d\pi(X)})^*, \tag{3.3.13}$$

and hence

$$d\pi(X)^* = -d\pi(X). \tag{3.3.14}$$

In other words, $\mathfrak{g}$ is represented by skew-Hermitian operators on V. The following is a Lie algebra variant of Schur's lemma.

Proposition 3.3.5. *Let $\mathfrak{g}$ be a Lie algebra, V a complex inner product space, and $\alpha : \mathfrak{g} \to \operatorname{End}(V)$ a Lie algebra representation of $\mathfrak{g}$ by skew-Hermitian operators on V. Then α is irreducible if and only if the following holds:*

$$\begin{gathered} A \in \operatorname{End}(V), \quad \alpha(X)A = A\alpha(X) \ \ \textit{for all} \ \ X \in \mathfrak{g} \\ \Longrightarrow A \ \ \textit{is a scalar multiple of the identity.} \end{gathered} \tag{3.3.15}$$

The proof is as for Lemma 2.1.4. One sees that if A commutes with $\alpha(X)$, so do $A_1 = A + A^*$ and $A_2 = (A - A^*)/i$, and the eigenspaces of A_j are invariant.

This gives one implication. For the converse, observe that if $V_0 \subset V$ is invariant under α, so is $V_0^\perp$, so the orthogonal projection of V nto V_0 commutes with $\alpha(X)$, for all X.

Exercises

1. Recall that a representation π of G on V is a smooth map $\pi : G \to \mathcal{L}(V)$. Show that, for $X \in \mathfrak{g}$,

$$d\pi(X) = D\pi(e)X(e),$$

with $D\pi(e) : T_eG \to \mathcal{L}(V)$.

2. Suppose the representation $\pi : G \to \mathcal{L}(V)$ is a one-to-one map. Show that $d\pi : \mathfrak{g} \to \mathcal{L}(V)$ is injective.
Hint. Use Proposition 3.3.3.

3. Show that, for $g \in G$, $\pi : G \to \mathcal{L}(V)$ a representation, $X \in \mathfrak{g}$, so

$$X(g) = DL_g(e)X(e) \in T_gG,$$

we have

$$D\pi(g)X(g) = \pi(g)d\pi(X).$$

4. Deduce from Exercises 2–3 that if the representation $\pi : G \to \mathcal{L}(V)$ is one-to-one, then

$$D\pi(g) : T_gG \longrightarrow \mathcal{L}(V) \;\text{ is injective, } \forall\, g \in G.$$

Hence π is a sooth embedding of G into $\mathcal{L}(V)$.

3.4. The adjoint representation

Here we consider a particularly important representation of a Lie group G on its Lie algebra $\mathfrak{g}$, the *adjoint representation*, defined as follows. Take

$$K_g : G \to G, \quad K_g(x) = gxg^{-1}, \tag{3.4.1}$$

and set

$$\mathrm{Ad}(g) = DK_g(e) : T_eG \to T_eG \approx \mathfrak{g}. \tag{3.4.2}$$

Note that

$$K_{gh} = K_g \circ K_h \Longrightarrow \mathrm{Ad}(gh) = \mathrm{Ad}(g)\,\mathrm{Ad}(h). \tag{3.4.3}$$

Proposition 3.4.1. *For $g \in G,\ X \in \mathfrak{g}$,*

$$\mathrm{Exp}(t\ \mathrm{Ad}(g)X) = g\ \mathrm{Exp}(tX)\, g^{-1}. \tag{3.4.4}$$

Proof. Both sides of (3.4.4) are one-parameter subgroups of G. Call them $\gamma(t)$ and $\sigma(t)$, respectively. It follows from (3.1.13) that $\gamma'(0) = \mathrm{Ad}(g)X$. Meanwhile, since $\sigma(t) = K_g(\mathrm{Exp}(tX))$, the chain rule plus (3.4.2) gives $\sigma'(0) = \mathrm{Ad}(g)X$. The uniqueness result established in (3.1.9) then implies $\gamma(t) \equiv \sigma(t)$. □

Let us take $g = \mathrm{Exp}\, sY$. By (3.1.7) the right side of (3.4.4) is then equal to

$$\begin{aligned} g\gamma_X(t)g^{-1} &= \mathcal{F}_Y^{-s}(g\gamma_X(t)) \\ &= \mathcal{F}_Y^{-s} \circ \mathcal{F}_X^t g \\ &= \mathcal{F}_Y^{-s} \circ \mathcal{F}_X^t \circ \mathcal{F}_Y^s\, e \\ &= \mathcal{F}_{X(s)}^t\, e, \qquad X(s) = \mathcal{F}_{Y\#}^s X, \end{aligned} \tag{3.4.5}$$

the last identity using (A.5.1). Consequently, comparing the left side of (3.4.4), and noting that $\mathcal{F}_{X(s)}^t e = \mathrm{Exp}(tX(s))$, we have

$$\mathrm{Exp}(t\,\mathrm{Ad}(\mathrm{Exp}\, sY)X) = \mathrm{Exp}(t\mathcal{F}_{Y\#}^s X),$$

hence

$$\mathrm{Ad}(\mathrm{Exp}\, sY)X = \mathcal{F}_{Y\#}^s X. \tag{3.4.6}$$

Taking the s-derivative at $s = 0$ and using (A.5.3)–(A.5.5), we have the following important conclusion.

Proposition 3.4.2. *For $X, Y \in \mathfrak{g}$,*

$$\frac{d}{ds}\ \mathrm{Ad}(\mathrm{Exp}\, sY)X\big|_{s=0} = [Y, X]. \tag{3.4.7}$$

According to (3.3.1), the left side of (3.4.7) is the Lie algebra representation derived from Ad, i.e., $d\,\mathrm{Ad}(Y)X$. We use the notation $\mathrm{ad}(Y)$ instead of $d\,\mathrm{Ad}(Y)$, a notation already brought forward in (3.1.23), and express the conclusion of Proposition 3.4.2 as follows:

$$\mathrm{ad}(Y)X = [Y, X]. \tag{3.4.8}$$

Having examined the adjoint representation in the setting of abstract Lie groups, let us take a second look in the concrete setting where G is a matrix group, e.g., $G = \mathrm{Gl}(n, \mathbb{F})$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Then K_g in (3.4.1) extends to a map linear in $x \in \mathrm{M}(n, \mathbb{F})$, and we simply have

$$\mathrm{Ad}(g)X = gXg^{-1}, \quad X \in \mathfrak{g} \subset \mathrm{M}(n, \mathbb{F}). \tag{3.4.9}$$

See Exercise 3 in §3.2. If $g = \operatorname{Exp} sY = \gamma_Y(s)$, we have

$$\mathrm{Ad}(\operatorname{Exp} sY)X = \gamma_Y(s)\, X\, \gamma_Y(-s), \quad X, Y \in \mathfrak{g} \subset \mathrm{M}(n, \mathbb{F}). \tag{3.4.10}$$

Since the matrix product $(A, B) \mapsto AB$ is bilinear on $\mathrm{M}(n, \mathbb{F})$, we can apply the Leibniz rule to differentiate such a product, and obtain

$$\frac{d}{ds}\, \mathrm{Ad}(\operatorname{Exp} sY)X\big|_{s=0} = YX - XY, \quad X, Y \in \mathfrak{g} \subset \mathrm{M}(n, \mathbb{F}). \tag{3.4.11}$$

Thus the Lie bracket on $\mathfrak{g} \subset \mathrm{M}(n, \mathbb{F})$ is seen to be the matrix commutator. This result is consistent with (3.3.5), applied to the "identity" representation $G \hookrightarrow \mathrm{Gl}(n, \mathbb{F})$ of G on $\mathbb{F}^n$. Compare also the presentation in §3.2.

The adjoint representation can be used to tell whether G is unimodular. In fact, take a nonzero $\omega_0 \in \Lambda^N T_e^* G$ ($N = \dim G$). and define a left Haar measure via $\omega_0 = L_g^* \omega(g)$. We have

$$R_{g^{-1}}^* \omega = R_{g^{-1}}^* L_g^* \omega = K_g^* \omega = \det \mathrm{Ad}(g)\, \omega. \tag{3.4.12}$$

(Otherwise said, $K_g^* \omega(e) = \Lambda^N DK_g(e)^t \omega_0 = \det \mathrm{Ad}(g) \omega_0$.) Hence G is unimodular if and only if $\det \mathrm{Ad}(g) = 1$ for all $g \in G$.

We can make use of the identity

$$\mathrm{Ad}(\operatorname{Exp} X) = e^{\mathrm{ad}\, X}, \tag{3.4.13}$$

which is a special case of (3.3.7), to formulate the unimodularity condition in purely Lie algebra terms. In view of Proposition 3.1.6, when G is connected, $\det \mathrm{Ad}(g) = 1$ for all $g \in G$ if and only if $\det \mathrm{Ad}(\operatorname{Exp} X) = 1$ for all $X \in \mathfrak{g}$. Now, for a general linear map A on a finite-dimensional vector space,

$$\det e^A = e^{\operatorname{Tr} A}, \tag{3.4.14}$$

so we have:

Proposition 3.4.3. *A connected Lie group G is unimodular if and only if* $\operatorname{Tr} \mathrm{ad}(X) = 0$ *for all* $X \in \mathfrak{g}$.

We give an example of a Lie group that is not unimodular, namely the 2-dimensional group Aff(1), known as the "$ax+b$-group." As a set, $\mathrm{Aff}(1) = \mathbb{R}^+ \times \mathbb{R}$; it acts on $\mathbb{R}$ by $(a,b)\cdot x = ax+b$, so the group law is

$$(a,b)(a',b') = (aa', b+ab'). \tag{3.4.15}$$

This group is isomorphic to the group of matrices

$$\left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a > 0,\, b \in \mathbb{R} \right\}. \tag{3.4.16}$$

Compare (1.1.19)–(1.1.22). The Lie algebra of Aff(1) is isomorphic to the matrix Lie subalgebra of $\mathrm{M}(2,\mathbb{R})$ spanned by

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{3.4.17}$$

We have

$$[X,Y] = Y, \quad \mathrm{Tr}\,\mathrm{ad}\, X = 1. \tag{3.4.18}$$

It is instructive to compute left and right invariant measures on this group, or equivalently left and right invariant 2-forms:

$$\omega_L = \varphi(x,y)\, dx\wedge dy, \quad \omega_R = \psi(x,y)\, dx \wedge dy,$$

on $\{(x,y) : x > 0, y \in \mathbb{R}\}$. Using

$$L_{a,b}(x,y) = (ax, ay+b), \quad R_{a,b}(x,y) = (ax, bx+y),$$

we have

$$\begin{aligned} L_{a,b}^*\omega_L &= a^2\varphi(ax, ay+b)\, dx\wedge dy, \\ R_{a,b}^*\omega_R &= a\psi(ax, bx+y)\, dx\wedge dy. \end{aligned}$$

Invariance is achieved by setting $\varphi(x,y) = 1/x^2$ and $\psi(x,y) = 1/x$, so we have

$$\omega_L = x^{-2}\, dx\wedge dy, \quad \omega_R = x^{-1}\, dx\wedge dy.$$

We next express the formula (1.3.19) for the derivative of the matrix exponential in terms of Ad. As shown in (1.3.19), if we consider

$$\mathrm{Exp} : \mathrm{M}(n,\mathbb{R}) \longrightarrow \mathrm{Gl}(n,\mathbb{R}), \quad \mathrm{Exp}\, X = e^X, \tag{3.4.19}$$

then

$$D\,\mathrm{Exp}(X)Y = e^X \int_0^1 e^{-\sigma X} Y e^{\sigma X}\, d\sigma. \tag{3.4.20}$$

Now, by (3.4.9) and (3.4.13),

$$e^{-\sigma X} Y e^{\sigma X} = \mathrm{Ad}(e^{-\sigma X})Y = e^{-\sigma\, \mathrm{ad}\, X} Y, \tag{3.4.21}$$

so we can rewrite (3.4.20) as

$$D\,\mathrm{Exp}(X)Y = e^X\, \Xi(\mathrm{ad}\, X)Y, \tag{3.4.22}$$

where

$$\Xi(z) = \int_0^1 e^{-\sigma z}\, d\sigma = \frac{1-e^{-z}}{z} \tag{3.4.23}$$

is an entire holomorphic function and $\Xi(A)$ is defined as in (1.3.20)–(1.3.21) for a linear transformation A on a finite-dimensional vector space; in this case $V = \mathfrak{g}$ and $A = \operatorname{ad} X$.

We now point out analogues of (3.4.19)–(3.4.23) valid for a general Lie group G, arising by defining e^{tX} for $x \in \mathfrak{g}$ as an operator on functions:

$$e^{tX}u(x) = u(\mathcal{F}_X^t x) = u(x \cdot \operatorname{Exp} tX). \tag{3.4.24}$$

Note that, for $u \in C^\infty(G)$,

$$\begin{aligned} \frac{d}{dt} e^{tX}u(x) &= Du(\mathcal{F}_X^t x)\, X(\mathcal{F}_X^t x) \\ &= Xu(\mathcal{F}_X^t x) \\ &= e^{tX}Xu(x). \end{aligned} \tag{3.4.25}$$

We claim that

$$e^{tX}Xu(x) = Xe^{tX}u(x). \tag{3.4.26}$$

Indeed, since $Xv(y) = (d/ds)v(\mathcal{F}_X^s y)|_{s=0}$,

$$\begin{aligned} X(e^{tX}u)(x) &= \frac{d}{ds}(e^{tX}u)(\mathcal{F}_X^s x)|_{s=0} \\ &= \frac{d}{ds}u(\mathcal{F}_X^t \mathcal{F}_X^s x)|_{s=0} \\ &= \frac{d}{ds}u(\mathcal{F}_X^s \mathcal{F}_X^t x)|_{s=0} \\ &= Xu(\mathcal{F}_X^t x), \end{aligned} \tag{3.4.27}$$

yielding (3.4.26). Thus

$$\frac{d}{dt}e^{tX}u(x) = e^{tX}Xu(x) = Xe^{tX}u(x). \tag{3.4.28}$$

From here, the derivation of (1.3.18)–(1.3.19) readily extends to yield

$$\begin{aligned} D\operatorname{Exp}(X)Y &= \frac{d}{dt}e^{X+tY}\big|_{t=0} \\ &= e^X \int_0^1 e^{-\sigma X} Y e^{\sigma X}\, d\sigma. \end{aligned} \tag{3.4.29}$$

Actually, (3.4.29) works for general smooth vector fields X and Y on a smooth manifold, assuming their flows are everywhere defined. However, the next step,

$$e^{-\sigma X} Y e^{\sigma X} = e^{-\sigma \operatorname{ad} X} Y, \tag{3.4.30}$$

extending (3.4.21), is problematic for such general vector fields, particularly if one wants to treat the right hand side as a convergent power series. We establish that (3.4.30) holds if G is a Lie group, with Lie algebra $\mathfrak{g}$, and $X, Y \in \mathfrak{g}$.

To get this, note that

$$Y(\sigma) = e^{-\sigma X} Y e^{\sigma X} \tag{3.4.31}$$

commutes with $L(g)$ for all $g \in G$, hence is a smooth curve in $\mathfrak{g}$. Also,

$$\begin{aligned} Y'(\sigma) &= -X e^{-\sigma X} Y e^{\sigma X} + e^{-\sigma X} Y e^{\sigma X} X \\ &= -[X, Y(\sigma)] \\ &= -\operatorname{ad} X(Y(\sigma)), \end{aligned} \tag{3.4.32}$$

the first identity by (3.4.28). Since $\mathfrak{g}$ is finite dimensional, $\operatorname{ad} X$ is a bounded linear operator on $\mathfrak{g}$, so the unique solution to (3.4.32) that is a smooth curve in $\mathfrak{g}$ satisfying $Y(0) = Y$ is

$$Y(\sigma) = e^{-\sigma \operatorname{ad} X} Y. \tag{3.4.33}$$

To proceed, we now have

$$\begin{aligned} D \operatorname{Exp}(X) Y &= e^X \int_0^1 e^{-\sigma \operatorname{ad} X} Y \, d\sigma \\ &= e^X \Xi(\operatorname{ad} X) Y, \end{aligned} \tag{3.4.34}$$

as in (3.4.22). This will be useful in §3.5.

Exercises

1. Let G be a Lie group. Recall $R(g)u(x) = u(xg)$. By (3.4.24), for $Z \in \mathfrak{g}$,

$$R(\operatorname{Exp}(tZ))u = e^{tZ}u,$$

By (3.4.4),

$$\operatorname{Exp}\big(t\ \operatorname{Ad}(\operatorname{Exp} sY)X\big) = \operatorname{Exp}(sY)\ \operatorname{Exp}(tX)\ \operatorname{Exp}(-sY).$$

Deduce that

$$R\big(\operatorname{Exp}\big(t\operatorname{Ad}(\operatorname{Exp} sY)X\big)\big)u = e^{sY}e^{tX}e^{-sY}u,$$

2. Show that for $Z \in \mathfrak{g},\ u \in C_0^\infty(G)$,

$$dR(Z)u = Zu.$$

Deduce from Exercise 1 and (3.4.25) that

$$dR\big(\operatorname{Ad}(\operatorname{Exp}\ sY)X\big)u = e^{sY}Xe^{-sY}u.$$

Again using (3.4.25), show that

$$\frac{d}{ds}\,dR\big(\operatorname{Ad}(\operatorname{Exp}\ sY)X\big)u\Big|_{s=0} = (YX - XY)u,$$

hence re-deriving (3.4.7),

$$\frac{d}{ds}\ \operatorname{Ad}(\operatorname{Exp}\ sY)X\big|_{s=0} = [Y,X].$$

3. Note that $\operatorname{ad} : \mathfrak{g} \to \mathcal{L}(\mathfrak{g})$ satisfies

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

which is equivalent to the Jacobi identity (3.1.26), and is also a special case of Proposition 3.3.2. Deduce that

$$\tau : \mathfrak{g} \longrightarrow \mathbb{R}, \quad \tau(X) = \operatorname{Tr}\operatorname{ad} X$$

satisfies

$$\tau([X,Y]) = 0, \quad \forall\, X, Y \in \mathfrak{g}.$$

4. Rewrite (3.4.34) as

$$D\operatorname{Exp}(X)Y = DL_g(e)\,\Xi(\operatorname{ad} X)Y, \quad g = \operatorname{Exp}(X).$$

5. Show that

$$\Xi(z) = 0 \Longleftrightarrow z = 2\pi i k,\ k \in \mathbb{Z} \setminus 0.$$

Deduce that $\Xi(\operatorname{ad} X) \in \mathcal{L}(\mathfrak{g})$ fails to be invertible if and only if $\operatorname{ad} X$ has an eigenvalue in $\{2\pi i k : k \in \mathbb{Z} \setminus 0\}$.

6. Assume G has a bi-invariant metric tensor. Show that K_g in (3.4.1) is

a group of isometries of G, and that $\mathrm{Ad}(g)$ in (3.4.2) is a group of linear isometries of $\mathfrak{g}$.

3.5. The Campbell-Hausdorff formula

The Campbell-Hausdorff formula has the form

$$\mathrm{Exp}(X)\ \mathrm{Exp}(Y) = \mathrm{Exp}\big(\mathcal{C}(X,Y)\big), \tag{3.5.1}$$

where G is any Lie group, with Lie algebra $\mathfrak{g}$, and $\mathrm{Exp} : \mathfrak{g} \to G$ is the exponential map defined by (3.1.12)); X and Y are elements of $\mathfrak{g}$ in a sufficiently small neighborhood U of zero. The map $\mathcal{C} : U \times U \to \mathfrak{g}$ has a "universal" form, independent of $\mathfrak{g}$. We give a demonstration similar to one in [**22**].

We begin with the case $G = \mathrm{Gl}(n, \mathbb{R})$, and produce an explicit formula for the matrix-valued analytic function $X(s)$ of s in the identity

$$e^{X(s)} = e^X e^{sY}, \tag{3.5.2}$$

near $s = 0$. Note that this function satisfies the ODE

$$\frac{d}{ds} e^{X(s)} = e^{X(s)} Y. \tag{3.5.3}$$

We can produce an ODE for $X(s)$ by using the following formula, derived in (1.3.19):

$$\frac{d}{ds} e^{X(s)} = e^{X(s)} \int_0^1 e^{-\tau X(s)} X'(s) e^{\tau X(s)}\, d\tau. \tag{3.5.4}$$

As shown in (3.4.22), we can rewrite this as

$$\frac{d}{ds} e^{X(s)} = e^{X(s)} \Xi\big(\mathrm{ad}\ X(s)\big) X'(s), \tag{3.5.5}$$

with

$$\Xi(z) = \int_0^1 e^{-\tau z}\, d\tau = \frac{1 - e^{-z}}{z}. \tag{3.5.6}$$

Comparing (3.5.3) and (3.5.5), we obtain

$$\Xi\big(\mathrm{ad}\ X(s)\big) X'(s) = Y, \quad X(0) = X. \tag{3.5.7}$$

We can obtain a more convenient ODE for $X(s)$ as follows. Note that

$$e^{\mathrm{ad}\ X(s)} = \mathrm{Ad}\ e^{X(s)} = \mathrm{Ad}\ e^X \cdot \mathrm{Ad}\ e^{sY} = e^{\mathrm{ad}\ X}\ e^{s\ \mathrm{ad}\ Y}. \tag{3.5.8}$$

Now let $\Psi(\zeta)$ be holomorphic near $\zeta = 1$ and satisfy

$$\Psi(e^a) = \frac{1}{\Xi(a)} = \frac{a}{1 - e^{-a}}, \tag{3.5.9}$$

explicitly,

$$\Psi(\zeta) = \frac{\zeta\ \log \zeta}{\zeta - 1}, \tag{3.5.10}$$

for $|\zeta - 1| < 1$. (Note that the singularity at $\zeta = 1$ is removable.) It follows that

$$\Psi\big(e^{\text{ad }X}e^{s\text{ ad }Y}\big)\Xi\big(\text{ad }X(s)\big) = I, \tag{3.5.11}$$

so we can transform (3.5.7) to

$$X'(s) = \Psi\big(e^{\text{ad }X}e^{s\text{ ad }Y}\big)Y, \quad X(0) = X. \tag{3.5.12}$$

Integrating gives the Campbell-Hausdorff formula for $X(s)$ in (3.5.2):

$$X(s) = X + \int_0^s \Psi\big(e^{\text{ad }X}e^{t\text{ ad }Y}\big)Y\ dt. \tag{3.5.13}$$

This is valid for $\|sY\|$ small enough, if also X is close enough to 0.

Taking the $s = 1$ case, we can rewrite this formula as

$$e^X e^Y = e^{\mathcal{C}(X,Y)}, \quad \mathcal{C}(X,Y) = X + \int_0^1 \Psi\big(e^{\text{ad }X}e^{t\text{ ad }Y}\big)Y\ dt. \tag{3.5.14}$$

The formula (3.5.14) gives a power series in ad X and ad Y which is norm-summable provided

$$\|\text{ad }X\| \le x, \quad \|\text{ad }Y\| \le y, \tag{3.5.15}$$

with $e^{x+y} - 1 < 1$, i.e.,

$$x + y < \log 2. \tag{3.5.16}$$

We can extend the analysis above to the case where X and Y belong to the Lie algebra $\mathfrak{g}$ of a Lie group G. As shown in §3.4, if $X(s)$ is a smooth curve in $\mathfrak{g}$, then (3.5.5) continues to hold. Since $\operatorname{ad} X$ and $\operatorname{ad} Y$ are bounded linear transformations on $\mathfrak{g}$, the argument involving (3.5.7)–(3.5.16) extends. We have

$$\mathcal{F}_X^t \mathcal{F}_Y^t = \mathcal{F}_{\mathcal{C}(t,X,Y)}^t, \tag{3.5.17}$$

with

$$\mathcal{C}(t,X,Y) = X + \int_0^1 \Psi\big(e^{\text{ad }tX}e^{\text{ad }stY}\big)Y\ ds, \tag{3.5.18}$$

provided $\|\text{ad }tX\| + \|\text{ad }tY\| < \log 2$, the operator norm $\|\text{ad }X\|$ being computed using any convenient norm on $\mathfrak{g}$. This yields the Campbell-Hausdorff formula for general Lie groups.

Another way to describe the extension of the Campbell-Hausdorff formula to general Lie groups is given by studying e^{tX} as operators for $X \in \mathfrak{g}$, given by (3.4.24). In this approach, we take $X, Y \in \mathfrak{g}$, near 0, and look for $\mathcal{C}(X,Y) \in \mathfrak{g}$ such that

$$e^X e^Y u(x) = e^{\mathcal{C}(X,Y)}u(x). \tag{3.5.19}$$

The construction of $\mathcal{C}(X,Y)$ uses the same formulas as in (3.5.2)–(3.5.14). Again we have

$$\mathcal{C}(X,Y) = X + \int_0^1 \Psi(e^{\mathrm{ad}\,X} e^{t\,\mathrm{ad}\,Y})Y\,dt. \tag{3.5.20}$$

Note that the left and right sides of (3.5.19) are equal respectively to

$$u\big(x(\operatorname{Exp} X)(\operatorname{Exp} Y)\big) \quad \text{and} \quad u(x \cdot \operatorname{Exp} \mathcal{C}(X,Y)), \tag{3.5.21}$$

so from (3.5.19) we again deduce

$$(\operatorname{Exp} X)(\operatorname{Exp} Y) = \operatorname{Exp} \mathcal{C}(X,Y). \tag{3.5.22}$$

One remarkable property of Lie groups that follows readily from the Campbell-Hausdorff formula is the existence of a natural real analytic structure on any Lie group G. (Recall we originally assumed G has a C^∞ structure.) This comes about as follows. Pick a neighborhood U of the origin 0 in the Lie algebra $\mathfrak{g}$ of G sufficiently small that

$$\operatorname{Exp} : U \longrightarrow G \tag{3.5.23}$$

is a diffeomorphism of U onto a neighborhood $\mathcal{O}$ of $e \in G$. Then, for each $p \in G$, define

$$\psi_p : U \longrightarrow G, \quad \psi_p(X) = p \operatorname{Exp}(X). \tag{3.5.24}$$

Proposition 3.5.1. *The coordinate cover $\{\psi_p : p \in G\}$ gives G the structure of a real-analytic manifold, on which the maps $(g,h) \mapsto gh$ and $g \mapsto g^{-1}$ are real analytic.*

Proof. We need to show that, if p and q are sufficiently close, then $\psi_q^{-1} \circ \psi_p$ is real analytic on a neighborhood of 0 in $\mathfrak{g}$. In fact, in such a case,

$$\psi_q^{-1} \circ \psi_p(X) = Y \Longrightarrow \operatorname{Exp}(Y) = q^{-1}p \operatorname{Exp}(X), \tag{3.5.25}$$

and hence, if $Z_{pq} = \operatorname{Exp}^{-1}(q^{-1}p)$, we have

$$\psi_q^{-1} \circ \psi_p(X) = \mathcal{C}(Z_{pq}, X). \tag{3.5.26}$$

The analyticity in X then follows from the explicit formula (3.5.18).

The formula (3.5.18) immediately gives analyticity of $(g,h) \mapsto gh$ for g and h in a small neighborhood of e. We now want to show that, for $p, q \in G$ fixed, $(p \operatorname{Exp} X)(q \operatorname{Exp} Y)$ is analytic in X and Y (near $0 \in \mathfrak{g}$). To see this, write

$$\begin{aligned}(p \operatorname{Exp} X)(q \operatorname{Exp} Y) &= pq(q^{-1} \operatorname{Exp} X\, q) \operatorname{Exp} Y \\ &= pq \operatorname{Exp}(\operatorname{Ad}(q^{-1})X) \operatorname{Exp} Y \\ &= pq \operatorname{Exp}(\mathcal{C}(\operatorname{Ad}(q^{-1})X, Y)),\end{aligned} \tag{3.5.27}$$

the second identity by (3.4.4). This gives the desired analyticity. The analyticity of $g \mapsto g^{-1}$ is established similarly. □

It is customary to write down a few terms in the series expansion for $\mathcal{C}(X,Y)$. We note that

$$\begin{aligned}\Psi(1+z) = (1+z)\frac{\log(1+z)}{z} &= (1+z)\Big(1-\frac{z}{2}+\frac{z^2}{3}-\cdots\Big)\\ &= 1+\sum_{k\ge 1}\frac{(-1)^{k-1}}{k(k+1)}\,z^k \\ &= 1+\frac{z}{2}-\frac{z^2}{6}+\cdots.\end{aligned} \tag{3.5.28}$$

If we set $\operatorname{ad} X = \xi$ and $\operatorname{ad} Y = \eta$, we have

$$\begin{aligned}\Psi(e^{\xi}e^{t\eta}) &= \Psi\big((I+\xi+\tfrac12\xi^2+\cdots)(I+t\eta+\tfrac12 t^2\eta^2+\cdots)\big)\\ &= \Psi(I+\xi+t\eta+\tfrac12\xi^2+t\xi\eta+\tfrac12 t^2\eta^2+\cdots)\\ &= I+\tfrac12\xi+\tfrac12 t\eta+\tfrac1{12}\xi^2+\tfrac13 t\xi\eta-\tfrac16 t\eta\xi+\tfrac1{12}t^2\eta^2+\cdots.\end{aligned} \tag{3.5.29}$$

Noting that $\eta(Y)=[Y,Y]=0$, we see that

$$\int_0^1 \Psi(e^{\operatorname{ad} X}e^{t\operatorname{ad} Y})Y\,dt = Y+\frac12\xi(Y)+\frac1{12}\xi^2(Y)-\frac1{12}\eta\xi(Y)+\cdots, \tag{3.5.30}$$

and hence

$$\mathcal{C}(X,Y)=X+Y+\frac12[X,Y]+\frac1{12}[X,[X,Y]]-\frac1{12}[Y,[X,Y]]+\cdots. \tag{3.5.31}$$

We complement (3.5.31) with a complete power series expansion, as follows. We have

$$\Psi(e^{\xi}e^{t\eta}) = I+\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k(k+1)}\big(e^{\xi}e^{t\eta}-I\big)^k, \tag{3.5.32}$$

and

$$(e^{\xi}e^{t\eta}-I)^k = \Big(\sum_{(\ell,m)\in\mathcal{S}_1}\frac{\xi^\ell}{\ell!}\,\frac{t^m\eta^m}{m!}\Big)^k, \tag{3.5.33}$$

where we set $\mathcal{S}_1=\{(\ell,m):\ell,m\ge 0,\ \ell+m>0\}$. More generally, set

$$\mathcal{S}_k=\{(\ell_1,\dots,\ell_k,m_1,\dots,m_k):\ell_j\ge 0, m_j\ge 0, \ell_j+m_j>0\}. \tag{3.5.34}$$

Then we can expand the right side of (3.5.33), to obtain

$$(e^{\xi}e^{t\eta}-I)^k = \sum_{\mathcal{S}_k} t^{m_1+\cdots+m_k}\frac{\xi^{\ell_1}}{\ell_1!}\frac{\eta^{m_1}}{m_1!}\cdots\frac{\xi^{\ell_k}}{\ell_k!}\frac{\eta^{m_k}}{m_k!}. \tag{3.5.35}$$

Plugging this into (3.5.32) and then into (3.5.20), we obtain

(3.5.36)

$$\mathcal{C}(X,Y) = X + Y + \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k(k+1)} \sum_{\mathcal{S}_k} \frac{1}{m_1 + \cdots + m_k + 1} \frac{(\operatorname{ad} X)^{\ell_1}}{\ell_1!} \frac{(\operatorname{ad} Y)^{m_1}}{m_1!} \cdots \times \frac{(\operatorname{ad} X)^{\ell_k}}{\ell_k!} \frac{(\operatorname{ad} Y)^{m_k}}{m_k!} Y.$$

Exercises

For notational simplicity, we work on matrix groups. Take $X, Y \in M(n, \mathbb{R})$.

1. If $A, B \in M(n, \mathbb{R})$, show that

$$\|e^A e^B - I\| \le e^{\|A\|} e^{\|B\|} - 1,$$

and use this to show that $\Psi(e^{\operatorname{ad} X} e^{s \operatorname{ad} Y})$ is well defined for $s \in [0, 1]$ when (3.5.15)–(3.5.16) hold.
Hint. Write $e^A e^B - I$ as a double power series.

2. Recalling that $e^X e^Y = e^{\mathcal{C}(X,Y)}$, use

$$e^Y e^X = e^Y (e^X e^Y) e^{-Y} = e^Y e^{\mathcal{C}(X,Y)} e^{-Y}$$

to show that (for $\|X\|$ and $\|Y\|$ small)

$$\mathcal{C}(Y, X) = e^Y \mathcal{C}(X, Y) e^{-Y}.$$

3. Given $C \in M(n, \mathbb{R})$, show that

$$\begin{aligned} C(s) = e^{sY} C e^{-sY} &\Rightarrow C'(s) = \operatorname{ad} Y\, C(s) \\ &\Rightarrow C(s) = e^{s \operatorname{ad} Y} C. \end{aligned}$$

Deduce that, in the setting of Exercise 2,

$$\mathcal{C}(Y, X) = e^{\operatorname{ad} Y} \mathcal{C}(X, Y).$$

Show that also

$$\mathcal{C}(Y, X) = e^{-\operatorname{ad} X} \mathcal{C}(X, Y).$$

4. Let $\mathfrak{h} \subset M(n, \mathbb{R})$ be a linear subspace, and assume

$$X, Y \in \mathfrak{h} \Longrightarrow [X, Y] \in \mathfrak{h}.$$

Show that, if X and Y belong to a suitable neighborhood of 0 in $M(n, \mathbb{R})$, i.e., (3.5.15)–(3.5.16) hold, then

$$X, Y \in \mathfrak{h} \Longrightarrow \mathcal{C}(X, Y) \in \mathfrak{h}.$$

Compare Proposition 3.2.6. Obtain an alternative proof of the existence of a subgroup $H \subset Gl(n, \mathbb{R})$ such that (3.2.37) holds.

5. Given $X, Y, Z \in M(n, \mathbb{R})$, show that $(e^X e^Y) e^Z = e^X (e^Y e^Z)$ implies

$$\mathcal{C}(\mathcal{C}(X, Y), Z) = \mathcal{C}(X, \mathcal{C}(Y, Z)),$$

for X, Y, Z suitably small. Try to establish this with $M(n, \mathbb{R})$ replaced by a general Lie algebra.

6. Now let $\mathfrak{g}$ be a general Lie algebra, and assume $\sigma : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, i.e., σ is linear and $\sigma([X, Y]) = [\sigma(X), \sigma(Y)]$ for all $X, Y \in \mathfrak{g}$. Show that, for X and Y in a suitable neighborhood of 0 in $\mathfrak{g}$,

$$\sigma\mathcal{C}(X, Y) = \mathcal{C}(\sigma(X), \sigma(Y)).$$

7. Ponder how (3.5.36) provides an answer to the question: why do lots of people hate the Campbell-Hausdorff formula? Ponder how (3.5.14) might provide the cure.

8. We say a Lie algebra $\mathfrak{g}$ is a two-step nilpotent Lie algebra provided

$$X_1, X_2 \in \mathfrak{g} \Longrightarrow (\operatorname{ad} X_1)(\operatorname{ad} X_2) = 0.$$

If $\mathfrak{g}$ is the Lie algebra of a connected Lie group G, we say G is a two-step nilpotent Lie group. Show that, in such a case, the Campbell-Hausdorff formula yields

$$(\operatorname{Exp} X_1)(\operatorname{Exp} X_2) = \operatorname{Exp} \mathcal{C}(X_1, X_2)$$

with

$$\mathcal{C}(X_1, X_2) = X_1 + X_2 + \frac{1}{2}[X_1, X_2].$$

9. Consider the group $\mathcal{H}^3$ consisting of elements

$$\begin{pmatrix} 1 & p & t \\ & 1 & q \\ & & 1 \end{pmatrix}, \quad p, q, t \in \mathbb{R}.$$

Show that $\mathcal{H}^3$ is a two-step nilpotent Lie group.

10. More generally, consider the group $\mathcal{H}^{2k+1} \subset Gl(k+2, \mathbb{R})$ consisting of elements

$$\begin{pmatrix} 1 & p^t & t \\ & 1 & q \\ & & 1 \end{pmatrix}, \quad t \in \mathbb{R},\ p, q \in \mathbb{R}^k.$$

Show that $\mathcal{H}^{2k+1}$ is a two-step nilpotent Lie group. This group is called the $(2k+1)$-dimensional Heisenberg group.

3.6. More Lie group – Lie algebra connections

We establish an essential equivalence between Lie group and Lie algebra homomorphisms. Our first result is in some sense a generalization of Proposition 3.3.2. Let G and H be Lie groups, and suppose

$$\rho : G \longrightarrow H \tag{3.6.1}$$

is a smooth group homomorphism. Denote the Lie algebras by $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and set

$$\sigma = D\rho(e) : \mathfrak{g} \longrightarrow \mathfrak{h}. \tag{3.6.2}$$

Thus, if $X \in \mathfrak{g}$, generating the one-parameter group $\gamma_X(t)$, we have

$$\rho \circ \gamma_X(t) = \gamma_{\sigma(X)}(t). \tag{3.6.3}$$

Proposition 3.6.1. *The linear map σ in* (3.6.2) *is a Lie algebra homomorphism, i.e.,*

$$X, Y \in \mathfrak{g} \Longrightarrow \sigma([X,Y]) = [\sigma(X), \sigma(Y)]. \tag{3.6.4}$$

Proof. We make use of results on the adjoint representation established in §3.4. With K_g given by (3.4.1), we have

$$\rho \circ K_g(x) = \rho(gxg^{-1}) = K_{\rho(g)}(\rho(x)). \tag{3.6.5}$$

Regard each side of (3.6.5) as a smooth function of x, mapping G to H. Differentiate each side, using the chain rule, and evaluate the derivatives at $x = e$. This yields

$$D\rho(e) \circ DK_g(e) = DK_{\rho(g)}(e) \circ D\rho(e), \tag{3.6.6}$$

or

$$\sigma \circ \mathrm{Ad}(g) = \mathrm{Ad}(\rho(g)) \circ \sigma, \tag{3.6.7}$$

as maps from $\mathfrak{g}$ to $\mathfrak{h}$. Taking $g = \gamma_X(t)$ and using (3.6.3), we have

$$\sigma \circ \mathrm{Ad}(\gamma_X(t)) = \mathrm{Ad}(\gamma_{\sigma(X)}(t)) \circ \sigma. \tag{3.6.8}$$

Taking d/dt at $t = 0$ and using (3.4.7) then gives

$$\sigma \circ \mathrm{ad}\, X = \mathrm{ad}\, \sigma(X) \circ \sigma, \tag{3.6.9}$$

which is equivalent to (3.6.4). □

We aim for a converse to Proposition 3.6.1. Suppose $\sigma : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. We desire to obtain a Lie group homomorphism. To start things off, let U be a neighborhood of $0 \in \mathfrak{g}$ such that $\mathrm{Exp} : U \to G$ is a diffeomorphism onto a neighborhood $\mathcal{O}$ of $e \in G$, and assume the Cambell-Hausdorff formula (3.5.18) holds for $t \in [0,1]$, $X, Y \in U$. Let us define

$$\rho : \mathcal{O} \longrightarrow H \tag{3.6.10}$$

by

$$\rho(g) = \operatorname{Exp}(\sigma \circ \operatorname{Exp}^{-1}(g)). \tag{3.6.11}$$

Lemma 3.6.2. *Let $\mathcal{O}_1$ be a sufficiently small neighborhood of $e \in G$. In particular assume $\mathcal{O}_1 \subset \mathcal{O}$ has the properties*

$$g_1, g_2 \in \mathcal{O}_1 \Longrightarrow g_1^{-1} \in \mathcal{O}_1, \quad g_1 g_2 \in \mathcal{O}.$$

Then

$$g_1, g_2 \in \mathcal{O}_1 \Longrightarrow \rho(g_1 g_2) = \rho(g_1)\rho(g_2). \tag{3.6.12}$$

Proof. With $X_j = \operatorname{Exp}^{-1}(g_j)$, we have $\rho(g_1 g_2) = \operatorname{Exp}(\sigma \mathcal{C}(X_1, X_2))$. Now the Campbell-Hausdorff formula (3.5.18) implies that, if σ is a Lie algebra homomorphism, then, for X_j sufficiently close to 0,

$$\sigma \mathcal{C}(X_1, X_2) = \mathcal{C}(\sigma(X_1), \sigma(X_2)). \tag{3.6.13}$$

See Exercise 5 of §3.5. Hence

$$\begin{aligned} \rho(g_1 g_2) &= \operatorname{Exp} \mathcal{C}(\sigma(X_1), \sigma(X_2)) \\ &= \operatorname{Exp} \sigma(X_1) \ \operatorname{Exp} \sigma(X_2) \\ &= \rho(g_1)\rho(g_2), \end{aligned} \tag{3.6.14}$$

as asserted. □

A map $\rho : \mathcal{O} \to H$ as in Lemma 3.6.2 is called a local homomorphism of G to H. We have the following result.

Proposition 3.6.3. *Let $\mathcal{O}_1 \subset \mathcal{O}$ and $\rho : \mathcal{O} \to H$ be as in Lemma 3.6.2. If G is simply connected, then ρ extends uniquely from $\mathcal{O}_1$ to a real-analytic homomorphism $\rho : G \to H$.*

Proof. Put a left-invariant metric on G and assume for simplicity that $\mathcal{O} = B_\delta(e) = \{g \in G : \operatorname{dist}(g, e) < \delta\}$ and $\mathcal{O}_1 = B_{\varepsilon/2}(e)$, with ε and δ small enough. Given $g \in G$, let γ be a smooth path from e to g, parametrized by arc length, say with $\gamma(0) = e$, $\gamma(L) = g$. We first define $\rho_\gamma(g)$ as follows. Pick $g_j = \gamma(t_j)$ with $0 = t_0 < t_1 < \cdots < t_N = L$ and $|t_{j+1} - t_j| < \varepsilon/2$. Thus

$$g_0 = e, \quad g_N = g, \quad g_{j+1} = x_j g_j, \quad x_j \in \mathcal{O}_1, \quad g = x_{N-1} \cdots x_2 x_1. \tag{3.6.15}$$

See Figure 3.6.1. We set

$$\rho_\gamma(g) = \rho(x_{N-1}) \cdots \rho(x_2)\rho(x_1). \tag{3.6.16}$$

First we show that $\rho_\gamma(g)$ is well defined, independent of the partition $0 = t_0 < t_1 < \cdots < t_N = L$ described above. Any two such partitions have a common refinement, so it suffices to show that refining a given partition does not change the value of $\rho_\gamma(g)$ presented in (3.6.16). So say we add one point, $t_{j+1/2} \in (t_j, t_{j+1})$. Then the factor $\rho(x_j)$ in (3.6.16) gets replaced by

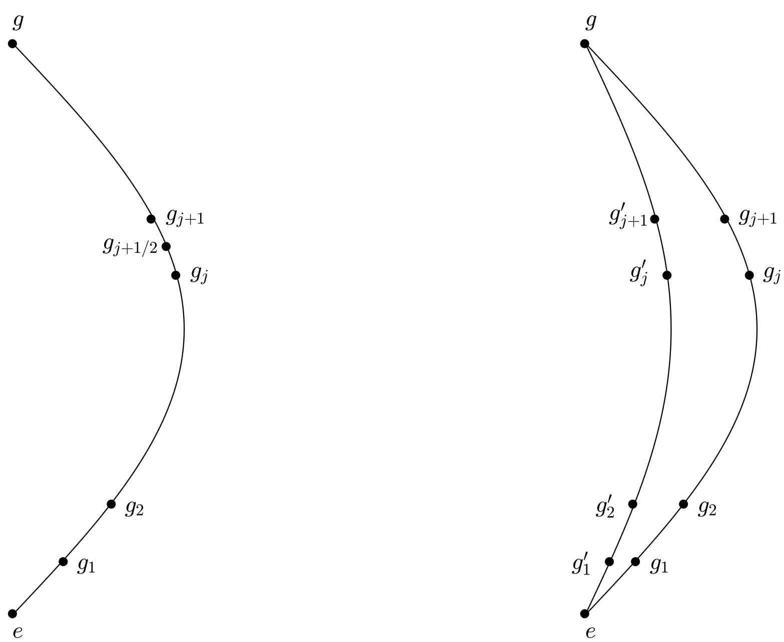

Figure 3.6.1. Defining $\rho_\gamma(g)$

$\rho(z_j)\rho(y_j)$, where $\gamma(t_{j+1/2}) = g_{j+1/2} = y_j g_j$ and $g_{j+1} = z_j g_{j+1/2}$. But x_j, y_j and z_j all belong to $\mathcal{O}_1$ and $x_j = y_j z_j$, so, by (3.6.12), $\rho(z_j)\rho(y_j) = \rho(x_j)$, and indeed (3.6.16) is not changed.

Now that we have $\rho_\gamma(g)$ well defined for a smooth path γ from e to g, we want to show that $\rho_\gamma(g)$ is independent of the path. This is where simple connectivity comes in. We will show that $\rho_\gamma(g) = \rho_\sigma(g)$ if γ and σ are smoothly homotopic paths from e to g. It suffices to show that $\rho_\gamma(g) = \rho_\sigma(g)$ when γ and σ are close enough, so assume $\sigma(t)$ is defined for $t \in [0, L]$, $\sigma(0) = e$, $\sigma(L) = g$, and assume $\operatorname{dist}(\sigma(t), \gamma(t)) < \delta/8$ for each $t \in [0, L]$. Pick a partition $0 = t_0 < t_1 < \cdots < t_N = L$ such that $|t_{j+1} - t_j| < \delta/8$. Let $g_j = \gamma(t_j)$ as in (3.6.15) and take $g'_j = \sigma(t_j)$, with

$$g'_0 = e, \quad g'_N = g, \quad g'_{j+1} = x'_j g'_j, \quad g = x'_{N-1} \cdots x'_2 x'_1. \tag{3.6.17}$$

See the right side of Figure 3.6.1. Here $\operatorname{dist}(x_j, e) < \delta/8$ and $\operatorname{dist}(x'_j, e) < \delta/8$. We also have

$$g'_j = z_j g_j, \quad \operatorname{dist}(z_j, e) < \frac{\delta}{8}. \tag{3.6.18}$$

In order to show that $\rho_\gamma(g) = \rho_\sigma(g)$, it will suffice to show that, for each k,

$$\rho_\sigma(g'_k) = \rho(z_k)\rho_\gamma(g_k), \tag{3.6.19}$$

and we do this by induction.

Clearly (3.6.19) holds for $k = 0$. Suppose it holds for $k = j - 1$. That is, we assume

$$\rho_\sigma(g'_{j-1}) = \rho(z_{j-1})\rho_\gamma(g_{j-1}), \tag{3.6.20}$$

and try to show

$$\rho_\sigma(g'_j) = \rho(z_j)\rho_\gamma(g_j), \tag{3.6.21}$$

assuming $j \le N$. In fact,

$$g'_j = x'_j z_{j-1} x_j^{-1} g_j, \quad \text{and} \quad g'_j = z_j g_j, \quad \text{so} \quad z_j = x'_j z_{j-1} x_j^{-1}, \tag{3.6.22}$$

and x'_j, x_j, z_{j-1} are all sufficiently close to the identity that (3.6.12) gives

$$\rho(z_j) = \rho(x'_j)\rho(z_{j-1})\rho(x_j^{-1}), \tag{3.6.23}$$

which does lead to (3.6.21) from (3.6.20).

At this point we can define

$$\rho : G \longrightarrow H, \quad \rho(g) = \rho_\gamma(g), \tag{3.6.24}$$

where γ is any smooth path from e to g, and we know that ρ is uniquely defined.

To show that ρ is a homomorphism in (3.6.24), take any $h \in G$ and write

$$h = y_{M-1} \cdots y_2 y_1, \quad y_j \in \mathcal{O}_1, \tag{3.6.25}$$

parallel to (3.6.15), with partial products $h_k = y_{k-1} \cdots y_2 y_1$ lying along a smooth curve from e to h. Then

$$gh = x_{N-1} \cdots x_2 x_1 y_{M-1} \cdots y_2 y_1 \tag{3.6.26}$$

has the form developed above, so the construction of ρ in (3.6.24) yields

$$\rho(gh) = \rho(x_{N-1}) \cdots \rho(x_2)\rho(x_1)\rho(y_{M-1}) \cdots \rho(y_2)\rho(y_1). \tag{3.6.27}$$

However the right side of (3.6.27) is equal to $\rho(g)\rho(h)$, so indeed

$$\rho(gh) = \rho(g)\rho(h), \quad \forall\, g, h \in G. \tag{3.6.28}$$

Finally, the analyticity of ρ on $\mathcal{O}_1$ follows from (3.6.11) and the analyticity of ρ near a general $g_0 \in G$ follows by writing $g = g_0 h$, $h \in \mathcal{O}_1$, and using $\rho(g) = \rho(g_0)\rho(h)$, plus analyticity of multiplication on G and H. $\square$

Thus we have a converse to Proposition 3.6.1:

Corollary 3.6.4. *If G is simply connected, then, for any Lie algebra homomorphism $\sigma : \mathfrak{g} \to \mathfrak{h}$, there is a unique Lie group homomorphism $\rho : G \to H$ such that $d\rho = \sigma$.*

In the setting of Corollary 3.6.4, we say the Lie algebra homomorphism σ exponentiates to the Lie group homomorphism ρ.

We return to the setting of Proposition 3.6.1 and apply it to the family of group automorphisms

$$K_g : G \longrightarrow G, \quad K_g(x) = gxg^{-1}. \tag{3.6.29}$$

Recall that

$$DK_g(e) = \mathrm{Ad}(g) : \mathfrak{g} \longrightarrow \mathfrak{g}. \tag{3.6.30}$$

Thus Proposition 3.6.1 applies to $H = G,\ \sigma = \mathrm{Ad}(g)$, to yield:

Proposition 3.6.5. *For each $g \in G$, $\mathrm{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra automorphism:*

$$\mathrm{Ad}(g)([X, Y]) = [\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y]. \tag{3.6.31}$$

Note that taking $g = \mathrm{Exp}\, tZ$ and applying d/dt at $t = 0$ simply recovers the Jacobi identity, in the following form (compare (3.1.23)–(3.1.26)):

$$\mathrm{ad}\, Z([X, Y]) = [\mathrm{ad}\, Z(X), Y] + [X, \mathrm{ad}\, Z(Y)]. \tag{3.6.32}$$

Exercises

1. Suppose $H \subset Gl(m, \mathbb{R})$ is a matrix Lie group. Show that in this case Proposition 3.6.1 follows from the combination of Proposition 3.3.2 and Proposition 3.2.5.

2. Suppose $G \subset Gl(n, \mathbb{R})$ is a matrix Lie group. Recall from Exercise 3 of §3.2 the formula

$$\mathrm{Ad}(g)A = gAg^{-1}, \quad A \in T_I G, \; g \in G.$$

Show that this leads directly to (3.6.31), i.e.,

$$\mathrm{Ad}(g)[A, B] = [\mathrm{Ad}(g)A, \mathrm{Ad}(g)B], \quad A, B \in T_I G.$$

3. Note that $\mathrm{ad}\, Z([X, Y]) = -\,\mathrm{ad}[X, Y](Z)$, and (3.1.24) says

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

Show that this is equivalent to (3.6.32).

3.7. Enveloping algebras

Associated to the Lie algebra $\mathfrak{g}$ of a Lie group G is an associative algebra $\mathfrak{U}(\mathfrak{g})$, called the universal enveloping algebra of $\mathfrak{g}$, defined as

$$\mathfrak{U}(\mathfrak{g}) = \bigotimes \mathfrak{g}_{\mathbb{C}}/J, \tag{3.7.1}$$

where $\mathfrak{g}_{\mathbb{C}}$ is the complexification of G and J is the two-sided ideal in the tensor algebra $\otimes\mathfrak{g}_{\mathbb{C}}$ generated by

$$\{XY - YX - [X, Y] : X, Y \in \mathfrak{g}\}. \tag{3.7.2}$$

It is easy to show that each element of $\mathfrak{U}(\mathfrak{g})$ defines a left-invariant differential operator on G. In fact, it can be shown that $\mathfrak{U}(\mathfrak{g})$ is isomorphic to the algebra of left-invariant differential operators on G. See §3.8 for further comments related to this.

Given a representation π of G on a finite-dimensional vector space V, there is also a representation of $\mathfrak{U}(\mathfrak{g})$, defined as follows. If

$$P = \sum_{\mu \le m} c_{i_1\cdots i_\mu} X_{i_1} \cdots X_{i_\mu}, \quad X_j \in \mathfrak{g}, \tag{3.7.3}$$

with $c_{i_1\cdots i_\mu} \in \mathbb{C}$, we have

$$d\pi(P) = \sum_{\mu \le m} c_{i_1\cdots i_\mu} d\pi(X_{i_1}) \cdots d\pi(X_{i_\mu}). \tag{3.7.4}$$

The following result is an immediate consequence of Proposition 3.3.5. As we will see it will be quite useful.

Proposition 3.7.1. *Suppose G is connected. Let $P \in \mathfrak{U}(\mathfrak{g})$ and assume*

$$PX = XP, \quad \forall\ X \in \mathfrak{g}. \tag{3.7.5}$$

If π is an irreducible unitary representation of G on V, then $d\pi(P)$ is a scalar multiple of the identity:

$$d\pi(P) = \lambda I.$$

3.8. The Poincaré-Birkhoff-Witt theorem

Given a Lie algebra $\mathfrak{g}$, the universal enveloping algebra $\mathfrak{U}(\mathfrak{g}) = \bigotimes \mathfrak{g}_{\mathbb{C}}/J$, where J is the two-sided ideal in $\bigotimes \mathfrak{g}_{\mathbb{C}}$ generated by $\{X \otimes Y - Y \otimes X - [X, Y] : X, Y \in \mathfrak{g}\}$, was introduced in §3.7. We can also form the space

$$\mathcal{P}(\mathfrak{g}) = \{P : \mathfrak{g}' \to \mathbb{C},\ \text{polynomial}\}. \tag{3.8.1}$$

There is a natural linear map

$$\beta : \mathcal{P}(\mathfrak{g}) \longrightarrow \mathfrak{U}(\mathfrak{g}), \tag{3.8.2}$$

given as follows. Say $n = \dim \mathfrak{g}$ and $\{X_1, \dots, X_n\}$ is a basis of $\mathfrak{g}$. Then $\{X_1^{\alpha_1} \cdots X_n^{\alpha_n}\}$ is a basis of $\mathcal{P}(\mathfrak{g})$, and we have

$$\beta(X_1^{\alpha_1} \cdots X_n^{\alpha_n}) = (X_1 \otimes \cdots \otimes X_1) \otimes \cdots \otimes (X_n \otimes \cdots \otimes X_n), \quad \text{mod } J, \tag{3.8.3}$$

where on the right side of (3.8.3) we have α_j factors of $X_j \otimes \cdots \otimes X_j$. The Poincaré-Birkhoff-Witt theorem is the following:

Theorem 3.8.1. *The map β in* (3.8.2) *is a linear isomorphism.*

To prove that β is surjective, we note that $\mathfrak{U}(\mathfrak{g})$ is spanned by monomials

$$X_{j_1} \otimes X_{j_2} \otimes \cdots \otimes X_{j_k} \quad (\text{mod } J). \tag{3.8.4}$$

The assertion that β is surjective is equivalent to the assertion that $\mathfrak{U}(\mathfrak{g})$ is actually spanned by monomials of the form (3.8.4) satisfying

$$j_1 \le j_2 \le \cdots \le j_k. \tag{3.8.5}$$

To see this, consider for example a monomial of the form (3.8.4) for which $j_1 > j_2$. We can rewrite it as

$$X_{j_2} \otimes X_{j_1} \otimes X_{j_3} \otimes \cdots \otimes X_{j_k} + [X_{j_1}, X_{j_2}] \otimes X_{j_3} \otimes \cdots \otimes X_{j_k} \quad (\text{mod } J),$$

that is, as a sum of two terms, the first of which is closer to satisfying the order criterion (3.8.5) and the second of which has lower order. A finite iteration rewrites each monomial (3.8.4) as a linear combination of monomials satisfying this order criterion (mod J), showing that β is surjective.

To complete the proof of Theorem 3.8.1, it remains to show that β is injective. To do this, we bring in another linear map:

$$\alpha : \mathfrak{U}(\mathfrak{g}) \longrightarrow \mathcal{D}_L(G), \tag{3.8.6}$$

where $\mathcal{D}_L(G)$ is the space of left-invariant differential operators on G. This is defined by

$$\alpha(X_{j_1} \otimes \cdots \otimes X_{j_k}) = X_{j_1} \cdots X_{j_k}, \tag{3.8.7}$$

where the right side is the product of first order differential operators $X_{j_1}, \dots, X_{j_k}$. Now

$$\alpha(X \otimes Y - Y \otimes X - [X, Y]) = XY - YX - [X, Y] = 0, \tag{3.8.8}$$

so α annihilates J, and hence (3.8.6) is well defined. Furthermore, we can compose α and β to get $\gamma = \alpha \circ \beta : \mathcal{P}(\mathfrak{g}) \to \mathcal{D}_L(G)$:

$$\begin{tikzcd} \mathcal{P}(\mathfrak{g}) \arrow[r, "\beta"] \arrow[dr, "\gamma"'] & \mathfrak{U}(\mathfrak{g}) \arrow[d, "\alpha"] \\ & \mathcal{D}_L(G) \end{tikzcd} \tag{3.8.9}$$

The formula looks tautological:

$$\gamma(X_1^{\alpha_1} \cdots X_n^{\alpha_n}) = X_1^{\alpha_1} \cdots X_n^{\alpha_n}, \tag{3.8.10}$$

but note that $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ on the left side of (3.8.10) is a polynomial function on $\mathfrak{g}'$, a product of linear functions $X_j : \mathfrak{g}' \to \mathbb{R}$, while $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ on the right side of (3.8.10) is a differential operator of order $|\alpha| = \alpha_1 + \cdots + \alpha_n$, a product of powers of first order differential operators X_j.

In order to prove that β is injective and complete the proof of Theorem 3.8.1, it suffices to prove:

Lemma 3.8.2. *The map* γ *in* (3.8.9) *is injective.*

Proof. The injectivity of γ is equivalent to the following assertion. Assume

$$\sum_{|\alpha| \le k} C_\alpha X_1^{\alpha_1} \cdots X_n^{\alpha_n} = 0 \ \text{ in } \ \mathcal{D}_L(G). \tag{3.8.11}$$

Then we assert that

$$C_\alpha = 0, \quad \forall\, \alpha. \tag{3.8.12}$$

To see this, use coordinates $(x_1, \dots, x_n)$ on $\mathfrak{g}$,

$$X = x_1 X_1 + \cdots x_n X_n, \tag{3.8.13}$$

and use the map $\operatorname{Exp} : \mathfrak{g} \to G$, a diffeomorphism of a neighborhood U of $0 \in \mathfrak{g}$ onto a neighborhood $\mathcal{O}$ of $e \in G$, to express the basis X_j of $\mathfrak{g}$ in these local exponential coordinates as

$$X_j = \sum_{\ell=1}^n A_{j\ell}(x) \frac{\partial}{\partial x_\ell}. \tag{3.8.14}$$

Note that

$$A_{j\ell}(0) = \delta_{j\ell}. \tag{3.8.15}$$

The hypothesis (3.8.11) implies that

$$\sum_{|\alpha| \le k} C_\alpha \Bigl(\sum_\ell A_{1\ell}(x) \frac{\partial}{\partial x_\ell}\Bigr)^{\alpha_1} \cdots \Bigl(\sum_\ell A_{n\ell}(x) \frac{\partial}{\partial x_\ell}\Bigr)^{\alpha_n} = 0. \tag{3.8.16}$$

Now the left side of (3.8.16) is a differential operator of order k:

$$\sum_{|\alpha|\le k} \widetilde{C}_\alpha(x)\partial_1^{\alpha_1}\cdots\partial_n^{\alpha_n}, \tag{3.8.17}$$

(with $\partial_j = \partial/\partial x_j$), and from (3.8.14)–(3.8.15) we obtain

$$\widetilde{C}_\alpha(0) = C_\alpha, \quad \forall\, |\alpha| = k. \tag{3.8.18}$$

Now if (3.8.14) holds, then $\widetilde{C}_\alpha \equiv 0$ for all α, so we deduce that $C_\alpha = 0$ whenever $|\alpha| = k$. Recording this in (3.8.11) then gives

$$\sum_{|\alpha|\le k-1} C_\alpha X_1^{\alpha_1}\cdots X_n^{\alpha_n} = 0 \ \text{ in } \ \mathcal{D}_L(G), \tag{3.8.19}$$

and iterating this argument finishes the proof of Lemma 3.8.2. □

Remark. It is also the case that α and γ in (3.8.6) and (3.8.9) are linear isomorphisms. This follows from the results proven above plus the result that

$$\alpha \ \text{ is surjective.} \tag{3.8.20}$$

For a proof of (3.8.20), see [**38**], p. 24. From these results it follows that α is an isomorphism of algebras. On the other hand, β and γ are not homomorphisms of algebras.

Chapter 4

The unitary groups $\mathrm{U}(n)$ and their representations

In this chapter we give a detailed description of the irreducible unitary representations of $\mathrm{U}(n)$ and related groups, making use of the tools developed in Chapters 2 and 3, such as Weyl orthogonality theorems and Lie algebra representations. We begin in §4.1 with a clasification of the representations of SU(2). We show that for each $k \in \mathbb{Z}^+$, there is precisely one equivalence class of irreducible unitary representations of SU(2) on $\mathbb{C}^{k+1}$. This is realized as

$$D_{k/2}(g)f(z) = f(g^{-1}z), \quad g \in SU(2),\ z \in \mathbb{C}^2, \tag{4.0.1}$$

acting on $\mathcal{P}_k \approx \mathbb{C}^{k+1}$, the space of polynomials homogeneous of degree k on $\mathbb{C}^2$, and this exhausts all the irreducible representations of SU(2). A key tool is to diagonalize a representation of SU(2) on the commutative subgroup $\mathbb{T}$, consisting of diagonal elements of SU(2). Having this result, we deduce the structure of irreducible unitary representations of the related groups, $\mathrm{U}(2), \mathrm{SO}(3)$, and SO(4).

This study provides a blueprint for a study of $\mathrm{U}(n)$, pursued in §§4.2–4.4. Given a unitary representation π of $\mathrm{U}(n)$ on V, we simultaneously diagonalize it on the "maximal torus" $\mathbb{T}$, consisting of the diagonal elements of $\mathrm{U}(n)$, forming a decomposition

$$V = \bigoplus_{\lambda \in \mathfrak{h}'} V_\lambda, \quad V_\lambda = \{v \in V : d\pi(X)v = i\lambda(X)v,\ X \in \mathfrak{h}\}, \tag{4.0.2}$$

where $\mathfrak{h}$ is the Lie algebra of $\mathbb{T}$. When $V_\lambda \neq 0$, we say λ is a weight for π, and a nonzero element $v \in V_\lambda$ is called a weight vector. In the special case where π is the adjoint representation, we say λ is a root and $v \in V_\lambda$ is a root vector. We put an order on $\mathfrak{h}'$, and show in §4.2 that each irreducible unitary representation of $\mathrm{U}(n)$ has a highest weight, that the associated space of highest weight vectors is one-dimensional, and that two such representations with the same highest weight are equivalent. This presents the problem of identifying just which elements of $\mathfrak{h}'$ arise as highest weights of irreducible representations.

Before tackling this highest weight question, we produce some examples of irreducible representations of $\mathrm{U}(n)$ in §4.3, including

$$S^k(g)f(z) = f(g^t z), \quad \overline{S}^k(g)f(z) = f(g^{-1}z), \tag{4.0.3}$$

acting on the space $\mathcal{P}_k$ of polynomials on $\mathbb{C}^n$, homogeneous of degree k. Note that $\overline{S}_k = D_{k/2}$ when $n = 2$. Another family of representations of $\mathrm{U}(n)$ is Λ^k, representing $\mathrm{U}(n)$ on the exterior algebra $\Lambda^k\mathbb{C}^n$ by

$$\Lambda^k(g)v_1 \wedge \cdots \wedge v_k = gv_1 \wedge \cdots \wedge gv_k. \tag{4.0.4}$$

In §4.4 we identify the class of highest weights, and classify the irreducible unitary representations of $\mathrm{U}(n)$ by the labels

$$\mathcal{D}_{(k_1,\dots,k_n)}, \quad k_\nu \in \mathbb{Z},\ k_1 \geq \cdots \geq k_n. \tag{4.0.5}$$

In this classification scheme,

$$S^k \approx \mathcal{D}_{(k,0,\dots,0)}, \quad \overline{S}^k \approx \mathcal{D}_{(0,\dots,0,-k)}, \tag{4.0.6}$$

and

$$\Lambda^k \approx \mathcal{D}_{(1,\dots,1,0,\dots,0)} \quad \text{(with } k \text{ ones)} \quad 0 \leq k \leq n. \tag{4.0.7}$$

In §4.5 we show that the irreducible unitary representations of $\mathrm{SU}(n)$ are precisely given by those of the form (4.0.5), this time with the equivalence

$$\mathcal{D}_{(k_1,\dots,k_n)} \approx \mathcal{D}_{(k_1+j,\dots,k_n+j)} \text{ on } \mathrm{SU}(n), \tag{4.0.8}$$

for all $j \in \mathbb{Z}$.

The analysis of representations in §§4.2–4.5 make use of the concept of $\mathrm{Gl}(n,\mathbb{C})$ as the complexification of $\mathrm{U}(n)$, and analytic continuation of representations from $\mathrm{U}(n)$ to $\mathrm{Gl}(n,\mathbb{C})$. Section 4.6 provides a second perspective on this process.

Sections 4.7–4.10 deal with the decomposition of various tensor product representations into irreducible pieces. In §4.7 we show that

$$S^k \otimes \overline{S}^\ell \approx \bigoplus_{0\leq\mu\leq k\wedge\ell} \mathcal{D}_{(k-\mu,0,\dots,0,\mu-\ell)}, \tag{4.0.9}$$

as representations of $\mathrm{U}(n)$. This extends the scope of the Clebsch-Gordon series, established in §4.1. Section 4.9 examines the linear subspace of

$$(\otimes^k \mathbb{C}^n) \otimes (\otimes^k \mathbb{C}^n)' \approx \mathrm{End}(\otimes^k \mathbb{C}^n) \tag{4.0.10}$$

on which $\mathrm{U}(n)$ acts trivially. This result is known as the First fundamental theorem of invariant theory (for unitary invariants). The analysis makes use of the fact that the groups S_k and $\mathrm{U}(n)$ act as a *dual pair* on $\otimes^k \mathbb{C}^n$, a notion introduced in §4.8.

We take up the decomposition of the representation $\otimes^k$ of $\mathrm{U}(n)$ on $\otimes^k \mathbb{C}^n$ in §4.10. In fact, this representation commutes with a natural representation τ of S_k, and these groups form a dual pair, as seen in §4.9. The major result presented in §4.10 is that there is an irreducible decomposition

$$\tau \cdot \otimes^k \approx \bigoplus_{\lambda \in F_{nk}} \mathcal{S}_\lambda \otimes \mathcal{D}_\lambda. \tag{4.0.11}$$

See §4.10 for the definitions of the objects that appear on the right side of (4.0.11). An element λ of F_{nk} is called a Young frame, and associated with it is a Young diagram. The representations $\mathcal{D}_\lambda$ of $\mathrm{U}(n)$ that arise in (4.0.11) are those depicted in (4.0.5), and $\mathcal{S}_\lambda$ are certain irreducible representations of S_k.

Section 4.11 takes up the Weyl integration formula, which implies

$$\int\limits_{\mathrm{U}(n)} f(g)\, dg = C_n \int\limits_{\mathbb{T}^n} f(D(\theta)) J(\theta)\, d\theta, \tag{4.0.12}$$

where $f \in C(G)$ is a central function, $\mathbb{T}^n$ is the set of diagonal elements of $\mathrm{U}(n)$, $D(\theta) = \mathrm{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$, and

$$J(\theta) = \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2. \tag{4.0.13}$$

Major examples of central functions are the characters $\chi_\pi(g) = \mathrm{Tr}\, \pi(g)$ of representations π of $\mathrm{U}(n)$. Section 4.12 produces the Weyl character formula

$$\chi_\lambda(D(\theta)) = \mathrm{Tr}\, \mathcal{D}_\lambda(D(\theta)) = \frac{A_{\lambda+\rho}(\theta)}{A_\rho(\theta)}, \tag{4.0.14}$$

where, for $\mu \in \mathfrak{h}'$,

$$A_\mu(\theta) = \sum_{\sigma \in S_n} (\mathrm{sgn}\, \sigma) e^{i\sigma \cdot \mu(\theta)}. \tag{4.0.15}$$

Here ρ is half the sum of the positive roots,

$$\rho = \frac{1}{2} \sum_{j<k} \omega_{jk}, \tag{4.0.16}$$

where $\{\omega_{jk}\}$ are the roots, introduced in §4.2, which are positive for $j < k$. We regard $\theta \in \mathfrak{h}$ to evaluate $\mu(\theta)$ for $\mu \in \mathfrak{h}'$. Evaluation at $\theta = 0$ gives the Weyl dimension formula

$$d_\lambda = \prod_{j<k} \frac{\langle \omega_{jk}, \lambda + \rho \rangle}{\langle \omega_{jk}, \rho \rangle}. \tag{4.0.17}$$

Section 4.13 evaluates the characters of the representations S^k and Λ^k, introduced in (4.0.3)–(4.0.4).

Section 4.14 looks at the character of the representation (4.0.11),

$$\operatorname{Tr} \tau(\sigma) \cdot \otimes^k g = \sum_{\lambda \in F_{nk}} \chi^S_\lambda(\sigma) \chi_\lambda(g), \tag{4.0.18}$$

for $\sigma \in S_k,\ g \in \mathrm{U}(n)$, with χ_λ given by (4.0.14), and derives a formula for $\chi^S_\lambda(\sigma)$, and for the dimension d^S_λ of the associated representation space.

Section 4.15 treats an integral that arises in random matrix theory, and establishes that

$$\int_{\mathrm{U}(n)} |\operatorname{Tr} g^k|^2 \, dg = k \wedge n. \tag{4.0.19}$$

Connections are made with results of §§4.10–4.11, and applications are indicated to formulas for

$$\int_{\mathrm{U}(n)} X_u(g) X_v(g) \, dg, \tag{4.0.20}$$

where $X_f(g) = \operatorname{Tr} f(g)$, with $f(g)$ defined by the spectral representation of $g \in \mathrm{U}(n)$, when $f : S^1 \to \mathbb{C}$ is a continuous function. Further material on this appears in §E.5.

4.1. Representations of SU(2) and related groups

Recall that SU(2) is the group of 2×2 complex unitary matrices of determinant 1, i.e.,

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} z_1 & z_2 \\ -\overline{z}_2 & \overline{z}_1 \end{pmatrix} : |z_1|^2 + |z_2|^2 = 1,\ z_j \in \mathbb{C} \right\}. \tag{4.1.1}$$

As a set, SU(2) is naturally identified with the unit sphere S^3 in $\mathbb{C}^2$. Its Lie algebra $\mathfrak{su}(2)$ consists of 2×2 complex skew adjoint matrices of trace zero. A basis of $\mathfrak{su}(2)$ is formed by

$$X_1 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad X_2 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad X_3 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{4.1.2}$$

Note the commutation relations

$$[X_1, X_2] = X_3, \quad [X_2, X_3] = X_1, \quad [X_3, X_1] = X_2. \tag{4.1.3}$$

We also recall that the group SO(3) is the group of linear isometries of $\mathbb{R}^3$ with determinant 1. Its Lie algebra so(3) is spanned by elements $J_\ell,\ \ell = 1, 2, 3$, which generate rotations about the x_ℓ-axis. One readily verifies that these satisfy the same commutation relations as in (4.1.3). Thus SU(2) and SO(3) have isomorphic Lie algebras. There is an explicit homomorphism

$$p : \mathrm{SU}(2) \longrightarrow \mathrm{SO}(3), \tag{4.1.4}$$

which exhibits SU(2) as a double cover of SO(3). One way to construct p is the following. The linear span $\mathfrak{g}$ of (4.1.2) over $\mathbb{R}$ is a three-dimensional real vector space, with an inner product given by $(X, Y) = -\operatorname{Tr} XY$. It is clear that the representation p of SU(2) by a group of linear transformations on $\mathfrak{g}$ given by $p(g)X = gXg^{-1}$ preserves this inner product and gives (4.1.4). Note that Ker $p = \{I, -I\}$. (Note also that $p(g) = \mathrm{Ad}(g)$.)

If we regard X_j as left-invariant vector fields on SU(2), set

$$\Delta = X_1^2 + X_2^2 + X_3^2, \tag{4.1.5}$$

a second-order, left-invariant differential operator. It follows easily from (4.1.3) that X_j and Δ commute:

$$\Delta X_j = X_j \Delta, \quad 1 \le j \le 3. \tag{4.1.6}$$

Suppose π is an irreducible unitary representation of SU(2) on V. Then π induces a skew-adjoint representation $d\pi$ of the Lie algebra $\mathfrak{su}(2)$, and an algebraic representation of the universal enveloping algebra. By (4.1.6), $d\pi(\Delta)$ commutes with $d\pi(X_j),\ j = 1, \dots, 3$. Thus, if π is irreducible, Proposition 3.7.1 implies

$$d\pi(\Delta) = -\lambda^2 I, \tag{4.1.7}$$

for some $\lambda \in \mathbb{R}$. (Since $d\pi(\Delta)$ is a sum of squares of skew-adjoint operators, it must be negative.) Let

$$L_j = d\pi(X_j). \tag{4.1.8}$$

Now we will diagonalize L_1 on V. Set

$$V_\mu = \{v \in V : L_1 v = i\mu v\}, \quad V = \bigoplus_{i\mu \in \text{ Spec } L_1} V_\mu. \tag{4.1.9}$$

The structure of π is defined by how L_2 and L_3 behave on V_μ. It is convenient to set

$$L_\pm = L_2 \mp iL_3, \tag{4.1.10}$$

i.e., $L_\pm = d\pi(X_\pm)$ where

$$X_+ = X_2 - iX_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_- = X_2 + iX_3 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}. \tag{4.1.11}$$

We have the following key identities, as a direct consequence of (4.1.3):

$$[X_1, X_\pm] = \pm iX_\pm, \quad \text{hence} \quad [L_1, L_\pm] = \pm iL_\pm. \tag{4.1.12}$$

Using this, we can establish the following:

Lemma 4.1.1. *We have*

$$L_\pm : V_\mu \longrightarrow V_{\mu\pm 1}. \tag{4.1.13}$$

In particular, if $i\mu \in$ Spec L_1, then either $L_+ = 0$ on V_μ or $i(\mu + 1) \in$ Spec L_1, and also either $L_- = 0$ on V_μ or $i(\mu - 1) \in$ Spec L_1.

Proof. Let $v \in V_\mu$. By (4.1.12) we have

$$L_1 L_\pm v = L_\pm L_1 v \pm iL_\pm v = i(\mu \pm 1)L_\pm v,$$

which establishes the lemma. The operators $L_\pm$ are called "ladder operators." □

To continue, if π is irreducible on V, we claim that Spec $(1/i)L_1$ must consist of a sequence

$$\text{Spec } \frac{1}{i} L_1 = \{\mu_0, \mu_0 + 1, \dots, \mu_0 + k = \mu_1\}, \tag{4.1.14}$$

with

$$L_+ : V_{\mu_0+j} \to V_{\mu_0+j+1} \quad \text{isomorphism, for} \quad 0 \le j \le k - 1, \tag{4.1.15}$$

and

$$L_- : V_{\mu_1-j} \to V_{\mu_1-j-1} \quad \text{isomorphism, for} \quad 0 \le j \le k - 1. \tag{4.1.16}$$

In fact, we can compute

$$L_- L_+ = L_2^2 + L_3^2 + i[L_3, L_2] = -\lambda^2 - L_1^2 - iL_1 \tag{4.1.17}$$

on V, and

$$L_+L_- = -\lambda^2 - L_1^2 + iL_1 \tag{4.1.18}$$

on V, so

$$\begin{aligned} L_-L_+ &= \mu(\mu+1) - \lambda^2 \quad \text{on} \quad V_\mu, \\ L_+L_- &= \mu(\mu-1) - \lambda^2 \quad \text{on} \quad V_\mu. \end{aligned} \tag{4.1.19}$$

Note that, since L_2 and L_3 are skew-adjoint, $L_+ = -L_-^*$, so

$$L_+L_- = -L_-^*L_-, \quad L_-L_+ = -L_+^*L_+.$$

Thus

$$\text{Ker } L_+ = \text{ Ker } L_-L_+, \quad \text{Ker } L_- = \text{ Ker } L_+L_-. \tag{4.1.20}$$

From (4.1.19) we see that, if $\mu_0 = \min \operatorname{Spec}(1/i)L_1$ and $\mu_1 = \max \operatorname{Spec}(1/i)L_1$, then, since $L_+ = 0$ on V_{μ_1} and $L_- = 0$ on V_{μ_0},

$$\mu_1(\mu_1+1) = \lambda^2 = \mu_0(\mu_0-1).$$

Hence

$$\mu_1 - \mu_0 = k \Longrightarrow \mu_0 = -\frac{k}{2}, \quad \mu_1 = \frac{k}{2}, \quad \lambda^2 = \frac{1}{4}k(k+2). \tag{4.1.21}$$

If L_+ is not injective on V_μ, then, by (4.1.19)–(4.1.20), $L_+ = 0$ on V_μ, so (by (4.1.19)),

$$\mu = -\frac{1}{2} \pm \frac{1}{2}\sqrt{1+4\lambda^2} = -\frac{1}{2} \pm \frac{1}{2}(k+1),$$

i.e., $\mu = k/2 = \mu_1$ or $\mu = -k/2 - 1 = \mu_0 - 1$, which is not allowed. Similarly, if L_- is not injective on V_μ, then

$$\mu = \frac{1}{2} \pm \frac{1}{2}\sqrt{1+4\lambda^2} = \frac{1}{2} \pm \frac{1}{2}(k+1),$$

i.e., $\mu = -k/2 = \mu_0$ or $\mu = k/2 + 1 = \mu_1 + 1$, which is not allowed. These observations establish (4.1.14)–(4.1.16).

Considering that $d\pi$ preserves the linear span of $\{v, L_+v, \dots L_+^{\mu_1-\mu_0}v\}$ for any nonzero $v \in V_{\mu_0}$, and that irreducibility implies this must be all of V, we have

$$\dim V_\mu = 1, \quad \mu_0 \le \mu \le \mu_1. \tag{4.1.22}$$

Hence we have

$$\dim V = k+1, \quad \lambda^2 = \frac{1}{4}k(k+2) = \frac{1}{4}(\dim V^2 - 1). \tag{4.1.23}$$

A nonzero element $v \in V$ such that $L_+v = 0$ is called a "highest weight vector" for the representation π of SU(2) on V. It follows from the analysis above that all highest weight vectors for an irreducible representation on V belong to the one-dimensional space V_{μ_1}.

The calculations above establish that an irreducible unitary representation π of SU(2) on V is determined uniquely up to equivalence by dim V. We are ready to prove:

Proposition 4.1.2. *There is precisely one equivalence class of irreducible unitary representation of* SU(2) *on* $\mathbb{C}^{k+1}$, *for each* $k = 0, 1, 2, \dots$.

We will realize each such representation, which is denoted $D_{k/2}$, on the space

$$\mathcal{P}_k = \{p(z) : p \text{ homogeneous polynomial of degree } k \text{ on } \mathbb{C}^2\}, \tag{4.1.24}$$

with SU(2) acting on $\mathcal{P}_k$ by

$$D_{k/2}(g)f(z) = f(g^{-1}z), \quad g \in SU(2),\ z \in \mathbb{C}^2. \tag{4.1.25}$$

Note that, for $X \in \mathrm{su}(2)$,

$$dD_{k/2}(X)f(z) = \frac{d}{dt} f\big(e^{-tX}z\big)\big|_{t=0} = -(\partial_1 f, \partial_2 f) \cdot X \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \tag{4.1.26}$$

where $\partial_j f = \partial f / \partial z_j$. A calculation gives

$$\begin{aligned} L_1 f(z) &= -\frac{i}{2}(z_1\partial_1 f - z_2\partial_2 f), \\ L_2 f(z) &= -\frac{1}{2}(z_2\partial_1 f - z_1\partial_2 f), \\ L_3 f(z) &= -\frac{i}{2}(z_2\partial_1 f + z_1\partial_2 f). \end{aligned} \tag{4.1.27}$$

In particular, for

$$\varphi_{kj}(z) = z_1^{k-j} z_2^j \in \mathcal{P}_k, \quad 0 \le j \le k, \tag{4.1.28}$$

we have

$$L_1 \varphi_{kj} = i\Big(-\frac{k}{2} + j\Big)\varphi_{kj}, \tag{4.1.29}$$

so

$$V = \mathcal{P}_k \Longrightarrow \text{ span } \varphi_{kj} = V_{-k/2+j}, \quad 0 \le j \le k. \tag{4.1.30}$$

Note that

$$L_+ f(z) = -z_2 \partial_1 f(z), \quad L_- f(z) = z_1 \partial_2 f(z), \tag{4.1.31}$$

so

$$L_+ \varphi_{kj} = -(k-j)\varphi_{k,j+1}, \quad L_- \varphi_{kj} = j\varphi_{k,j-1}. \tag{4.1.32}$$

We see that the structure of the representation $D_{k/2}$ of SU(2) on $\mathcal{P}_k$ is as described in (4.1.13)–(4.1.23). The last detail is to show that $D_{k/2}$ is irreducible. If not, then $\mathcal{P}_k$ splits into a direct sum of several irreducible subspaces, each of which have a one-dimensional space of highest weight vectors, annihilated by L_+. But as seen above, within $\mathcal{P}_k$, only multiples of z_2^k are annihilated by L_+, so the representation $D_{k/2}$ of SU(2) on $\mathcal{P}_k$ is irreducible.

Representations of $SO(3)$

We can deduce the classification of irreducible unitary representations of SO(3) from the result above as follows. We have the covering homomorphism (4.1.4), and Ker $p = \{\pm I\}$. Now each irreducible representation d_j of SO(3) defines an irreducible representation $d_j \circ p$ of SU(2), which must be equivalent to one of the representations $D_{k/2}$ described above. On the other hand, $D_{k/2}$ factors through to yield a representation of SO(3) if and only if $D_{k/2}$ is the identity on Ker p, i.e., if and only if $D_{k/2}(-I) = I$. Clearly this holds if and only if k is even, since

$$D_{k/2}(-I) = (-1)^k I.$$

Thus all the irreducible unitary representations of SO(3) are given by representations $\widetilde{D}_j$ on $\mathcal{P}_{2j}$, uniquely defined by

$$\widetilde{D}_j\big(p(g)\big) = D_j(g), \quad g \in \mathrm{SU}(2). \tag{4.1.33}$$

It is conventional to use D_j instead of $\widetilde{D}_j$ to denote such a representation of SO(3). Note that D_j represents SO(3) on a space of dimension $2j+1$, and

$$dD_j(\Delta) = -j(j+1). \tag{4.1.34}$$

Representations of $U(2)$

Also we can classify the irreducible representations of U(2), using the results on SU(2). To do this, use the exact sequence

$$1 \to K \to S^1 \times \mathrm{SU}(2) \to \mathrm{U}(2) \to 1, \tag{4.1.35}$$

where "1" denotes the trivial multiplicative group, and

$$K = \{(\omega, g) \in S^1 \times \mathrm{SU}(2) : g = \omega^{-1} I, \omega^2 = 1\} = \{\pm(1, I)\}. \tag{4.1.36}$$

The irreducible representations of $S^1 \times \mathrm{SU}(2)$ are given by

$$\pi_{mk}(\omega, g) = \omega^m D_{k/2}(g) \ \text{ on } \ \mathcal{P}_k, \tag{4.1.37}$$

with $m, k \in \mathbb{Z}$, $k \geq 0$. Those giving a complete set of irreducible representations of U(2) are those for which $\pi_{mk}(K) = I$, i.e., those for which $(-1)^m D_{k/2}(-I) = I$. Since $D_{k/2}(-I) = (-1)^k I$, we see the condition is that $m + k$ be an even integer.

For another perspective on the irreducible representations of U(2), note that (4.1.25), i.e., $\mathcal{D}_{k/2}(g)p(z) = p(z^{-1}z)$, for $p \in \mathcal{P}_k$, is clearly well defined for $g \in \mathrm{U}(2)$. If $g = \omega g_0$, with $|\omega| = 1$, $g_0 \in \mathrm{SU}(2)$, then

$$\mathcal{D}_{k/2}(\omega g_0) = \omega^{-k} \mathcal{D}_{k/2}(g_0) = \pi_{kk}(\omega, g_0),$$

so the general irreducible representation of U(2) has the form

$$\pi_{k+2j,k}(\omega, g_0) = (\det \omega g_0)^j \mathcal{D}_{k/2}(\omega g_0), \quad j \in \mathbb{Z},\ k \in \mathbb{Z}^+,$$

i.e.,

$$(\det g)^j\, \mathcal{D}_{k/2}(g), \quad j \in \mathbb{Z},\ k \in \mathbb{Z}^+. \tag{4.1.38}$$

Representations of $SO(4)$

We now consider the representations of SO(4). First note that SO(4) is covered by SU(2) × SU(2). To see this, equate the unit sphere $S^3 \subset \mathbb{R}^4$, with its standard metric, to SU(2), with a bi-invariant metric. Then SO(4) is the identity component in the isometry group of S^3. Meanwhile, SU(2) × SU(2) acts as a group of isometries, by

$$(g_1, g_2) \cdot x = g_1 x g_2^{-1}, \quad g_j \in \mathrm{SU}(2), \quad x \in \mathrm{SU}(2) \approx S^3. \tag{4.1.39}$$

Thus we have a map

$$\tau : \mathrm{SU}(2) \times \mathrm{SU}(2) \longrightarrow \mathrm{SO}(4). \tag{4.1.40}$$

This is a group homomorphism. Note that $(g_1, g_2) \in$ Ker τ implies $g_1 = g_2 = \pm I$. In fact, if $g_1 x g_2^{-1} = x$ for all $x \in \mathrm{SU}(2)$, taking $x = e$ implies $g_1 = g_2$, then Schur's lemma implies $g_1 = g_2$ is a scalar, and then $g_1 \in \mathrm{SU}(2)$ implies $g_1 = g_2 = \pm I$. Furthermore, a dimension count shows τ must be surjective, so

$$\mathrm{SO}(4) \approx \mathrm{SU}(2) \times \mathrm{SU}(2)/\{\pm(I, I)\}. \tag{4.1.41}$$

As shown in Proposition 2.8.11, if G_1 and G_2 are compact Lie groups, and $G = G_1 \times G_2$, then the set of all irreducible unitary representations of G, up to unitary equivalence, is given by

$$\{\pi(g) = \pi_1(g_1) \otimes \pi_2(g_2) : \pi_j \in \widehat{G}_j\}, \tag{4.1.42}$$

where $g = (g_1, g_2) \in G$ and $\widehat{G}_j$ parametrizes the irreducible unitary representations of G_j. In particular, the irreducible unitary representations of SU(2)×SU(2), up to equivalence, are precisely the representations of the form

$$\gamma_{k\ell}(g) = D_{k/2}(g_1) \otimes D_{\ell/2}(g_2), \quad k, \ell \in \{0, 1, 2, \dots\}, \tag{4.1.43}$$

acting on $\mathcal{P}_k \otimes \mathcal{P}_\ell \approx \mathbb{C}^{k+1} \otimes \mathbb{C}^{\ell+1}$. By (4.1.41), the irreducible unitary representations of SO(4) are given by all $\gamma_{k\ell}$ such that $k + \ell$ is even, since, for $p_0 = (-I, -I) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$, $\gamma_{k\ell}(p_0) = (-1)^{k+\ell} I$.

Tensor products and the Clebsch-Gordon series

We next consider the problem of decomposing the tensor product representations $D_{k/2} \otimes D_{\ell/2}$ of SU(2), i.e., the composition of (4.1.43) with the diagonal map $\mathrm{SU}(2) \hookrightarrow \mathrm{SU}(2) \times \mathrm{SU}(2)$, into irreducible representations. We may as well assume that $\ell \le k$. Note that $\pi_{k\ell} = D_{k/2} \otimes D_{\ell/2}$ acts on

$$\begin{aligned} \mathcal{P}_{k\ell} = \{ f(z,w) &: \text{polynomial on } \mathbb{C}^2 \times \mathbb{C}^2, \\ &\text{homogeneous of degree } k \text{ in } z,\ \ell \text{ in } w\}, \end{aligned} \tag{4.1.44}$$

as

$$\pi_{k\ell}(g) f(z,w) = f(g^{-1}z, g^{-1}w). \tag{4.1.45}$$

Parallel to (4.1.27) and (4.1.31), we have, on $\mathcal{P}_{k\ell}$,

$$\begin{aligned} L_1 f &= -\frac{i}{2}(z_1 \partial_{z_1} f - z_2 \partial_{z_2} f + w_1 \partial_{w_1} f - w_2 \partial_{w_2} f), \\ L_+ f &= -z_2 \partial_{z_1} f - w_2 \partial_{w_1} f, \\ L_- f &= z_1 \partial_{z_2} f + w_1 \partial_{w_2} f. \end{aligned} \tag{4.1.46}$$

To decompose $\mathcal{P}_{k\ell}$ into irreducible subspaces, we specify Ker L_+. In fact, a holomorphic function $f(z,w)$ annihilated by L_+ is of the form

$$f(z,w) = g(z_2, w_2, w_2 z_1 - z_2 w_1). \tag{4.1.47}$$

In more detail, if $L_+ f(z,w) = 0$, then $f(e^{tX_+}z, e^{tX_+}w)$ is independent of t, with X_+ as in (4.1.11), i.e.,

$$f(z_1 + tz_2, z_2, w_1 + tw_2, w_2) \ \text{ is independent of } \ t. \tag{4.1.48}$$

If $z_2 \ne 0$, take $t = -z_1/z_2$, to conclude

$$f(z_1, z_2, w_1, w_2) = f(0, z_2, w_1 - (z_1/z_2)w_2), w_2).$$

If $f \in \mathcal{P}_{k\ell}$ and $k \ge \ell$, this yields

$$f(z_1, z_2, w_1, w_2) = z_2^{k-\ell} f(0, 1, w_1 z_2 - z_1 w_2, z_2 w_2),$$

which also holds at $z_2 = 0$.

Thus the kernel of L_+ in $\mathcal{P}_{k\ell}$ is the linear span of

$$\psi_{k\ell\mu}(z,w) = z_2^{k-\mu} w_2^{\ell-\mu} (w_2 z_1 - z_2 w_1)^\mu, \quad 0 \le \mu \le \ell. \tag{4.1.49}$$

A calculation gives

$$L_1 \psi_{k\ell\mu} = \frac{i}{2}(k + \ell - 2\mu) \psi_{k\ell\mu}. \tag{4.1.50}$$

In fact, since $e^{-tX_1} z = (e^{-(i/2)t} z_1, e^{(i/2)t} z_2)$, and similarly for $e^{-tX_1} w$, we see that

$$\psi_{k\ell\mu}(e^{-tX_1} z, e^{-tX_1} w) = e^{(i/2)(k+\ell-2\mu)t} \psi_{k\ell\mu}(z,w), \tag{4.1.51}$$

which gives (4.1.50).

It follows that, for fixed $k, \ell,\ 0 \le \ell \le k$, and for each $\mu = 0, \dots, \ell,\ \psi_{k\ell\mu}$ is the highest weight vector of a representation equivalent to $D_{(k+\ell-2\mu)/2}$, so we have

$$
\begin{aligned}
D_{k/2} \otimes D_{\ell/2} &\approx \bigoplus_{\mu=0}^{\ell} D_{(k+\ell-2\mu)/2} \\
&= D_{(k-\ell)/2} \oplus D_{(k-\ell)/2+1} \oplus \cdots \oplus D_{(k+\ell)/2}.
\end{aligned}
\tag{4.1.52}
$$

This is called the Clebsch-Gordon series.

We make some general comments about decomposing a unitary representation π of SU(2) on V into irreducible pieces. First, one identifies

$$
K = \operatorname{Ker} L_+ \subset V. \tag{4.1.53}
$$

We know that π splits into mutually orthogonal irreducible pieces, $\pi_1 \oplus \cdots \oplus \pi_M$, on $V_1 \oplus \cdots \oplus V_M = V$, and K is spanned by the one-dimensional highest weight subspaces of each V_j, each of them eigenspaces of L_1. Hence

$$
L_1 : K \longrightarrow K, \tag{4.1.54}
$$

and of course $L_1|_K$ is skew-adjoint. To find the pieces V_j, one diagonalizes $L_1|_K$ (and each space V_j is spanned by the images of an eigenvector of $L_1|_K$ under L_- and its powers). This procedure can be seen to have been followed in the decomposition described above of $D_{k/2} \otimes D_{\ell/2}$ on $\mathcal{P}_{k\ell}$.

Note also how it has been convenient to analyze $\operatorname{Ker} L_+$ and eigenspaces of L_1 via passage back to the group SU(2), via

$$
\begin{aligned}
L_+ v = 0 &\Longleftrightarrow \pi(e^{tX_+})v = v, \quad \forall\, t, \\
L_1 v = i\mu v &\Longleftrightarrow \pi(e^{tX_1})v = e^{i\mu t} v, \quad \forall\, t.
\end{aligned}
\tag{4.1.55}
$$

Compare (4.1.48) and (4.1.51). Analogous observations will be useful in §4.2.

Characters of $D_{k/2}$

We turn our attention to the representations $D_{k/2}$ of $SU(2)$ on $\mathcal{P}_k$, given by (4.1.24)–(4.1.25), and discuss the characters

$$
\chi_{k/2}(g) = \operatorname{Tr} D_{k/2}(g). \tag{4.1.56}
$$

These are central functions, and they are uniquely determined by their values on the one-dimensional subgroup

$$
\mathbb{T} = \left\{ \begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{pmatrix} = u(\theta) : \theta \in \mathbb{R}/2\pi\mathbb{Z} \right\}, \tag{4.1.57}
$$

since every element of $SU(2)$ is conjugate to some element $u(\theta)$. Taking the basis $\varphi_{kj}(z) = z_1^{k-j} z_2^j \in \mathcal{P}_k$ as in (4.1.28)–(4.1.29), we see that

$$\begin{aligned} D_{k/2}(u(\theta))\varphi_{kj} &= e^{-i(k-j)\theta} e^{ij\theta} \varphi_{kj} \\ &= e^{i(2j-k)\theta} \varphi_{kj}. \end{aligned} \tag{4.1.58}$$

Hence

$$\chi_{k/2}(u(\theta)) = \sum_{j=0}^{k} e^{i(2j-k)\theta}. \tag{4.1.59}$$

One way to write this sum is as

$$\chi_{k/2}(u(\theta)) = \sum_{j=0}^{k} \cos(-k+2j)\theta. \tag{4.1.60}$$

In particular, we have

$$\chi_0(u(\theta)) = 1, \quad \chi_{1/2}(u(\theta)) = 2\cos\theta, \tag{4.1.61}$$

and, inductively,

$$\chi_{k/2+1}(u(\theta)) = \chi_{k/2}(u(\theta)) + 2\cos(k+2)\theta. \tag{4.1.62}$$

It follows that

$$\operatorname{Span}\{\chi_{k/2} \circ u : 0 \le k \le N\} = \operatorname{Span}\{\cos k\theta : 0 \le k \le N\}, \tag{4.1.63}$$

hence

$$\operatorname{Span}\{\chi_{k/2} \circ u : k \in \mathbb{Z}^+\} \ \text{ is dense in } \ C_e(\mathbb{T}), \tag{4.1.64}$$

the space of even continuous functions on $\mathbb{R}/2\pi\mathbb{Z}$. In connection with this, note that

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Longrightarrow J u(\theta) J^{-1} = u(-\theta), \tag{4.1.65}$$

so every continuous central function on $SU(2)$ restricts to an element of $C_e(\mathbb{T})$. This yields the following.

Proposition 4.1.3. *The space*

$$\operatorname{Span}\{\chi_{k/2} : k \in \mathbb{Z}^+\} \tag{4.1.66}$$

is dense in the space $C_{\mathcal{C}}(SU(2))$ of continuous central functions, hence also in the space $L^2_{\mathcal{C}}(SU(2))$ of L^2 central functions.

Proof. This follows from (4.1.64) plus the fact that, for each $f \in C_{\mathcal{C}}(SU(2))$,

$$\sup_{g \in SU(2)} |f(g)| = \sup_{g \in \mathbb{T}} |f(g)|. \tag{4.1.67}$$

□

This argument yields a second proof of Proposition 4.1.2:

Corollary 4.1.4. *The representations $D_{k/2}$ of $SU(2)$ on $\mathcal{P}_k$ form a complete set of irreducible unitary representations of $SU(2)$.*

Proof. This follows from Proposition 4.1.3 plus the orthogonality relations for the characters of irreducible representations established in §2.4. □

Note that (4.1.59) represents $\chi_{k/2}(u(\theta))$ as a geometric series. Summing this series gives the compact formula

$$\begin{aligned} \chi_{k/2}(u(\theta)) &= \frac{e^{i(k+1)\theta} - e^{-i(k+1)\theta}}{e^{i\theta} - e^{-i\theta}} \\ &= \frac{\sin(k+1)\theta}{\sin\theta}. \end{aligned} \tag{4.1.68}$$

We move on to an application of the character formula to an alternative derivation of the Clebsch-Gordon formula, (4.1.52).

Proposition 4.1.5. *Given $k, \ell \in \mathbb{Z}^+$, $g \in SU(2)$,*

$$\chi_{k/2}(g)\chi_{\ell/2}(g) = \chi_{|k-\ell|/2}(g) + \chi_{|k-\ell|/2+1}(g) \cdots + \chi_{(k+\ell)/2}(g). \tag{4.1.69}$$

Proof. It suffices to establish (4.1.69) for $g = u(\theta) \in \mathbb{T}$. With $\omega = e^{i\theta}$, write the character formula (4.1.58) as

$$\chi_{k/2}(\omega) = \sum_{j=0}^{k} \omega^{2j-k}. \tag{4.1.70}$$

Then

$$\chi_{k/2}(\omega)\chi_{\ell/2}(\omega) = \sum_{j=0}^{k} \sum_{m=0}^{\ell} \omega^{2(j+m)-k-\ell}. \tag{4.1.71}$$

There is no loss in generality in assuming $k \le \ell$. Then the double sum (4.1.71) can be rearranged to yield

$$\sum_{a=0}^{k}(a+1)\omega^{2a-k-\ell} + \sum_{a=k+1}^{\ell-1}(k+1)\omega^{2a-k-\ell} + \sum_{a=\ell}^{\ell+k}(k+\ell+1-a)\omega^{2a-k-\ell}. \tag{4.1.72}$$

We can set $b = k + \ell - a$ and write the last sum in (4.1.72) as

$$\sum_{b=0}^{k}(b+1)\omega^{k+\ell-2b}. \tag{4.1.73}$$

Now we can rewrite (4.1.72) as a sum of $k+1$ strings each string, labeled by $a \in \{0.\dots, k\}$, having the form

$$\begin{aligned} \sum \omega^M, \quad &\text{for} \quad -(k+\ell) + 2a \le M \le (k+\ell) - 2a, \\ &\text{proceeding by even integer increments.} \end{aligned} \tag{4.1.74}$$

Such a sum is equal to

$$\chi_{(k+\ell)/2-a}(\omega), \quad 0 \le a \le k, \tag{4.1.75}$$

i.e., for $k \le \ell$, to

$$\chi_{(\ell-k)/2}(\omega),\ \chi_{(\ell-k)/2+1}(\omega),\ \cdots,\ \chi_{(\ell+k)/2}(\omega). \tag{4.1.76}$$

This proves (4.1.69). $\square$

Exercises

1. Show that the elements $X_j \in \mathfrak{su}(2)$ given by (4.1.2) are an orthogonal set with respect to the inner product $(X, Y) = \operatorname{Tr} X^*Y$. Verify the commutation relations (4.1.3).

2. With X_k as in (4.1.2), consider

$$J_k = \operatorname{ad} X_k : \mathfrak{su}(2) \longrightarrow \mathfrak{su}(2).$$

Show that the matrix representations J_k with respect to the orthonormal basis $\{X_1, X_2, X_3\}$ are

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Deduce that

$$[J_1, J_2] = J_3, \quad [J_2, J_3] = J_1, \quad [J_3, J_1] = J_2.$$

Note that

$$e^{2\pi X_k} = -I, \quad e^{2\pi J_k} = I.$$

Relate this to (4.1.4).

3. Verify that, for each $j, k \in \{1, 2, 3\}$,

$$[X_j, X_k^2] = [X_j, X_k]X_k + X_k[X_j, X_k],$$

hence

$$\sum_k [X_j, X_k^2] = \sum_{k \neq j} \big\{[X_j, X_k]X_k + X_k[X_j, X_k]\big\}.$$

In particular,

$$\begin{aligned} \sum_k [X_1, X_k^2] &= \sum_{k=2}^{3} \big\{[X_1, X_k]X_k + X_k[X_1, X_k]\big\} \\ &= X_3X_2 + X_2X_3 - X_2X_3 - X_3X_2 = 0. \end{aligned}$$

4. Define the linear map $\sigma : \mathfrak{su}(2) \to \mathfrak{su}(2)$ by

$$\sigma(X_j) = X_{j+1}, \quad j, j+1 \in \mathbb{Z}/(3).$$

Show that σ is a Lie algebra automorphism, as is $\tau = \sigma^2$.
Show that these Lie algebra automorphisms act on the universal enveloping algebra, and that each preserves $\Delta = X_1^2 + X_2^2 + X_3^2$.
Deduce from this and from Exercise 3 that $[X_j, \Delta] = 0$ for each j.

5. Recall the isomorphism $\tau : SU(2) \times SU(2) \overset{\approx}{\to} SO(4)$, from (4.1.40). Using results from §1.2 (Chapter 1), show that

$$\tau\Bigl(\begin{pmatrix} u(\theta) & \\ & u(\varphi) \end{pmatrix}\Bigr) = \begin{pmatrix} R(\theta-\varphi) & \\ & R(\theta+\varphi) \end{pmatrix},$$

where

$$u(\theta) = \begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{pmatrix} \in SU(2), \quad R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in SO(2).$$

Deduce from (4.1.68) that if $\gamma_{k\ell} : SO(4) \to \mathcal{L}(\mathcal{P}_k \otimes \mathcal{P}_\ell)$ is as in (4.1.43), with $k+\ell$ even, then

$$\begin{aligned} &\operatorname{Tr} \gamma_{k\ell}\Bigl(\begin{pmatrix} R(\theta) & \\ & R(\varphi) \end{pmatrix}\Bigr) \\ &= \operatorname{Tr} D_{k/2}\Bigl(u\Bigl(\frac{\varphi+\theta}{2}\Bigr)\Bigr) \otimes D_{\ell/2}\Bigl(u\Bigl(\frac{\varphi-\theta}{2}\Bigr)\Bigr) \\ &= \frac{\sin{(k+1)(\varphi+\theta)/2}}{\sin(\varphi+\theta)/2} \cdot \frac{\sin{(\ell+1)(\varphi-\theta)/2}}{\sin(\varphi-\theta)/2}. \end{aligned}$$

4.2. Representations of U(n), I: roots and weights

Here we begin to take a detailed look at U(n) and its representations. Recall that the Lie algebra of U(n) is

$$\mathfrak{u}(n) = \{X \in \mathrm{M}(n, \mathbb{C}) : X^* = -X\}. \tag{4.2.1}$$

The complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$ of $\mathfrak{g} = \mathfrak{u}(n)$ is just $\mathrm{M}(n, \mathbb{C})$, which is the Lie algebra of $\mathrm{Gl}(n, \mathbb{C})$, which in turn can be regarded as the complexification of U(n). We can write

$$\mathrm{M}(n, \mathbb{C}) = \mathbb{C}\mathfrak{h} \oplus \mathfrak{n}_+ \oplus \mathfrak{n}_-. \tag{4.2.2}$$

Here

$$\mathfrak{h} = \{\mathrm{diag}(ia_1, \dots, ia_n) : a_j \in \mathbb{R}\} \tag{4.2.3}$$

is the Lie algebra of

$$\mathbb{T} = \{\mathrm{diag}(e^{ia_1}, \dots, e^{ia_n}) : a_j \in \mathbb{R}\} \subset \mathrm{U}(n). \tag{4.2.4}$$

In addition $\mathfrak{n}_+$ consists of strictly upper triangular matrices and $\mathfrak{n}_-$ of strictly lower triangular matrices, in $\mathrm{M}(n, \mathbb{C})$. It is clear that each of the three factors on the right side of (4.2.2) is a Lie algebra. The Lie algebra $\mathbb{C}\mathfrak{h}$ generates

$$D = \{\mathrm{diag}(c_1, \dots, c_n) : c_j \in \mathbb{C} \setminus 0\} \subset \mathrm{Gl}(n, \mathbb{C}), \tag{4.2.5}$$

while $\mathfrak{n}_+$ generates N_+, the group of upper triangular matrices in $\mathrm{Gl}(n, \mathbb{C})$ with ones on the diagonal, and $\mathfrak{n}_-$ generates N_-, the group of lower triangular matrices in $\mathrm{Gl}(n, \mathbb{C})$ with ones on the diagonal. There is the Gauss decomposition:

$$N_- D N_+ = G_{\mathrm{reg}} \quad \text{is dense in} \quad \mathrm{Gl}(n, \mathbb{C}), \tag{4.2.6}$$

or (in a weaker form)

$$G_{\mathrm{reg}} \text{ contains a neighborhood of the identity.} \tag{4.2.7}$$

The latter result follows fairly easily from the spanning property (4.2.2).

Convenient bases for the factors in (4.2.2) are provided by the matrices e_{jk}. Here we define e_{jk} to be the $n \times n$ matrix with a 1 in row j, column k, and zeros elsewhere. Equivalently, let $u_1, \dots, u_n$ denote the standard basis of $\mathbb{C}^n$. Then

$$e_{jk} u_\ell = \delta_{k\ell}\, u_j. \tag{4.2.8}$$

Then $\{e_{jk} : j < k\}$ spans $\mathfrak{n}_+$, $\{e_{jk} : j > k\}$ spans $\mathfrak{n}_-$, and, with

$$e_j = i e_{jj}, \tag{4.2.9}$$

the set $\{e_j : 1 \le j \le n\}$ spans $\mathfrak{h}$.

Suppose now that π is a unitary representation of U(n) on V, assumed to be finite dimensional. Since $\mathbb{T}$ is commutative, we can simultaneously diagonalize $\{\pi(h) : h \in \mathbb{T}\}$. Equivalently, we can simultaneously diagonalize $\{d\pi(X) : X \in \mathfrak{h}\}$. In other words,

$$V = \bigoplus_{\lambda \in \mathfrak{h}'} V_\lambda, \tag{4.2.10}$$

where, for $\lambda \in \mathfrak{h}'$,

$$V_\lambda = \{v \in V : d\pi(X)v = i\lambda(X)v, \ \forall \ X \in \mathfrak{h}\}. \tag{4.2.11}$$

If $\lambda \in \mathfrak{h}'$ and $V_\lambda \neq 0$, we call λ a *weight* for π, and a nonzero $v \in V_\lambda$ is called a *weight vector*. Note that the spaces V_λ in (4.2.10) are mutually orthogonal.

Let us apply this notion to the adjoint representation of $\mathrm{U}(n)$ on $\mathfrak{u}(n)_{\mathbb{C}} = \mathrm{M}(n, \mathbb{C})$. It is convenient to use the basis e_{jk} defined in (4.2.8). A computation gives $e_{ij}e_{k\ell} = \delta_{jk}e_{i\ell}$, hence

$$[e_{ij}, e_{k\ell}] = \delta_{jk}e_{i\ell} - \delta_{i\ell}e_{kj}. \tag{4.2.12}$$

In particular, $[e_{jj}, e_{k\ell}] = \delta_{jk}e_{j\ell} - \delta_{j\ell}e_{kj}$, and hence

$$X = \sum_j x_j e_j \in \mathfrak{h} \Longrightarrow [X, e_{jk}] = i(x_j - x_k)e_{jk}. \tag{4.2.13}$$

In other words, if we define

$$\omega_{jk} \in \mathfrak{h}', \quad \omega_{jk}(X) = x_j - x_k, \tag{4.2.14}$$

then ω_{jk} is a weight for the adjoint representation, with weight vector e_{jk}. We call ω_{jk} a *root*, and e_{jk} a *root vector*. Note the parallel between (4.2.13) and the commutator relation $[X_1, X_\pm] = \pm iX_\pm$, from (4.1.12).

Let us return to a general unitary representation π of $U(n)$ on V. The following can be compared with Lemma 4.1.1.

Proposition 4.2.1. *Set $E_{jk} = d\pi(e_{jk})$. Then*

$$E_{jk} : V_\lambda \longrightarrow V_{\lambda+\omega_{jk}}. \tag{4.2.15}$$

Thus if $\lambda \in \mathfrak{h}'$ is a weight for the representation π, then either E_{jk} annihilates V_λ or $\lambda + \omega_{jk}$ is a weight for π.

Proof. The commutation relation (4.2.13), which can be rewritten as

$$[X, e_{jk}] = i\omega_{jk}(X)e_{jk}, \quad X \in \mathfrak{h}, \tag{4.2.16}$$

leads to the identity

$$d\pi(X)E_{jk} = E_{jk}\big(d\pi(X) + i\omega_{jk}(X)I\big), \quad X \in \mathfrak{h}, \tag{4.2.17}$$

which implies (4.2.15). $\square$

Let us define an order on $\mathfrak{h}'$ as follows. Use the basis $\{e_j : 1 \le j \le n\}$ to make $\mathfrak{h} \approx \mathbb{R}^n$; then $\mathfrak{h}' \approx \mathbb{R}^n$. Given $\alpha, \beta \in \mathbb{R}^n$, we say $\alpha < \beta$ if the first nonzero entry of $\beta - \alpha$ is positive. With respect to this order, we have

$$\begin{aligned} \lambda + \omega_{jk} &> \lambda \ \text{ if } \ j < k, \\ \lambda + \omega_{jk} &< \lambda \ \text{ if } \ j > k. \end{aligned} \tag{4.2.18}$$

Hence we call E_{jk} a *raising operator* if $j < k$ (so $e_{jk} \in \mathfrak{n}_+$) and a *lowering operator* if $j > k$ (in which case $e_{jk} \in \mathfrak{n}_-$).

In (4.2.10) only finitely many weights appear. Thus there is a highest weight λ_m and a lowest weight λ_s. All the raising operators annihilate V_{λ_m} and all the lowering operators annihilate V_{λ_s}. Nonzero elements of V_{λ_m} are called *highest weight vectors.*

In view of this discussion, we have the following criterion for irreducibility. A converse will be established below.

Proposition 4.2.2. *Let π be a unitary representation of* $\mathrm{U}(n)$ *on V, finite dimensional. Consider the set $\mathcal{A}(\pi)$ of weight vectors annihilated by all raising operators. If $\mathcal{A}(\pi) \cup \{0\}$ is a linear space of dimension* 1*, then π is irreducible.*

Proof. Suppose $V = V_1 \oplus V_2$ with V_j invariant. We see from the previous paragraph that both V_1 and V_2 contain a nonzero element of $\mathcal{A}(\pi)$. □

Remark. A vector $v \in V$ is annihilated by all raising operators if and only if

$$d\pi(X_+)v = 0, \quad \forall\, X_+ \in \mathfrak{n}_+. \tag{4.2.19}$$

Let us note the following. Set

$$\mathcal{H}(\pi) = \bigcap_{j<k} \operatorname{Ker} E_{jk}. \tag{4.2.20}$$

From (4.2.17) it follows that

$$X \in \mathfrak{h} \Longrightarrow d\pi(X) : \mathcal{H}(\pi) \to \mathcal{H}(\pi), \tag{4.2.21}$$

and of course $\{d\pi(X)|_{\mathcal{H}(\pi)} : X \in \mathfrak{h}\}$ forms a commuting family of skew-adjoint operators, so they are simultaneously diagonalizable on $\mathcal{H}(\pi)$, i.e., $\mathcal{H}(\pi)$ is spanned by weight vectors. Thus the hypothesis that $\mathcal{A}(\pi) \cup \{0\}$ is a linear space of dimension 1 is equivalent to the hypothesis that $\dim \mathcal{H}(\pi) = 1$.

We next bring in the notion of contragredient representations. If π is a representation of a Lie group G on a finite dimensional space V, we define its contragredient representation $\overline{\pi}$ on V' by

$$\langle v, \overline{\pi}(g)w\rangle = \langle \pi(g^{-1})v, w\rangle, \quad v \in V,\ w \in V', \tag{4.2.22}$$

as in (2.3.12). Suppose π is unitary and V is given an orthonormal basis, so $\pi(g)$ is given by a unitary matrix $(\pi_{jk}(g))$. Then the matrix entries of $\overline{\pi}(g)$, with respect to the dual basis of V', are just the complex conjugates of those of π. If π is irreducible, so is $\overline{\pi}$.

Now assume π is an irreducible representation of $\mathrm{U}(n)$ on V, with contragredient representation $\overline{\pi}$ on V'. Let $\xi_0 \in \mathcal{A}(\pi) \subset V$ (i.e., ξ_0 is a weight vector annihilated by all raising operators) and let $\eta_0 \in \mathcal{A}^b(\overline{\pi})$ (i.e., η_0 is a weight vector for $\overline{\pi}$ annihilated by all *lowering* operators). Assume ξ_0 and η_0 are nonvanishing. Say ξ_0 has weight $\lambda \in \mathfrak{h}'$ and η_0 has weight $-\mu \in \mathfrak{h}'$. We form

$$\psi(X) = \langle d\pi(X)\xi_0, \eta_0\rangle = -\langle \xi_0, d\overline{\pi}(X)\eta_0\rangle, \quad X \in \mathrm{M}(n, \mathbb{C}). \tag{4.2.23}$$

Note that

$$\begin{gathered} X_+ \in \mathfrak{n}_+ \Longrightarrow \psi(X_+) = 0, \\ X_- \in \mathfrak{n}_- \Longrightarrow \psi(X_-) = 0, \\ H \in \mathfrak{h} \Longrightarrow \psi(H) = i\lambda(H)\langle \xi_0, \eta_0\rangle = i\mu(H)\langle \xi_0, \eta_0\rangle. \end{gathered} \tag{4.2.24}$$

We aim to show that $\langle \xi_0, \eta_0\rangle \neq 0$, which will imply that $\lambda = \mu$. First, it is convenient to bring in the following group level analogue of (4.2.23). Thus, with π a representation of $\mathrm{U}(n)$ on V, and with nonzero $\xi_0 \in \mathcal{A}(\pi)$, $\eta_0 \in \mathcal{A}^b(\overline{\pi})$ as before, set

$$\alpha(g) = \langle \pi(g)\xi_0, \eta_0\rangle = \langle \xi_0, \overline{\pi}(g^{-1})\eta_0\rangle. \tag{4.2.25}$$

As we will show in §4.5, a finite-dimensional representation π of $\mathrm{U}(n)$ always extends to a holomorphic representation of $\mathrm{Gl}(n, \mathbb{C})$. (Another proof is given in §4.6.) Hence $\alpha(g)$ is well defined for $g \in \mathrm{Gl}(n, \mathbb{C})$ and is holomorphic in g. At this point it is useful to note that $d\pi(X_+)\xi_0 = 0$ for all $X_+ \in \mathfrak{n}_+$ (cf. (4.2.19)), and

$$d\pi(X_+)\xi_0 = 0 \Longrightarrow \pi(e^{tX_+})\xi_0 = e^{td\pi(X_+)}\xi_0 = \xi_0, \quad \forall\, t.$$

Hence, since $\mathrm{Exp} : \mathfrak{n}_+ \to N_+$ has range containing a neighborhood of $e \in N_+$,

$$\pi(\zeta_+)\xi_0 = \xi_0, \quad \forall\, \zeta_+ \in N_+.$$

Similarly, $\overline{\pi}(\zeta_-^{-1})\eta_0 = \eta_0$ for all $\zeta_- \in N_-$. Hence, parallel to (4.2.24), we have, for all $g \in \mathrm{Gl}(n, \mathbb{C})$,

$$\begin{aligned} \alpha(g\zeta_+) &= \alpha(g), \quad \zeta_+ \in N_+, \\ \alpha(\zeta_- g) &= \alpha(g), \quad \zeta_- \in N_-, \end{aligned} \tag{4.2.26}$$

Also

$$\begin{aligned}\alpha(g\delta) &= e^{i\lambda(H)}\alpha(g), \quad \delta = e^H \in \mathbb{T}, \\ \alpha(\delta g) &= e^{i\mu(H)}\alpha(g),\end{aligned} \tag{4.2.27}$$

since $\pi(e^H)\xi_0 = e^{i\lambda(H)}\xi_0$ and $\overline{\pi}(e^{-H})\eta_0 = e^{i\mu(H)}\eta_0$. More generally,

$$\begin{aligned}\alpha(\delta g) &= e^{i(\lambda(H_1)+i\lambda(H_2))}\alpha(g), \quad \delta = e^{H_1+iH_2} \in D, \\ \alpha(g\delta) &= e^{i(\mu(H_1)+i\mu(H_2))}\alpha(g).\end{aligned} \tag{4.2.28}$$

We have from (4.2.26)–(4.2.28) that

$$\begin{aligned}&\zeta_\pm \in N_\pm,\ \delta = e^{H_1+iH_2} \in D \\ &\Longrightarrow \alpha(\zeta_-\delta\zeta_+) = \alpha(\delta) = e^{i(\lambda(H_1)+i\lambda(H_2))}\alpha(e).\end{aligned} \tag{4.2.29}$$

We are now prepared to prove:

Lemma 4.2.3. *Given that π is irreducible,*

$$\langle \xi_0, \eta_0\rangle \neq 0. \tag{4.2.30}$$

Hence $\lambda = \mu$.

Proof. We have $\langle \xi_0, \eta_0\rangle = \alpha(e)$. By (4.2.29), (4.2.7) and holomorphy, if $\alpha(e) = 0$ then $\alpha(g) \equiv 0$. Consider

$$V_0 = \{\xi \in V : \langle \pi(g)\xi, \eta_0\rangle = 0,\ \forall\, g \in \mathrm{Gl}(n, \mathbb{C})\}. \tag{4.2.31}$$

Then V_0 is an invariant linear subspace of V and $\xi_0 \in V_0$, so $V_0 \neq 0$. Irreducibility forces $V_0 = V$, but this is clearly false, since $\eta_0 \neq 0$, and the contradiction forces (4.2.30) to hold. $\square$

Having $\lambda = \mu$, we can rewrite (4.2.28) as

$$\alpha(g\delta) = \alpha(\delta g) = e^{i(\lambda(H_1)+i\lambda(H_2))}\alpha(g), \quad \delta = e^{H_1+iH_2} \in D. \tag{4.2.32}$$

We next prove:

Proposition 4.2.4. *If π is an irreducible representation of* U(n) *on V, then $\mathcal{H}(\pi)$, given in* (4.2.20)*, is a one-dimensional linear space. Hence the highest weight vector for π is unique, up to a constant multiple.*

Proof. Suppose $\xi_1 \in \mathcal{H}(\pi)$ is a weight vector. The argument above also shows $\langle \xi_1, \eta_0\rangle \neq 0$. Normalize so $\langle \xi_1, \eta_0\rangle = \langle \xi_0, \eta_0\rangle$. Then computations parallel to (4.2.23)–(4.2.29) give $\langle \pi(g)\xi_1, \eta_0\rangle \equiv \alpha(g)$, so

$$\langle \pi(g)(\xi_1 - \xi_0), \eta_0\rangle = 0, \quad \forall\, g, \tag{4.2.33}$$

hence

$$W = \mathrm{Span}\{\pi(g)(\xi_1 - \xi_0) : g \in U(n)\} \perp \eta_0. \tag{4.2.34}$$

Since $\pi(g) : W \to W$ and π is irreducible, this implies $\xi_1 = \xi_0$. $\square$

We next show that inequivalent irreducible representations of $\mathrm{U}(n)$ have distinct highest weights.

Proposition 4.2.5. *If π and π' are irreducible representations of* $\mathrm{U}(n)$ *with the same highest weight, then $\pi \approx \pi'$.*

Proof. Suppose π' also has highest weight λ. Pick $\xi_0' \in \mathcal{A}(\pi'),\ \eta_0' \in \mathcal{A}^b(\overline{\pi}')$ and arrange that $\langle \xi_0, \eta_0\rangle = \langle \xi_0', \eta_0'\rangle$. Consider

$$\beta(g) = \langle \pi'(g)\xi_0', \eta_0'\rangle. \tag{4.2.35}$$

We have $\beta(e) = \alpha(e) \neq 0$, and results parallel to (4.2.26)–(4.2.29) for β imply

$$\beta(\zeta_-\delta\zeta_+) = \alpha(\zeta_-\delta\zeta_+), \quad \forall\ \zeta_- \in N_-,\ \delta \in D,\ \zeta_+ \in N_+. \tag{4.2.36}$$

As both α and β are holomorphic on $\mathrm{Gl}(n,\mathbb{C})$ and N_-DN_+ contains a neighborhood of $e \in \mathrm{Gl}(n,\mathbb{C})$, it follows that $\alpha \equiv \beta$ on $\mathrm{Gl}(n,\mathbb{C})$ and a fortiori $\alpha \equiv \beta$ on $\mathrm{U}(n)$. But if π and π' are not equivalent the Weyl orthogonality relations imply $\alpha \perp \beta$ in $L^2(\mathrm{U}(n))$, so the proposition is proven □

It remains to characterize which elements $\lambda \in \mathfrak{h}'$ are highest weights of irreducible representations of $\mathrm{U}(n)$. We take this up in §4.4.

Exercises

1. A matrix $A \in Gl(n, \mathbb{C})$ is said to have an LU factorization if it can be written $A = LU$, with L lower triangular and U upper triangular. Show that this exists if and only if $A \in G_{\text{reg}}$, as defined in (4.2.6).

2. The following result is Proposition 1.6.7 of [**42**].

Proposition. Take $A \in M(n, \mathbb{C})$, and for $\ell \in \{1, \dots, n\}$, let $A^{(\ell)}$ denote the $\ell \times \ell$ matrix forming the upper left corner of A, i.e.,

$$A^{(1)} = (a_{11}), \quad A^{(2)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \dots, \ A^{(n)} = A.$$

Assume each $A^{(\ell)}$ is invertible, i.e., $\det A^{(\ell)} \neq 0$ for $1 \le \ell \le n$. Then A has an LU factorization.

Show that this proposition implies the denseness result stated in (4.2.6). See if you can prove this proposition (or consult §1.6 of [**42**]).

3. Show that the map $E : \mathbb{C}\mathfrak{h} \oplus \mathfrak{n}_+ \oplus \mathfrak{n}_- \to Gl(n, \mathbb{C})$, given by

$$E(H, N_+, N_-) = e^H e^{N_+} e^{N_-}$$

maps a neighborhood of 0 in $M(n, \mathbb{C})$ diffeomorphically onto a neighborhood of the identity I in $Gl(n, \mathbb{C})$, and use this to establish that G_{reg} contains a neighborhood of the identity in $Gl(n, \mathbb{C})$, as stated in (4.2.7).
Hint. Show that $DE(0, 0, 0) = I$.

4.3. Representations of U(n), II: some basic examples

Here we consider some basic examples of representations of U(n). First, define representations S^ℓ and $\overline{S}^\ell$ of U(n) on

$$\mathcal{P}_\ell = \text{ space of polynomials on } \mathbb{C}^n \text{ homogeneous of degree } \ell, \tag{4.3.1}$$

by

$$S^\ell(g)f(z) = f(g^t z), \quad \overline{S}^\ell(g)f(z) = f(g^{-1}z). \tag{4.3.2}$$

Note that (4.3.2) extends to $g \in \mathrm{Gl}(n, \mathbb{C})$, and we have

$$d\overline{S}^\ell(X)f(z) = \frac{d}{dt}f(e^{-tX}z)\big|_{t=0} = \frac{d}{dt}f(z-tXz)\big|_{t=0}, \quad X \in \mathrm{M}(n, \mathbb{C}). \tag{4.3.3}$$

Hence

$$\begin{aligned} d\overline{S}^\ell(e_{jk})p(z) &= \frac{d}{dt}p(z_1, \dots, z_j - tz_k, \dots, z_n)\big|_{t=0} \\ &= -z_k \frac{\partial p}{\partial z_j}, \end{aligned} \tag{4.3.4}$$

and in particular

$$d\overline{S}^\ell(e_j)p(z) = -iz_j \frac{\partial p}{\partial z_j}. \tag{4.3.5}$$

We see that, for $z^\alpha = z_1^{\alpha_1} \cdots z_n^{\alpha_n}$ with $|\alpha| = \alpha_1 + \cdots + \alpha_n = \ell$,

$$d\overline{S}^\ell(e_j)z^\alpha = -i\alpha_j z^\alpha. \tag{4.3.6}$$

Thus z^α is a weight vector for $\overline{S}^\ell$, with weight $-\alpha$. The highest weight is $(0, \dots, 0, -\ell)$, with weight vector z_n^ℓ. It is clear from (4.3.4) that the only weight vector annihilated by all raising operators is z_n^ℓ. Hence $\overline{S}^\ell$ is irreducible.

Note that

$$dS^\ell(X)f(z) = \frac{d}{dt}f(e^{tX^t}z)\big|_{t=0} = -d\overline{S}^\ell(X^t). \tag{4.3.7}$$

Hence

$$dS^\ell(e_{jk})p(z) = -d\overline{S}^\ell(e_{kj})p(z) = z_j \frac{\partial p}{\partial z_k}, \quad dS^\ell(e_j)p(z) = iz_j \frac{\partial p}{\partial z_j}. \tag{4.3.8}$$

In particular z^α is a weight vector for S^ℓ, with weight α. The highest weight is $(\ell, 0, \dots, 0)$, with weight vector z_1^ℓ. This is the only weight vector annihilated by all raising operators, so S^ℓ is also irreducible.

Next, we define representations Λ^ℓ of U(n) on $\Lambda^\ell\mathbb{C}^n$ $(0 \le \ell \le n)$ by

$$\Lambda^\ell(g)\, v_1 \wedge \cdots \wedge v_\ell = gv_1 \wedge \cdots \wedge gv_\ell. \tag{4.3.9}$$

This is also well defined for $g \in \mathrm{Gl}(n,\mathbb{C})$, and we have, for $X \in \mathrm{M}(n,\mathbb{C})$,

$$\begin{aligned} & d\Lambda^\ell(X)\, v_1 \wedge \cdots \wedge v_\ell \\ & = \frac{d}{dt}\, e^{tX} v_1 \wedge \cdots \wedge e^{tX} v_\ell \big|_{t=0} \\ & = Xv_1 \wedge v_2 \wedge \cdots \wedge v_\ell + \cdots + v_1 \wedge \cdots \wedge v_{\ell-1} \wedge Xv_\ell. \end{aligned} \tag{4.3.10}$$

In this case, with $u_1, \dots, u_n$ as before denoting the standard basis of $\mathbb{C}^n$, if we set

$$u_J = u_{j_1} \wedge \cdots \wedge u_{j_\ell}, \quad J = (j_1, \dots, j_\ell), \tag{4.3.11}$$

with $j_1 < \cdots < j_\ell$, then

$$\begin{aligned} E_{jk} u_J = u_{j_1} \wedge \cdots \wedge u_{j_{\nu-1}} \wedge\, u_j \wedge u_{j_{\nu+1}} \wedge \cdots \wedge u_{j_\ell}, & \quad \text{if } k = j_\nu, \\ 0, & \quad \text{if } k \notin \{j_1, \dots, j_\ell\}, \end{aligned} \tag{4.3.12}$$

and

$$\begin{aligned} d\Lambda^\ell(e_j) u_J = i u_J & \quad \text{if} \quad j \in \{j_1, \dots, j_\ell\}, \\ 0 & \quad \text{if} \quad j \notin \{j_1, \dots, j_\ell\}. \end{aligned} \tag{4.3.13}$$

Thus u_J is a weight vector for Λ^ℓ, of weight $\gamma(J)$, where $\gamma(J)_j = 1$ if $j \in \{j_1, \dots, j_\ell\}$, 0 otherwise. Also from (4.3.12) it follows that the only weight vector annihilated by all raising operators is $u_1 \wedge \cdots \wedge u_\ell$. Hence Λ^ℓ is irreducible, with highest weight $(1, \dots, 1, 0, \dots, 0)$ (with ℓ ones).

We record the dimensions of the representation spaces described above. A look at a standard basis shows that

$$\dim \Lambda^\ell \mathbb{C}^n = \binom{n}{\ell}. \tag{4.3.14}$$

As for $\mathcal{P}_\ell \approx S^\ell \mathbb{C}^n$, we have

$$\dim S^\ell \mathbb{C}^n = \#\{\beta \ge 0 : z^\beta = z_1^{\beta_1} \cdots z_n^{\beta_n},\ |\beta| = \ell\}. \tag{4.3.15}$$

If we set $\vartheta_n(\ell) = \dim S^\ell \mathbb{C}^n$, we can see that

$$\vartheta_{n+1}(\ell) = \vartheta_n(\ell) + \vartheta_n(\ell - 1) + \cdots + \vartheta_n(0). \tag{4.3.16}$$

It is shown in §8.A that

$$\dim S^\ell \mathbb{C}^{n+1} = \binom{n+\ell}{n}. \tag{4.3.17}$$

We next reconsider the adjoint representation of $\mathrm{U}(n)$ on $\mathrm{M}(n,\mathbb{C})$, given by

$$\mathrm{Ad}(g)X = gXg^{-1}, \tag{4.3.18}$$

and the derived representation ad of $\mathfrak{u}(n)$ on $\mathrm{M}(n,\mathbb{C})$, and its extension to the representation $\mathbb{C}\mathfrak{u}(n) = \mathrm{M}(n,\mathbb{C})$ on $\mathrm{M}(n,\mathbb{C})$, given by

$$\mathrm{ad}\, X(Y) = [X, Y]. \tag{4.3.19}$$

These representations are not irreducible. We have a decomposition into invariant subspaces

$$M(n,\mathbb{C}) = \{cI\} \oplus \mathrm{M}_0(n,\mathbb{C}),\\ \mathrm{M}_0(n,\mathbb{C}) = \{X \in \mathrm{M}(n,\mathbb{C}) : \mathrm{Tr}\, X = 0\}. \tag{4.3.20}$$

Ad acts trivially on $\{cI\}$. We claim it acts irreducibly on $\mathrm{M}_0(n,\mathbb{C})$. The analysis below will establish this.

Using (4.2.12)–(4.2.14), we have the weight space (aka root space) decomposition

$$\begin{aligned} \mathrm{M}(n,\mathbb{C}) &= \mathbb{C}\mathfrak{h} \oplus \bigoplus_{j\neq k} \mathrm{Span}(e_{jk}) \\ &= \mathfrak{g}_0 \oplus \bigoplus_{j\neq k} \mathfrak{g}_{\omega_{jk}}, \end{aligned} \tag{4.3.21}$$

where, for $X = \sum x_j e_j$, $\omega_{jk}(X) = x_j - x_k$. Recall from (4.2.15) that $E_{\ell m} = \mathrm{ad}\, e_{\ell m}$ satisfies

$$E_{\ell m} : \mathfrak{g}_{\omega_{jk}} \longrightarrow \mathfrak{g}_{\omega_{jk}+\omega_{\ell m}}. \tag{4.3.22}$$

Now $\omega_{jk}(X)+\omega_{\ell m}(X) = x_j - x_k + x_\ell - x_m$, so, given that $\ell < m$ and $j \neq k$,

$$\omega_{jk} + \omega_{\ell m} \ \text{ is a root } \iff k = \ell \ \text{ or } \ j = m. \tag{4.3.23}$$

Furthermore,

$$\begin{aligned} E_{\ell m} e_{jk} = [e_{\ell m}, e_{jk}] &= \delta_{mj} e_{\ell k} - \delta_{\ell k} e_{jm} \\ &= 0 \ \text{ provided } \ m \neq j \text{ and } \ell \neq k, \end{aligned} \tag{4.3.24}$$

and also

$$\begin{aligned} m = j \Rightarrow E_{\ell m} e_{jk} = e_{\ell k} - \delta_{\ell k} e_{mm} &= e_{\ell k} \text{ if } \ell \neq k \\ &\qquad e_{\ell\ell} - e_{mm} \text{ if } \ell = k, \end{aligned} \tag{4.3.25}$$

and

$$\begin{aligned} \ell = k \Rightarrow E_{\ell m} e_{jk} = \delta_{mj} e_{\ell\ell} - e_{jm} &= \ -e_{jm} \text{ if } m \neq j, \\ &\qquad e_{\ell\ell} - e_{mm} \text{ if } m = j. \end{aligned} \tag{4.3.26}$$

In conclusion, we deduce that

$$\begin{aligned} E_{\ell m} e_{jk} = 0,\ \forall\, \ell < m &\iff m \neq j \text{ and } \ \ell \neq k,\ \forall\, \ell < m \\ &\iff j = 1 \ \text{ and } \ k = n. \end{aligned} \tag{4.3.27}$$

Hence the subspace of $\oplus_{j\neq k}\mathfrak{g}_{\omega_{jk}}$ annihilated by all raising operators is $\mathfrak{g}_{\omega_{1n}} = \mathrm{Span}(e_{1n})$, with weight $\omega_{1n} = (1, 0, \dots, 0, -1)$.

It remains to investigate which elements of $\mathfrak{g}_0 = \mathbb{C}\mathfrak{h}$ are annihilated by all raising operators. In fact, by (4.2.13), for $X \in \mathfrak{h}$, $-[e_{\ell m}, X] = [X, e_{\ell m}] = i\omega_{\ell m}(X) e_{\ell m}$, hence

$$E_{\ell m}\Big(\sum x_j e_j\Big) = -i(x_\ell - x_m) e_{\ell m}, \tag{4.3.28}$$

which is 0 for all $\ell < m$ if and only if $x_1 = \cdots = x_n$.

It follows from these arguments that $\mathcal{H}(\mathrm{Ad})$ is spanned by e_{1n} and $e_{11} + \cdots + e_{nn}$. These are weight vectors with weights

$$(1, 0, \dots, 0, -1) \quad \text{and} \quad (0, \dots, 0). \tag{4.3.29}$$

This establishes the irreducibility of Ad on each of the two factors in (4.3.20).

The irreducibility of the representation Ad of $\mathrm{U}(n)$ on $\mathrm{M}_0(n, \mathbb{C})$ is equivalent to the irreducibility of ad, representing $\mathrm{M}(n, \mathbb{C})$ on $\mathrm{M}_0(n, \mathbb{C})$, In turn, since $\{cI\}$ is the center of $\mathrm{M}(n, \mathbb{C})$, this is equivalent to the irreducibility of ad, representing $\mathrm{M}_0(n, \mathbb{C})$ on $\mathrm{M}_0(n, \mathbb{C})$.

Generally, if $\mathfrak{g}$ is a Lie algebra, the representation ad of $\mathfrak{g}$ on $\mathfrak{g}$ has an invariant linear subspace $\mathfrak{h} \subset \mathfrak{g}$ if and only if

$$X \in \mathfrak{g}, \ Y \in \mathfrak{h} \Longrightarrow [X, Y] \in \mathfrak{h}, \tag{4.3.30}$$

i.e., if and only if $\mathfrak{h}$ is an *ideal* of $\mathfrak{g}$. If $\mathfrak{g}$ has no proper ideals, we say $\mathfrak{g}$ is a *simple* Lie algebra. Hence the content of the irreducibility of the action of $\mathrm{U}(n)$ on $\mathrm{M}_0(n, \mathbb{C})$ derived above is that

$$\mathrm{M}_0(n, \mathbb{C}) \ \text{ is a simple complex Lie algebra.} \tag{4.3.31}$$

Exercises

1. Consider the following alternative analyses of the weight vectors for the representations $\overline{S}^{\ell}$, S^{ℓ}, Λ^{ℓ}, and Ad of $U(n)$.

(a) $\overline{S}^{\ell}(\mathrm{diag}(e^{i\theta_1},\dots,e^{i\theta_n}))z^{\alpha} = e^{-i\alpha\cdot\theta}z^{\alpha}$.

So z^{α} is a weight vector for $\overline{S}^{\ell}$, with weight $-\alpha$.

Highest weight $(0,\dots,0,-\ell)$.

(b) $S^{\ell}(\mathrm{diag}(e^{i\theta_1},\dots,e^{i\theta_n}))z^{\alpha} = e^{i\alpha\cdot\theta}z^{\alpha}$.

So z^{α} is a weight vector for S^{ℓ}, with weight α.

Highest weight $(\ell,0,\dots,0)$.

(c) $\Lambda^{\ell}(\mathrm{diag}(e^{i\theta_1},\dots,e^{i\theta_n}))u_J = e^{i\gamma(J)\cdot\theta}u_J$, with

$$\begin{aligned}\gamma(J)_j = 1, &\quad \text{if } \ j \in \{j_1,\dots,j_\ell\} \quad (J = (j_1,\dots,j_\ell)),\\ 0, &\quad \text{if not.}\end{aligned}$$

So $u_J = u_{j_1}\wedge\cdots\wedge u_{j_\ell}$ is a weight vector for Λ^{ℓ}, with weight $\gamma(J)$.

Highest weight $(1,\dots,1,0,\dots,0)$ (ℓ ones).

(d) $\mathrm{Ad}(\mathrm{diag}(e^{i\theta_1},\dots,e^{i\theta_n}))e_{jk} = e^{i(\theta_j-\theta_k)}e_{jk}$.

So e_{jk} is a weight vector for Ad, with weight $\omega_{jk}(\theta) = \theta_j - \theta_k$.

Highest weight $(1,0,\dots,0,-1)$.

4.4. Representations of U(n), III: identification of highest weights

In this section we characterize which elements of $\mathfrak{h}'$ are highest weights of irreducible representations of U(n) and hence parametrize the set of such representations. As in §4.2, we use the basis $\{e_j : 1 \le j \le n\}$ of $\mathfrak{h}$ and the dual basis of $\mathfrak{h}'$ to identify these spaces with $\mathbb{R}^n$, so $\lambda \in \mathfrak{h}'$ is written as $\lambda = (\lambda_1, \dots, \lambda_n)$. Here is our main result.

Theorem 4.4.1. *The elements of $\mathfrak{h}'$ that are highest weights of an irreducible representation of* U(n) *are precisely given by*

$$\{(k_1, \dots, k_n) : k_\nu \in \mathbb{Z},\ k_1 \ge \cdots \ge k_n\}. \tag{4.4.1}$$

Hence the set of equivalence classes of irreducible unitary representations of U(n) *is in natural one-to-one correspondence with the set* (4.4.1).

First we show that if $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathfrak{h}'$ is a highest weight, then it must have the form (4.4.1). Since $i\lambda : \mathfrak{h} \to i\mathbb{R}$ must exponentiate to a homomorphism $\mathbb{T} \to S^1 \subset \mathbb{C}$, we must have $\lambda = (k_1, \dots, k_n)$, $k_\nu \in \mathbb{Z}$. The fact that $k_1 \ge \cdots \ge k_n$ is a consequence of the following.

Lemma 4.4.2. *If $\lambda = (\lambda_1, \dots, \lambda_n)$ is a weight of a representation π of* U(n) *on V, so is $\lambda(\sigma) = (\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)})$, for each $\sigma \in S_n$.*

Proof. Let E_σ denote the permutation matrix, $E_\sigma u_k = u_{\sigma(k)}$, where $\{u_1, \dots, u_n\}$ is the standard basis of $\mathbb{C}^n$; thus $E_\sigma \in \mathrm{U}(n)$. It is readily verified that

$$E_\sigma^{-1} \operatorname{diag}(c_1, \dots, c_n)\, E_\sigma = \operatorname{diag}(c_{\sigma(1)}, \dots, c_{\sigma(n)}). \tag{4.4.2}$$

Now, given $\lambda = (\lambda_1, \dots, \lambda_n)$, the weight space V_λ has the characterization

$$v \in V_\lambda \Longleftrightarrow \pi\big(\operatorname{diag}(c_1, \dots, c_n)\big)v = (c_1^{\lambda_1} \cdots c_n^{\lambda_n})v. \tag{4.4.3}$$

It follows that

$$\pi(E_\sigma^{-1}) : V_\lambda \longrightarrow V_{\lambda(\sigma)} \tag{4.4.4}$$

is an isomorphism. $\square$

It remains to show that each element of the form (4.4.1) is the highest weight of an irreducible representation of U(n). First note that if $(k_1, \dots, k_n) \in \mathfrak{h}'$ is the highest weight of π, then, for each $j \in \mathbb{Z}$,

$${}^j\pi(g) = (\det g)^j\, \pi(g) \ \text{ has highest weight } (k_1 + j, \dots, k_n + j), \tag{4.4.5}$$

with the same weight vector as π, as follows readily from (4.4.3).

Thus it suffices to construct an irreducible representation of $U(n)$ with highest weight $(k_1, \dots, k_n)$ satisfying $k_\nu \in \mathbb{Z}$ and $k_1 \ge \cdots \ge k_n \ge 0$. In this case we can write

$$(k_1, \dots, k_n) = j_1\gamma_1 + \cdots + j_n\gamma_n, \quad j_\nu \in \mathbb{Z}^+, \tag{4.4.6}$$

where γ_ℓ is the highest weight of the representation Λ^ℓ of $\mathrm{U}(n)$ discussed in (4.3.9)–(4.3.13), i.e.,

$$\gamma_\ell = (1, \dots, 1, 0, \dots, 0) \quad \text{(with } \ell \text{ ones)}. \tag{4.4.7}$$

The following gives the key construction.

Proposition 4.4.3. *A weight of the type* (4.4.6) *occurs as the highest weight of an irreducible component of the representation*

$$(\Lambda^1)^{\otimes j_1} \otimes \cdots \otimes (\Lambda^n)^{\otimes j_n} \tag{4.4.8}$$

of $\mathrm{U}(n)$ *on* $(\Lambda^1\mathbb{C}^n)^{\otimes j_1} \otimes \cdots \otimes (\Lambda^n\mathbb{C}^n)^{\otimes j_n}$.

Here $V^{\otimes j}$ denotes the j-fold tensor product $V \otimes \cdots \otimes V$. More generally than Proposition 4.4.3, we have the following.

Proposition 4.4.4. *Suppose π_j is a unitary representation of $U(n)$ on V_j, with highest weight λ_j. Then the representation*

$$\pi_1 \otimes \cdots \otimes \pi_K \quad \textit{on} \quad V_1 \otimes \cdots \otimes V_K \tag{4.4.9}$$

has highest weight $\lambda_1 + \cdots + \lambda_K$.

Proof. Indeed, suppose we have weight space decompositions

$$V_j = \bigoplus_{\mu \in \mathcal{S}_j \subset \mathfrak{h}'} V_{j\mu} \tag{4.4.10}$$

for π_j. Then $V_1 \otimes \cdots \otimes V_K$ is spanned by

$$V_{1\mu_1} \otimes \cdots \otimes V_{K\mu_K}, \quad \mu_\nu \in \mathcal{S}_\nu, \tag{4.4.11}$$

and we claim this consists of weight vectors for $\pi_1 \otimes \cdots \otimes \pi_K$, of weight $\mu_1 + \cdots + \mu_K$. Indeed, if $v_{j\mu_j} \in V_{j\mu_j}$ and $\vartheta(\theta) = \mathrm{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$, then

$$\begin{aligned} &\pi_1 \otimes \cdots \otimes \pi_K(\vartheta(\theta))(v_{1\mu_1} \otimes \cdots \otimes v_{K\mu_K}) \\ &\quad = \pi_1(\vartheta(\theta))v_{1\mu_1} \otimes \cdots \otimes \pi_K(\vartheta(\theta))v_{K\mu_K} \\ &\quad = e^{i\mu_1(\theta)} \cdots e^{i\mu_K(\theta)}\, v_{1\mu_1} \otimes \cdots \otimes v_{K\mu_K}. \end{aligned} \tag{4.4.12}$$

□

To return to Proposition 4.4.3, we have that (4.4.6) is the highest weight of the representation (4.4.8). Now when a representation π of $\mathrm{U}(n)$ on V is decomposed into irreducible factors, the weights that occur in these factors are precisely the weights that occur in π, so an irreducible factor of (4.4.8) has the desired highest weight. This finishes the proof of Theorem 4.4.1.

We will denote by $\mathcal{D}_{(k_1,\dots,k_n)}$ an irreducible unitary representation of $\mathrm{U}(n)$ with highest weight $(k_1, \dots, k_n)$, satisfying (4.4.1). In particular, from §4.3 we have

$$S^\ell \approx \mathcal{D}_{(\ell,0,\dots,0)}, \quad \overline{S}^\ell \approx \mathcal{D}_{(0,\dots,0,-\ell)}, \tag{4.4.13}$$

and

$$\Lambda^\ell \approx \mathcal{D}_{(1,\dots,1,0,\dots,0)}, \quad \text{(with } \ell \text{ ones)}, \quad 0 \le \ell \le n. \tag{4.4.14}$$

It is useful to record explicitly the content of (4.4.5) in this notation:

$$\mathcal{D}_{(k_1+j,\dots,k_n+j)}(g) = (\det g)^j\, \mathcal{D}_{(k_1,\dots,k_n)}(g). \tag{4.4.15}$$

Also from §4.3 we have

$$\mathrm{Ad} \approx \mathcal{D}_{(1,0,\dots,0,-1)} \oplus \mathcal{D}_{(0,\dots,0)}. \tag{4.4.16}$$

One simple corollary of Theorem 4.4.1 and Lemma 4.4.2 is the following.

Proposition 4.4.5. *All the one-dimensional representations of* U(n) *are equivalent to the representations*

$$\theta_j(g) = (\det g)^j, \tag{4.4.17}$$

for some $j \in \mathbb{Z}$*, in turn equivalent to* $\mathcal{D}_{(j,\dots,j)}$.

Proof. A representation of U(n) on V when dim $V = 1$ has only one weight, say $\lambda = (k_1, \dots, k_n)$, with $k_1 \ge \cdots \ge k_n$. By Lemma 4.4.2, each $(k_{\sigma(1)}, \dots, k_{\sigma(n)})$ must also be a weight. This forces $k_1 = \cdots = k_n = j$ (say), which gives (4.4.17). $\square$

4.5. Connections between representations of U(n), SU(n), and Gl($n, \mathbb{C}$)

Here we compare finite-dimensional representations of the three groups U(n), SU(n), and Gl($n, \mathbb{C}$). We first show that any such representation of U(n) extends to Gl($n, \mathbb{C}$), as a holomorphic representation. (See §4.6 for another proof.) To this end, let π be a representation of U(n) on V, dim $V < \infty$. We have a Lie algebra representation

$$d\pi : \mathfrak{u}(n) \longrightarrow \mathrm{End}(V), \tag{4.5.1}$$

which extends to a Lie algebra representation

$$d\pi : \mathrm{M}(n, \mathbb{C}) \longrightarrow \mathrm{End}(V), \tag{4.5.2}$$

which is also $\mathbb{C}$-linear. By Corollary 3.6.4, this exponentiates to a representation

$$\pi : \widetilde{\mathrm{Gl}}(n, \mathbb{C}) \longrightarrow \mathrm{Gl}(V), \tag{4.5.3}$$

where $\widetilde{\mathrm{Gl}}$ is the universal cover of Gl($n, \mathbb{C}$). In order to obtain

$$\pi : \mathrm{Gl}(n, \mathbb{C}) \longrightarrow \mathrm{Gl}(V), \tag{4.5.4}$$

we need to show that π in (4.5.3) has the property

$$\pi(g) = I, \quad \forall\ g \in \mathrm{Ker}\,\beta, \tag{4.5.5}$$

where

$$\beta : \widetilde{\mathrm{Gl}}(n, \mathbb{C}) \longrightarrow \mathrm{Gl}(n, \mathbb{C}) \tag{4.5.6}$$

is the natural covering map. To see this let

$$\alpha : \widetilde{\mathrm{U}}(n) \longrightarrow \mathrm{U}(n) \tag{4.5.7}$$

denote the natural projection of the universal group $\widetilde{\mathrm{U}}(n)$ onto U(n). We have a commutative diagram

$$\begin{array}{ccc} \widetilde{\mathrm{U}}(n) & \longrightarrow & \widetilde{\mathrm{Gl}}(n, \mathbb{C}) \\ \alpha \downarrow & & \beta \downarrow \\ \mathrm{U}(n) & \longrightarrow & \mathrm{Gl}(n, \mathbb{C}) \end{array}$$

The following result is key:

Lemma 4.5.1. *We have*

$$\mathrm{Ker}\,\alpha = \mathrm{Ker}\,\beta. \tag{4.5.8}$$

Proof. Since $\operatorname{Ker}\alpha$ and $\operatorname{Ker}\beta$ are naturally isomorphic to the fundamental groups of $\mathrm{U}(n)$ and $\mathrm{Gl}(n,\mathbb{C})$, it suffices to note:

$$\mathrm{U}(n) \hookrightarrow \mathrm{Gl}(n,\mathbb{C}) \quad \text{is a deformation retract.} \tag{4.5.9}$$

As for this result, this follows by polar decomposition:

$$\mathrm{Gl}(n,\mathbb{C}) \approx \mathrm{U}(n) \times \mathcal{P}(n), \tag{4.5.10}$$

where $\mathcal{P}(n)$ denotes the set of positive-definite operators on $\mathbb{C}^n$. If $A \in \mathrm{Gl}(n,\mathbb{C})$, we have

$$A = UP, \quad U \in \mathrm{U}(n), \quad P \in \mathcal{P}(n), \tag{4.5.11}$$

uniquely defined by

$$P = (A^*A)^{1/2}, \quad U = AP^{-1}. \tag{4.5.12}$$

$\square$

Returning to (4.5.5), we see this is true when π represents $\mathrm{U}(n)$, since, on $\widetilde{\mathrm{U}}(n)$, we have $\pi(g) = I$ for $g \in \operatorname{Ker}\alpha$, and (4.5.8) holds. Thus (4.5.4) is established. Since

$$\mathrm{Exp} : \mathrm{M}(n,\mathbb{C}) \longrightarrow \mathrm{Gl}(n,\mathbb{C}) \tag{4.5.13}$$

is holomorphic, we also have that π in (4.5.3) is holomorphic on $\mathrm{Gl}(n,\mathbb{C})$.

Our next topic in this section is the comparison of the unitary irreducible representations of $\mathrm{U}(n)$ and $\mathrm{SU}(n)$. The key to this study comes from the exact sequence of groups

$$1 \longrightarrow K_n \longrightarrow S^1 \times \mathrm{SU}(n) \longrightarrow \mathrm{U}(n) \longrightarrow 1, \tag{4.5.14}$$

where $(\omega, g) \mapsto \omega g$ and

$$K_n = \{(\omega, g) \in S^1 \times \mathrm{SU}(n) : g = \omega^{-1} I,\ \omega^n = 1\}, \tag{4.5.15}$$

a cyclic group of order n, generated by

$$(\zeta^{-1}, \zeta I), \quad \zeta = e^{2\pi i/n}. \tag{4.5.16}$$

Let $\{\sigma_\alpha : \alpha \in \mathcal{I}\}$ denote a complete set of irreducible unitary representations of $\mathrm{SU}(n)$. By Proposition 2.8.11, a complete set of irreducible unitary representations of $S^1 \times \mathrm{SU}(n)$ is given by $\{\pi_{m\alpha} : m \in \mathbb{Z}, \alpha \in \mathcal{I}\}$, defined by

$$\pi_{m\alpha}(\omega, g) = \omega^m \sigma_\alpha(g). \tag{4.5.17}$$

Such a representation of $S^1 \times \mathrm{SU}(n)$ produces a representation of $\mathrm{U}(n)$ if and only if $\pi_{m\alpha}(K_n) = I$, i.e., if and only if

$$\sigma_\alpha(\zeta I) = \zeta^m I, \tag{4.5.18}$$

when $\zeta = e^{2\pi i/n}$. Now, since ζI is in the center of $\mathrm{SU}(n)$, it follows that, for any $\alpha \in \mathcal{I}$, $\sigma_\alpha(\zeta I)$ is a scalar that is an nth root of unity, i.e.,

$$\sigma_\alpha(\zeta I) = \zeta^\mu I, \quad \mu = \mu(\alpha) \in \mathbb{Z}. \tag{4.5.19}$$

Then $\pi_{m\alpha}$ in (4.5.17) gives a representation of $\mathrm{U}(n)$ if and only if

$$m = \mu(\alpha) \mod n. \tag{4.5.20}$$

Since we have already produced a complete set of irreducible unitary representations of $\mathrm{U}(n)$, it is appropriate to turn this around. We have the following.

Proposition 4.5.2. *Each irreducible unitary representation of* $\mathrm{U}(n)$ *restricts to an irreducible unitary representation of* $\mathrm{SU}(n)$*, and all irreducible unitary representations of* $\mathrm{SU}(n)$ *are obtained in this fashion. Furthermore, two irreducible unitary representations* π_1 *and* π_2 *of* $\mathrm{U}(n)$ *restrict to the same representation of* $\mathrm{SU}(n)$ *if and only if, for some* $j \in \mathbb{Z}$*,*

$$\pi_2(g) = (\det g)^j \, \pi_1(g), \quad \forall \ g \in \mathrm{U}(n). \tag{4.5.21}$$

Hence the set of equivalence classes of irreducible unitary representations of $\mathrm{SU}(n)$ *is parametrized by*

$$\{(d_1, \dots, d_{n-1}, 0) : d_\nu \in \mathbb{Z}, \ d_1 \ge d_2 \ge \cdots \ge d_{n-1} \ge 0\}. \tag{4.5.22}$$

Proof. It remains to show that if π_ℓ are irreducible and $\pi_1 = \pi_2$ on $\mathrm{SU}(n)$, then there exists $j \in \mathbb{Z}$ such that (4.5.21) holds. To see this, suppose π_ℓ are as in (4.5.17), which we rewrite as

$$\pi_\ell(\omega g) = \omega^{m_\ell} \sigma_\ell(g), \quad \omega \in S^1, \ g \in \mathrm{SU}(n), \tag{4.5.23}$$

where σ_ℓ is an irreducible representation of $\mathrm{SU}(n)$, so, as in (22.18),

$$\sigma_\ell(\zeta I) = \zeta^{m_\ell} I, \tag{4.5.24}$$

where $\zeta = e^{2\pi i/n}$. We have

$$\pi_1 = \pi_2 \ \text{ on } \ \mathrm{SU}(n) \Longleftrightarrow \sigma_1 \equiv \sigma_2, \tag{4.5.25}$$

which implies

$$\pi_1(\omega g) = \omega^{m_1 - m_2} \pi_2(\omega g), \quad \forall \, \omega \in S^1, \ g \in \mathrm{SU}(n). \tag{4.5.26}$$

We claim

$$n | (m_1 - m_2), \ \text{ i.e., } m_1 - m_2 = nj, \ j \in \mathbb{Z}. \tag{4.5.27}$$

Since $\det \omega g = \omega^n$, this would give (4.5.21). To verify (4.5.27), we note that (4.5.24)–(4.5.25) give

$$\zeta^{m_1 - m_2} = 1, \tag{4.5.28}$$

from which (4.5.27) follows. This proves (4.5.21). ☐

Recall from §4.4 the notation $\mathcal{D}_{(k_1,\dots,k_n)}$ for an irreducible representation of U(n) with highest weight $(k_1,\dots,k_n)$. We keep this notation for the restriction to SU(n), noting that

$$\mathcal{D}_{(k_1,\dots,k_n)} \approx \mathcal{D}_{(k_1+j,\dots,k_n+j)} \quad \text{on} \quad \mathrm{SU}(n), \quad \forall\, j \in \mathbb{Z}. \tag{4.5.29}$$

Note that the representations $D_{k/2}$ of SU(2) produced in §4.1, by (4.1.25), are given in this nomenclature (cf. (4.4.13), in light of (4.3.2)) as

$$\begin{aligned} D_{k/2} &= \mathcal{D}_{(0,-k)} \\ &\approx \mathcal{D}_{(k,0)}, \quad \text{on} \quad \mathrm{SU}(2), \end{aligned} \tag{4.5.30}$$

the last equivalence (special to $n = 2$) by (4.5.29).

4.6. Analytic continuation from $\mathrm{U}(n)$ to $\mathrm{Gl}(n,\mathbb{C})$ revisited

The following result proved useful in the analysis of the irreducible representations of $\mathrm{U}(n)$ in §4.2.

Theorem 4.6.1. *If π is a representation of $\mathrm{U}(n)$ on a finite dimensional complex vector space V, then π extends to a holomorphic representation of $\mathrm{Gl}(n,\mathbb{C})$ on V.*

This was proven in §4.5. The proof given there extended $d\pi$, $\mathbb{C}$-linearly, from the Lie algebra $\mathfrak{u}(n)$ of $\mathrm{U}(n)$ to the Lie algebra $\mathrm{M}(n,\mathbb{C})$ of $\mathrm{Gl}(n,\mathbb{C})$. Then it was shown that this Lie algebra representation arose from a representation of $\mathrm{Gl}(n,\mathbb{C})$ itself, and not just its universal covering group. This latter step involved some topology. Here we give another proof of Theorem 4.6.1, deriving it from Proposition 4.6.2, below, which is of independent interest.

To set up Proposition 4.6.2, we define the representation $T^{p,q}$ of $\mathrm{U}(n)$ on $T^{p,q}(\mathbb{C}^n) = (\otimes^p\mathbb{C}^n)\otimes(\otimes^q\mathbb{C}^n)$ by

$$\begin{aligned} &T^{p,q}(g)v_1\otimes\cdots\otimes v_p\otimes w_1\otimes\cdots\otimes w_q \\ &= gv_1\otimes\cdots\otimes gv_p\otimes(g^{-1})^t w_1\otimes\cdots\otimes(g^{-1})^t w_q. \end{aligned} \tag{4.6.1}$$

Note that

$$g\in\mathrm{U}(n)\Longleftrightarrow(g^{-1})^t=\overline{g}. \tag{4.6.2}$$

Next, we define the representation T_K of $\mathrm{U}(n)$ on $T_K(\mathbb{C}^n)=\oplus_{p+q\le K}T^{p,q}(\mathbb{C}^n)$ by

$$T_K(g)\Big(\bigoplus_{p+q\le K} v_{pq}\Big) = \bigoplus_{p+q\le K} T^{p,q}(g)v_{pq}. \tag{4.6.3}$$

The following result is closely related to the "easy" proof of the Peter-Weyl theorem for compact matrix groups given in §2.3.

Proposition 4.6.2. *If π is a finite dimensional representation of $\mathrm{U}(n)$ on V, then there exists $K<\infty$ such that π is contained in T_K.*

The content of Proposition 4.6.2 is that, for some K, there is a linear subspace $W\subset T_K(\mathbb{C}^n)$, invariant under the action of T_K, and a linear isomorphism $J: V\to W$ such that

$$\pi(g) = J^{-1}T_K(g)J, \tag{4.6.4}$$

for all $g\in\mathrm{U}(n)$. Given this result, Theorem 4.6.1 has the following simple proof. The formula (4.6.1) is clearly well defined for $g\in\mathrm{Gl}(n,\mathbb{C})$, holomorphic in g, and this formula together with (4.6.3) provides an explicit extension of T_K from $\mathrm{U}(n)$ to $\mathrm{Gl}(n,\mathbb{C})$.

We claim that (4.6.4) extends π from $\mathrm{U}(n)$ to $\mathrm{Gl}(n,\mathbb{C})$. To see this, we start with the fact that

$$T_K(g) : W \longrightarrow W, \tag{4.6.5}$$

for all $g \in \mathrm{U}(n)$. We want to show that (4.6.5) holds for all $g \in \mathrm{Gl}(n,\mathbb{C})$. Indeed, the validity of (4.6.5) for all $g \in \mathrm{U}(n)$ implies

$$dT_K(X) : W \longrightarrow W, \tag{4.6.6}$$

for all $X \in \mathfrak{u}(n)$, i.e., $X = -X^*$. Since T_K is holomorphic, $dT_K : M(n,\mathbb{C}) \to \mathcal{L}(T_K(\mathbb{C}^n))$ is $\mathbb{C}$-linear, hence (4.6.6) holds for all $X \in M(n,\mathbb{C})$. Thus

$$T_K(e^{tX}) = e^{t\, dT_K(X)} : W \longrightarrow W, \tag{4.6.7}$$

for all $X \in M(n,\mathbb{C})$, so (4.6.5) holds for all g in a neighborhood of I in $\mathrm{Gl}(n,\mathbb{C})$, hence for all $g \in \mathrm{Gl}(n,\mathbb{C})$. Thus we have Theorem 4.6.1, once we have proved Proposition 4.6.2.

To prove Proposition 4.6.2, we can produce hermitian inner products so that π and T_K are unitary representations. Also π breaks up into irreducible pieces, and it suffices to treat each piece. Thus we can assume π is irreducible. Let us assume such π is not contained in T_K for any K and obtain a contradiction.

Let $\mathcal{L}$ denote the linear span of the matrix entries of T_K, as K varies over $\mathbb{N}$. If π is not contained in any T_K, it is not equivalent to any of the irreducible representations into which T_K breaks up, so by the Weyl orthogonality relations it follows that the matrix entries of π must be orthogonal to each element of $\mathcal{L}$, in $L^2(\mathrm{U}(n))$. However, from the construction (4.6.1)–(4.6.3) it is clear that $\mathcal{L}$ is an algebra of continuous functions on $\mathrm{U}(n)$, invariant under complex conjugation, and $\mathcal{L}$ separates the points of $\mathrm{U}(n)$. Hence, by the Stone-Weierstrass theorem, $\mathcal{L}$ is dense in $C(\mathrm{U}(n))$, and a fortiori dense in $L^2(\mathrm{U}(n))$. This contradiction proves Proposition 4.6.2.

We complement Theorem 4.6.1 with the following uniqueness result.

Proposition 4.6.3. *In the setting of Theorem 4.6.1, the extension of π on $U(n)$ to a holomorphic representation of $Gl(n,\mathbb{C})$ on V is unique.*

Proof. Suppose π_1 and π_2 are holomorphic representations of $Gl(n,\mathbb{C})$ that agree on $U(n)$:

$$\pi_1(g) = \pi_2(g) = \pi(g), \quad \forall\, g \in U(n). \tag{4.6.8}$$

The representations π_j have derived representations

$$d\pi_j : M(n,\mathbb{C}) \longrightarrow \mathcal{L}(V), \tag{4.6.9}$$

Lie algebra homomorphisms, and

$$d\pi_1(X) = d\pi_2(X) = d\pi(X), \quad \forall\, X = -X^* \in M(n, \mathbb{C}). \tag{4.6.10}$$

Furthermore, the fact that π_j are holomorphic implies

$$d\pi_j : M(n, \mathbb{C}) \longrightarrow \mathcal{L}(V) \ \text{ are } \mathbb{C}\text{-linear}. \tag{4.6.11}$$

Since each $Y \in M(n, \mathbb{C})$ can be written as $Y = X_1 + iX_2$, with $X_j^* = -X_j$, this forces

$$d\pi_1 = d\pi_2 : M(n, \mathbb{C}) \longrightarrow \mathcal{L}(V). \tag{4.6.12}$$

Hence

$$\pi_1(e^{tY}) = e^{t\, d\pi_1(Y)} = e^{t\, d\pi_2(Y)} = \pi_2(e^{tY}), \quad \forall\, Y \in M(n, \mathbb{C}). \tag{4.6.13}$$

Thus $\pi_1(g) = \pi_2(g)$ for all g in a neighborhood of I in $Gl(n, \mathbb{C})$, and this implies

$$\pi_1(g) = \pi_2(g), \quad \forall\, g \in Gl(n, \mathbb{C}). \tag{4.6.14}$$

$\square$

4.7. Decomposition of $S^k \otimes \overline{S}^\ell$

Here we consider how to decompose the representation

$$S_\ell^k = S^k \otimes \overline{S}^\ell \tag{4.7.1}$$

of U(n) into irreducible pieces. This representation acts on $\mathcal{P}_k \otimes \mathcal{P}_\ell$, which we can identify with the space of polynomials $p(z, w)$, homogeneous of degree k in z and ℓ in w. We have

$$S_\ell^k(g)p(z, w) = p(g^t z, g^{-1} w), \tag{4.7.2}$$

for $g \in \mathrm{U}(n)$, extending holomorphically to $g \in \mathrm{Gl}(n, \mathbb{C})$. This induces an action $dS_\ell^k(X)$ on such polynomials, for X in $\mathfrak{u}(n)$, and its complexification $\mathrm{M}(n, \mathbb{C})$. Parallel to (4.3.4) and (4.3.8) we have

$$dS_\ell^k(e_{\mu\nu}) = z_\mu \frac{\partial}{\partial z_\nu} - w_\nu \frac{\partial}{\partial w_\mu}. \tag{4.7.3}$$

To decompose S_ℓ^k into irreducible pieces, it will be helpful to identify the set of elements of $\mathcal{P}_k \otimes \mathcal{P}_\ell$ annihilated by all raising operators, i.e., by all operators of the form (4.7.3) with $\mu < \nu$. The following result accomplishes this.

Lemma 4.7.1. *If $p(z, w)$ is a polynomial annihilated by all operators of the form (4.7.3) with $\mu < \nu$, then*

$$p(z, w) = q(z_1, w_n, z \cdot w), \tag{4.7.4}$$

for some polynomial q on $\mathbb{C}^3$, where $z \cdot w = z_1 w_1 + \cdots + z_n w_n$.

Proof. The polynomials we are considering can be characterized by

$$p(g^t z, g^{-1} w) = p(z, w), \quad \forall g \in N_+.$$

In particular, $p(z, w)$ is invariant under the action of one-parameter subgroups:

$$z_\nu \mapsto z_\nu + t z_\mu, \quad w_\mu \mapsto w_\mu - t w_\nu, \quad \mu < \nu. \tag{4.7.5}$$

Suppose $z_1 \neq 0$ and take, successively for $\nu = 2, \dots, n$, parameters t such that $z_\nu + t z_1 = 0$. We deduce that

$$p(z, w) = p((z_1, 0, \dots, 0), \tilde{w}), \tag{4.7.6}$$

where $\tilde{w}$ differs from w only in the first coordinate. Note that $z \cdot w = z_1 \tilde{w}_1$, since $g^t z \cdot g^{-1} w = z \cdot w$ for all $g \in \mathrm{Gl}(n, \mathbb{C})$. Next, taking $\nu = n$ and $\mu \in \{1, \dots, n-1\}$ in (4.7.5), if $w_n \neq 0$ we can transform $\tilde{w}$ to a vector whose first $n - 1$ coordinates vanish, while leaving unchanged its last coordinate, and also leaving unchanged all coordinates of $(z_1, 0, \dots, 0)$ but the last. Hence (4.7.5) implies

$$p(z, w) = p((z_1, 0, \dots, 0, \zeta), (0, \dots, 0, w_n)). \tag{4.7.7}$$

Again $z \cdot w = (z_1, 0, \dots, 0, \zeta) \cdot (0, \dots, 0, w_n) = \zeta w_n$, so

$$\zeta = \frac{z \cdot w}{w_n}. \tag{4.7.8}$$

Consequently, for $p \in \mathcal{P}_k \otimes \mathcal{P}_\ell$, (23.7) yields

$$p(z, w) = w_n^{\ell - k} p((z_1 w_n, 0, \dots, 0, z \cdot w), (0, \dots, 0, 1)),$$

for $z_1 \neq 0, w_n \neq 0$, an identity that clearly extends to $z_1 = 0$. If $\ell \geq k$, it also extends to $w_n = 0$, yielding (4.7.4).

If $\ell < k$, we can argue in the opposite order, obtaining the following analogue of (4.7.7):

$$p(z, w) = p((z_1, 0, \dots, 0), (\zeta, 0, \dots, 0, w_n)),$$

this time with $\zeta = (z \cdot w)/z_1$. Hence, for $p \in \mathcal{P}_k \otimes \mathcal{P}_\ell$,

$$p(z, w) = z_1^{k-\ell} p((1, 0, \dots, 0), (z \cdot w, 0, \dots, 0, z_1 w_n)),$$

for $z_1 \neq 0, w_n \neq 0$. This extends to $w_n = 0$ and also, given $k > \ell$, to $z_1 = 0$, and again we have (4.7.4). □

It follows from Lemma 4.7.1 that the space $\mathcal{Z}_{k\ell}$ of elements of $\mathcal{P}_k \otimes \mathcal{P}_\ell$ annihilated by all raising operators is spanned by

$$\psi_{k\ell\mu}(z, w) = z_1^{k-\mu} w_n^{\ell-\mu} (z_1 w_1 + \cdots + z_n w_n)^\mu, \quad 0 \leq \mu \leq k \wedge \ell. \tag{4.7.9}$$

Each of these elements is a weight vector for S_ℓ^k. In fact,

$$\begin{aligned} dS_\ell^k(e_\nu)\psi_{k\ell\mu}(z, w) &= i\Big(z_\nu \frac{\partial}{\partial z_\nu} - w_\nu \frac{\partial}{\partial w_\nu}\Big)\psi_{k\ell\mu}(z, w) \\ &= i[(k-\mu)\delta_{\nu 1} - (\ell - \mu)\delta_{\nu n}]\psi_{k\ell\mu}(z, w). \end{aligned} \tag{4.7.10}$$

The weight so obtained is

$$(k - \mu, 0, \dots, 0, \mu - \ell), \quad 0 \leq \mu \leq k \wedge \ell. \tag{4.7.11}$$

Alternatively, note that if $g = \mathrm{diag}(c_1, \dots, c_n)$ then

$$\begin{aligned} S_\ell^k(g)\psi_{k\ell\mu}(z, w) &= \psi_{k\ell\mu}(g^t z, g^{-1} w) \\ &= c_1^{k-\mu} c_n^{-(\ell-\mu)} z_1^{k-\mu} w_n^{\ell-\mu} (z \cdot w)^\mu \\ &= c_1^{k-\mu} c_n^{\mu-\ell} \psi_{k\ell\mu}(z, w), \end{aligned} \tag{4.7.12}$$

again leading to (4.7.11). These calculations establish the following.

Proposition 4.7.2. *For $k, \ell \geq 0$ we have*

$$S^k \otimes \overline{S}^\ell \approx \bigoplus_{0 \leq \mu \leq k \wedge \ell} \mathcal{D}_{(k-\mu,0,\dots,0,\mu-\ell)}, \tag{4.7.13}$$

as representations of $\mathrm{U}(n)$*. The highest weight vectors for the irreducible components on the right side of* (4.7.13) *are given by* (4.7.9).

Remark 1. For $n = 2$, this result captures the Clebsch-Gordon series (4.1.52), in view of the identities in (4.5.30).

Remark 2. In case $k = \ell = 1$, we have $S^1 \otimes \overline{S}^1 \approx \mathrm{Ad}$, analyzed in §20. Compare this case of (4.7.13) with (4.4.16).

Note that the highest weight that occurs in (4.7.13) is $(k, 0, \dots, 0, -\ell)$. We specifically identify the subspace of $\mathcal{P}_k \otimes \mathcal{P}_\ell$ on which S^k_ℓ acts like $\mathcal{D}_{(k,0,\dots,0,-\ell)}$.

Proposition 4.7.3. *The irreducible component of $\mathcal{P}_k \otimes \mathcal{P}_\ell$ containing the highest weight vector $\psi_{k\ell 0}(z, w) = z_1^k w_n^\ell$ is given by*

$$\mathcal{P}^\#_{k\ell} = \Big\{ p(z, w) \in \mathcal{P}_k \otimes \mathcal{P}_\ell : \sum_{j=1}^n \frac{\partial^2 p}{\partial z_j \partial w_j} = 0 \Big\}. \tag{4.7.14}$$

Proof. That $\mathcal{P}^\#_{k\ell}$ is invariant under the action of $\mathrm{U}(n)$ follows from the fact that the operator $\sum \partial^2/\partial z_j \partial w_j$ commutes with all the operators in (4.7.3). Now we consider which elements of $\mathcal{P}^\#_{k\ell}$ are annihilated by all raising operators, i.e., we identify the intersection of $\mathcal{P}^\#_{k\ell}$ with the linear span of the elements $\psi_{k\ell\mu}$ given by (4.7.9). A calculation gives

$$\sum_{j=1}^n \frac{\partial^2}{\partial z_j \partial w_j} \psi_{k\ell\mu}(z, w) = \mu(n - 1 + k + \ell - \mu) \, \frac{\psi_{k\ell\mu}(z, w)}{z \cdot w}. \tag{4.7.15}$$

Hence the only element in $\mathcal{P}^\#_{k\ell}$ annihilated by all raising operators is (up to a scalar multiple) $\psi_{k\ell 0}(z, w) = z_1^k w_n^\ell$. This establishes irreducibility of the action of $\mathrm{U}(n)$ on $\mathcal{P}^\#_{k\ell}$ and finishes the proof. □

In conclusion, we see that

$$\mathcal{D}_{(k,0,\dots,0,-\ell)} \ \text{ is realized on } \ \mathcal{P}^\#_{k\ell}. \tag{4.7.16}$$

Let us specialize to $n = 3$. We have representations

$$\mathcal{D}_{(k,0,-\ell)} \ \text{ of } \ \mathrm{U}(3) \ \text{ on } \ \mathcal{P}^\#_{k\ell}. \tag{4.7.17}$$

Multiplying by $(\det g)^j$ gives representations

$$\mathcal{D}_{(k+j,j,j-\ell)} \quad \text{of} \quad \mathrm{U}(3) \quad \text{on} \quad \mathcal{P}^{\#}_{k\ell}. \tag{4.7.18}$$

The results of §4.4 show that (4.7.18) produces a complete set of irreducible representations of U(3).

We can produce an alternative realization of (4.7.16) as follows. An element of $\mathcal{P}_k \otimes \mathcal{P}_\ell$ can be written

$$p(z,w) = \sum A^{j_1\cdots j_k}_{i_1\cdots i_\ell} z_{j_1} \cdots z_{j_k} w_{i_1} \cdots w_{i_\ell}, \tag{4.7.19}$$

with $A^{\cdots}_{\cdots}$ symmetric in the js and the is. A computation gives

$$\sum_{j=1}^n \frac{\partial^2}{\partial z_j \partial w_j} p(z,w) = \sum_{\nu,i,j} A^{\nu j_2\cdots j_k}_{\nu i_2\cdots i_\ell} z_{j_2} \cdots z_{j_k} w_{i_2} \cdots w_{i_\ell}. \tag{4.7.20}$$

In other words, we have $\mathrm{U}(n)$ acting on

$$\mathcal{P}_k \otimes \mathcal{P}_\ell \approx (S^k\mathbb{C}^n) \otimes (S^\ell\mathbb{C}^n)', \tag{4.7.21}$$

and

$$\mathcal{P}^{\#}_{k\ell} \approx \bigl\{A^{j_1\cdots j_k}_{i_1\cdots i_\ell} \in (S^k\mathbb{C}^n) \otimes (S^\ell\mathbb{C}^n)' : A^{\nu j_2\cdots j_k}_{\nu i_2\cdots i_\ell} \equiv 0\bigr\}, \tag{4.7.22}$$

where the summation convention is indicated over ν.

Let us return to (4.7.13) and set $k = \ell = 1$, so (4.7.13) gives

$$S^1 \otimes \overline{S}^1 \approx \mathcal{D}_{(1,0,\dots,0,-1)} \oplus \mathcal{D}_{(0,\dots,0)}. \tag{4.7.23}$$

Note that S^1 acts on $\mathbb{C}^n$ and $\overline{S}^1$ acts on $(\mathbb{C}^n)'$, and via $\mathbb{C}^n \otimes (\mathbb{C}^n)' \approx \mathrm{M}(n,\mathbb{C})$, we have

$$S^1 \otimes \overline{S}^1 \approx \mathrm{Ad}, \tag{4.7.24}$$

so (4.7.23) is equivalent to (4.4.16).

Contrast this with the decomposition of $\mathbb{C}^n \otimes \mathbb{C}^n$ into symmetric and antisymmetric 2-tensors. This yields

$$S^1 \otimes S^1 \approx \Lambda^2 \oplus S^2 \approx \mathcal{D}_{(1,1,0,\dots,0)} \oplus \mathcal{D}_{(2,0,\dots,0)}. \tag{4.7.25}$$

The decomposition of $\otimes^k\mathbb{C}^n$ into irreducible spaces for larger k will be studied in §4.10.

4.8. Commutants, double commutants, and dual pairs

In this section we make some general observations on decomposing a unitary representation π of a compact Lie group G on a finite-dimensional space V into irreducible pieces. Recall that π is irreducible if and only if the set of operators on V commuting with $\pi(g)$ for all $g \in G$ consists of scalar multiples of the identity. In the general case, it is useful to look at

$$\mathcal{A} = \text{algebra generated by } \pi(g) : g \in G, \quad \mathcal{A} \subset \operatorname{End}(V), \tag{4.8.1}$$

and its *commutant*, defined by

$$\mathcal{A}' = \{B \in \operatorname{End}(V) : BA = AB, \ \forall\ A \in \mathcal{A}\}. \tag{4.8.2}$$

As we know, V can be decomposed into irreducible subspaces. Say

$$V = \bigoplus_{j=1}^{k} n_j V_j, \tag{4.8.3}$$

where $n_j V_j = V_j \oplus \cdots \oplus V_j$ (n_j factors), with π acting irreducibly on each V_j (call the irreducible representation U_j). Arrange the decomposition (4.8.3) so that distinct js correspond to inequivalent U_js. We can write $n_j V_j = V_j \otimes W_j,\ W_j \approx \mathbb{C}^{n_j}$, i.e.,

$$V = \bigoplus_{j=1}^{k} V_j \otimes W_j, \tag{4.8.4}$$

with $\pi_j = \pi|_{V_j \otimes W_j}$ given by

$$\pi_j(g) = U_j(g) \otimes I, \tag{4.8.5}$$

and

$$U_j \ \text{ irreducible on } \ V_j, \quad j_1 \neq j_2 \Rightarrow U_{j_1} \ \text{ inequivalent to } \ U_{j_2}. \tag{4.8.6}$$

The following result records some important structure.

Proposition 4.8.1. *In the set-up described above,*

$$\mathcal{A} = \Big\{ \bigoplus_{j=1}^{k} A_j \otimes I : A_j \in \operatorname{End}(V_j) \Big\}. \tag{4.8.7}$$

If we set

$$\mathcal{B} = \mathcal{A}', \tag{4.8.8}$$

then

$$\mathcal{B} = \Big\{ \bigoplus_{j=1}^{k} I \otimes B_j : B_j \in \operatorname{End}(W_j) \Big\}. \tag{4.8.9}$$

Furthermore,

$$\mathcal{B}' = \mathcal{A}. \tag{4.8.10}$$

Proof. It is immediate from (4.8.5) that every element of $\mathcal{A}$ has the form given on the right side of (4.8.7). For the converse, pick a basis $u_1^{(j)}, \dots, u_{d_j}^{(j)}$ for V_j and, for $\ell, m \in \{1, \dots, d_j\}$, let $e_{\ell m}^{(j)} \in \mathrm{End}(V_j)$ be given by $e_{\ell m}^{(j)} u_\mu^{(j)} = \delta_{\mu m} u_\ell^{(j)}$. Also write $(U_{\ell m}^{(j)}(g))$ as the matrix representation of $U_j(g) \in \mathrm{End}(V_j)$ with respect to this basis. It follows from the Weyl orthogonality relations (2.2.6)–(2.2.7) that

$$d_j \int_G \overline{U}_{\ell m}^{(j)}(g)\, \pi(g)\, dg = e_{\ell m}^{(j)} \otimes I, \tag{4.8.11}$$

on $V_j \otimes W_j$, and vanishes on $V_\ell \otimes W_\ell$ for $\ell \neq j$, so every element on the right side of (4.8.7) is a limit of superpositions of elements of $\mathcal{A}$, hence an element of $\mathcal{A}$ (since the linear subspace $\mathcal{A}$ of the finite-dimensional space $\mathrm{End}(V)$ must be closed). Thus we have (4.8.7).

To prove (4.8.9), first note that whenever $\mathcal{A}$ is given by (4.8.7), then clearly the right side of (4.8.9) is contained in $\mathcal{A}'$. We establish the reverse inclusion. Let P_j be the orthogonal projection of V onto $V_j \otimes W_j$. By (4.8.7), $\mathcal{P}_j \in \mathcal{A}$. Hence $B \in \mathcal{A}' \Rightarrow BP_j = P_j B$, $1 \le j \le k$, i.e., B leaves each $V_j \otimes W_j$ invariant; say $B|_{V_j \otimes W_j} = \widetilde{B}_j$. We have

$$\widetilde{B}_j : V_j \otimes W_j \to V_j \otimes W_j, \quad \widetilde{B}_j(A \otimes I) = (A \otimes I)\widetilde{B}_j, \quad \forall\, A \in \mathrm{End}(V_j). \tag{4.8.12}$$

Taking $A = e_{\ell\ell}^{(j)}$ we see that $\widetilde{B}_j$ leaves invariant each space $(u_\ell^{(j)}) \otimes W_j$. Taking a basis $w_1^{(j)}, \dots, w_{n_j}^{(j)}$ of W_j, we have

$$\widetilde{B}_j(u_\ell^{(j)} \otimes w_m^{(j)}) = \sum_\mu \beta_{\ell m}^\mu\, u_\ell^{(j)} \otimes w_\mu^{(j)}. \tag{4.8.13}$$

If we next take $A = e_{\ell\nu}^{(j)} \otimes I$ and compute $\widetilde{B}_j A(u_\nu^{(j)} \otimes w_m^{(j)})$ and $A\widetilde{B}_j(u_\nu^{(j)} \otimes w_m^{(j)})$ and compare, we see that $\beta_{\ell m}^\mu = \beta_{\nu m}^\mu$, i.e., $\beta_{\ell m}^\mu$ is independent of ℓ. Hence $\widetilde{B}_j = I \otimes B_j$ with $B_j \in \mathrm{End}(W_j)$, proving (4.8.9). The way we got (4.8.9) from (4.8.7) immediately gives (4.8.10). □

The result (4.8.10) is a special case of a result known as the double commutant theorem. It holds when $\mathcal{A}$ is a subalgebra of $\mathrm{End}(V)$ closed under adjoints (we say $\mathcal{A}$ is a C^*-algebra). In fact there is a far ranging extension to a special class of C^*-algebras (called von Neumann algebras) valid when V is an infinite dimensional Hilbert space. See [**33**].

Now we add structure by bringing in two groups, acting on V.

Proposition 4.8.2. *Let G and K be compact Lie groups, π a unitary representation of G on V, τ a unitary representation of K on V. Let*

$$\mathcal{A} = \textit{algebra generated by } \pi(g),\ g \in G,$$
$$\mathcal{B} = \textit{algebra generated by } \tau(k),\ k \in K. \tag{4.8.14}$$

Assume

$$\mathcal{A}' = \mathcal{B}. \tag{4.8.15}$$

Let $\mathcal{S}_\pi = \{\alpha_j\}$ denote the set of irreducible unitary representations of G contained in π, and let $\mathcal{S}_\tau = \{\beta_j\}$ denote the set of irreducible unitary representations of K contained in τ (up to equivalence). Then there exists a bijective map $Q : \mathcal{S}_\pi \to \mathcal{S}_\tau$ and a decomposition

$$V = \bigoplus_{j=1}^{\ell} V_j \otimes W_j, \quad \ell = \#(\mathcal{S}_\pi) = \#(\mathcal{S}_\tau), \tag{4.8.16}$$

such that

$$\pi(g) = \bigoplus_{j=1}^{\ell} \alpha_j(g) \otimes I, \quad \tau(k) = \bigoplus_{j=1}^{\ell} I \otimes \beta_{Q(j)}(k). \tag{4.8.17}$$

Proof. The representation π of G decomposes as in (4.8.3). The orthogonal projections P_j of V on $n_j V_j$ are the minimal projections in the center of $\mathcal{A}$, so such minimal projections match up bijectively with $\mathcal{S}_\pi$. Similarly $\mathcal{S}_\tau$ is in one-to-one correspondence with the minimal orthogonal projections in the center of $\mathcal{B}$. Now we are given that $\mathcal{A}' = \mathcal{B}$, and hence, by Proposition 4.8.1, $\mathcal{B}' = \mathcal{A}$. Hence central projections in $\mathcal{A}$ are precisely projections in $\mathcal{A} \cap \mathcal{B}$ and similarly for central projections in $\mathcal{B}$. Thus both $\mathcal{S}_\pi$ and $\mathcal{S}_\tau$ are in one-to-one correspondence with the same set of projections.

Let us focus on the range of the projection P_j, relabeling this space as V, so π contains n_1 copies of one irreducible representation (say α_1) of G and τ contains m_1 copies of one irreducible representation (say β_1) of K. Our final claim is that in such a case

$$V \approx V_1 \otimes W_1, \tag{4.8.18}$$

with

$$\pi(g) = \alpha_1(g) \otimes I, \quad \tau(k) = I \otimes \beta_1(k), \tag{4.8.19}$$

given (4.8.15). In fact Proposition 4.8.1 gives a tensor product decomposition (4.8.18) such that (4.8.19) holds for $\pi(g)$. Then the fact that $\mathcal{A}'$ is given by (4.8.9) also puts $\tau(k)$ in the general form indicated in (4.8.19), i.e.,

$$\tau(k) = I \otimes \gamma(k).$$

It remains to dispose of the possibility that γ is a sum of several copies of an irreducible representation (i.e., of β_1). Indeed, again by Proposition 4.8.1, the commutant of the set of operators generated by $\alpha_1(g) \otimes I$ is all of $I \otimes \mathrm{End}(W_1)$, which (by hypothesis) is the algebra generated by $I \otimes \gamma(k)$, so γ cannot decompose into several irreducibles. $\square$

When compact G and K act on V as in Proposition 4.8.2, with (4.8.15) holding, we say G and K act as a *dual pair* on V. A key family of examples of dual pairs will be given in §4.9.

Note that, in the setting of Proposition 4.8.2, $\pi(g)\tau(k)$ gives a representation of $G \times K$ on V, and (4.8.17) gives

$$\pi(g)\tau(k) = \bigoplus_{j=1}^{\ell} \alpha_j(g) \otimes \beta_{Q(j)}(k). \tag{4.8.20}$$

In particular, taking traces gives

$$\operatorname{Tr} \pi(g)\tau(k) = \sum_{j=1}^{\ell} \chi_{\alpha_j}(g)\, \chi_{\beta_{Q(j)}}(k). \tag{4.8.21}$$

4.9. The first fundamental theorem of invariant theory

The group $\mathrm{U}(n)$ acts on $\otimes^k \mathbb{C}^n$ via

$$\otimes^k g(v_1 \otimes \cdots \otimes v_k) = gv_1 \otimes \cdots \otimes gv_k, \quad g \in \mathrm{U}(n),\ v_\nu \in \mathbb{C}^n. \tag{4.9.1}$$

In addition, the permutation group S_k acts on $\otimes^k \mathbb{C}^n$ via

$$\tau(\sigma)(v_1 \otimes \cdots \otimes v_k) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}, \quad \sigma \in S_k. \tag{4.9.2}$$

It is clear that $\otimes^k g$ commutes with $\tau(\sigma)$ for each $g \in \mathrm{U}(n),\ \sigma \in S_k$, so we get a representation of $S_k \times \mathrm{U}(n)$ on $\otimes^k \mathbb{C}^n$. The following is the key result of this section.

Proposition 4.9.1. *The groups S_k and $\mathrm{U}(n)$ act as a dual pair on $\otimes^k \mathbb{C}^n$.*

To restate this, let

$$\begin{aligned} \mathcal{A} &= \text{algebra generated by } \tau(\sigma),\ \sigma \in S_k, \quad \mathcal{A} \subset \mathrm{End}(\otimes^k \mathbb{C}^n), \\ \mathcal{B} &= \text{algebra generated by } \otimes^k g,\ g \in \mathrm{U}(n). \end{aligned} \tag{4.9.3}$$

It is clear that $\mathcal{B} \subset \mathcal{A}'$ and $\mathcal{A} \subset \mathcal{B}'$, as we have already mentioned. To prove Proposition 4.9.1, we will show that

$$\mathcal{A}' = \mathcal{B}. \tag{4.9.4}$$

In view of Proposition 4.8.1, this gives also

$$\mathcal{B}' = \mathcal{A}. \tag{4.9.5}$$

Our treatment follows [**34**].

To begin our analysis of $\mathcal{A}'$, we note that

$$\mathrm{End}(\otimes^k \mathbb{C}^n) \approx \otimes^k \mathrm{End}(\mathbb{C}^n), \tag{4.9.6}$$

via

$$A_1 \otimes \cdots \otimes A_k(v_1 \otimes \cdots \otimes v_k) = A_1 v_1 \otimes \cdots \otimes A_k v_k. \tag{4.9.7}$$

In fact, (4.9.7) yields a homomorphism $\otimes^k \mathrm{End}(\mathbb{C}^n) \to \mathrm{End}(\otimes^k \mathbb{C}^n)$. One verifies that this map is injective, hence bijective, since the dimensions of the two sides of (4.9.6) are equal. We let $\sigma \in S_k$ act on $\otimes^k \mathrm{End}(\mathbb{C}^n)$ by

$$T(\sigma)\, A_1 \otimes \cdots \otimes A_k = A_{\sigma(1)} \otimes \cdots \otimes A_{\sigma(k)}. \tag{4.9.8}$$

Lemma 4.9.2. *Given $X \in \mathrm{End}(\otimes^k \mathbb{C}^n),\ \sigma \in S_k$,*

$$T(\sigma)X = \tau(\sigma) X \tau(\sigma)^{-1}. \tag{4.9.9}$$

Proof. It suffices to check (4.9.9) for $X = A_1 \otimes \cdots \otimes A_k$. Then

$$\begin{aligned}\tau(\sigma)(A_1 \otimes \cdots \otimes A_k)\tau(\sigma)^{-1}(v_1 \otimes \cdots \otimes v_k) \\ &= \tau(\sigma)\, A_1 \otimes \cdots \otimes A_k(v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}) \\ &= \tau(\sigma)\, A_1 v_{\sigma^{-1}(1)} \otimes \cdots \otimes A_k v_{\sigma^{-1}(k)} \\ &= A_{\sigma(1)} v_1 \otimes \cdots \otimes A_{\sigma(k)} v_k,\end{aligned} \tag{4.9.10}$$

which gives (4.9.9). ∎

At this point we have

$$\begin{aligned}\mathcal{A}' &= \{X : \tau(\sigma)X = X\tau(\sigma),\ \forall\, \sigma \in S_k\} \\ &= \{X : T(\sigma)X = X,\ \forall\, \sigma \in S_k\} \\ &= S^k \operatorname{End}(\mathbb{C}^n).\end{aligned} \tag{4.9.11}$$

The next lemma is an exercise in linear algebra.

Lemma 4.9.3. *If W is a finite-dimensional vector space,*

$$S^k W = \mathit{Span}\,\{a \otimes \cdots \otimes a : a \in W\}. \tag{4.9.12}$$

Hence we have from (4.9.11)

$$\mathcal{A}' = \operatorname{Span}\{A \otimes \cdots \otimes A : A \in \operatorname{End}(\mathbb{C}^n)\}. \tag{4.9.13}$$

By comparison,

$$\mathcal{B} = \operatorname{Span}\{g \otimes \cdots \otimes g : g \in \mathrm{U}(n)\}. \tag{4.9.14}$$

To prove (4.9.4), it remains to show that the spaces (4.9.13) and (4.9.14) coincide. To see this, note that, for $Y \in \mathfrak{u}(n)$,

$$d\otimes^k(Y)(v_1 \otimes \cdots \otimes v_k) = Y v_1 \otimes v_2 \otimes \cdots \otimes v_k + \cdots + v_1 \otimes \cdots \otimes v_{k-1} \otimes Y v_k \tag{4.9.15}$$

has the property that

$$d \otimes^k (Y) \in \mathcal{B}, \tag{4.9.16}$$

since this is a limit of difference quotients of elements of $\mathcal{B}$, by (4.9.14). Then (4.9.16) holds for all $Y \in \mathbb{C}\mathfrak{u}(n) = \operatorname{End}(\mathbb{C}^n)$, and exponentiating this gives

$$A \otimes \cdots \otimes A \in \mathcal{B}, \quad \forall\, A \in \mathrm{Gl}(n, \mathbb{C}). \tag{4.9.17}$$

Since $\mathrm{Gl}(n, \mathbb{C})$ is dense in $\operatorname{End}(\mathbb{C}^n)$, we have $A \otimes \cdots \otimes A \in \mathcal{B}$ for all $A \in \operatorname{End}(\mathbb{C}^n)$, so in view of (4.9.13) we now have $\mathcal{A}' = \mathcal{B}$, as advertised in (4.9.4), and Proposition 4.9.1 is proven.

It is useful to restate the result (4.9.5) as follows. Define the representation ϑ_{nk} of $\mathrm{U}(n)$ on $\operatorname{End}(\otimes^k \mathbb{C}^n)$ by

$$\vartheta_{nk}(g)A = (\otimes^k g)A(\otimes^k g^{-1}), \quad g \in \mathrm{U}(n),\ A \in \operatorname{End}(\otimes^k \mathbb{C}^n). \tag{4.9.18}$$

Denote by $\mathcal{E}_{nk}$ the subspace of $\mathrm{End}(\otimes^k\mathbb{C}^n)$ on which ϑ_{nk} acts trivially. Then (4.9.5) implies that $\mathcal{E}_{nk}$ is spanned by the operators $\tau(\sigma)$, $\sigma \in S_k$, given by (4.9.2). To elaborate, (4.9.2) yields a linear map

$$\tau_{nk}^{\#} : \ell^1(S_k) \longrightarrow \mathrm{End}(\otimes^k\mathbb{C}^n), \tag{4.9.19}$$

as a special case of the construction in §2.7, and

$$\mathcal{E}_{nk} = \text{Range of } \ \tau_{nk}^{\#}. \tag{4.9.20}$$

Note that

$$\vartheta_{nk} \approx (\otimes^k) \otimes (\overline{\otimes}^k), \tag{4.9.21}$$

acting on $(\otimes^k\mathbb{C}^n) \otimes (\otimes^k\mathbb{C}^n)'$, via

$$g\cdot(v_1\otimes\cdots\otimes v_k\otimes w_1\otimes\cdots\otimes w_k) = gv_1\otimes\cdots\otimes gv_k\otimes g'w_1\otimes\cdots\otimes g'w_k, \tag{4.9.22}$$

where $g' = (g^t)^{-1}$, so $g \in \mathrm{U}(n) \Rightarrow g' = \overline{g}$. Hence $\mathcal{E}_{nk}$ is isomorphic to the space

$$E_{nk} \subset (\otimes^k\mathbb{C}^n) \otimes (\otimes^k\mathbb{C}^n)' \ \text{ where } \ \mathrm{U}(n) \ \text{ acts trivially.} \tag{4.9.23}$$

The following restatement of (4.9.20) is known as the first fundamental theorem of invariant theory (for unitary invariants).

Proposition 4.9.4. *The space $E_{nk} \subset (\otimes^k\mathbb{C}^n) \otimes (\otimes^k\mathbb{C}^n)'$ on which $\mathrm{U}(n)$ acts trivially is spanned by $\{t_\sigma : \sigma \in S_k\}$, where*

$$t_\sigma(w_1 \otimes \cdots \otimes w_k, v_1 \otimes \cdots \otimes v_k) = \langle v_1, w_{\sigma(1)}\rangle \cdots \langle v_k, w_{\sigma(k)}\rangle, \tag{4.9.24}$$

with $v_\nu \in \mathbb{C}^n$, $w_\nu \in (\mathbb{C}^n)'$, and the standard identification of $V \otimes V'$ with the space of bilinear maps on $V' \times V$.

It follows from Proposition 2.1.5 that the orthogonal projection of $(\otimes^k\mathbb{C}^n)\otimes(\otimes^k\mathbb{C}^n)'$ onto E_{nk} is given by

$$P_{nk} = \int\limits_{\mathrm{U}(n)} (\otimes^k g) \otimes (\otimes^k \overline{g})\, dg. \tag{4.9.25}$$

Hence

$$\dim E_{nk} = \mathrm{Tr}\, P_{nk} = \int\limits_{\mathrm{U}(n)} |\,\mathrm{Tr}\, g|^{2k}\, dg. \tag{4.9.26}$$

A computation of this dimension is of interest; "half" the cases are elementary:

Proposition 4.9.5. *If $k \le n$, then the map $\tau_{nk}^{\#}$ in (4.9.19) is injective. Hence*

$$k \le n \Longrightarrow \dim E_{nk} = k! \tag{4.9.27}$$

Proof. Let $\{u_1, \ldots, u_n\}$ denote the standard basis of $\mathbb{C}^n$. If $k \le n$, then the elements

$$\tau(\sigma)\, u_1 \otimes \cdots \otimes u_k = u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(k)}, \quad \sigma \in S_k, \tag{4.9.28}$$

are linearly independent in $\otimes^k \mathbb{C}^n$, which implies injectivity of $\tau^{\#}_{nk}$. $\square$

If $k > n$, neither the left nor the right side of (4.9.26) is easy to evaluate. We will make further comments on this in §4.10.

4.10. Decomposition of $\otimes^k \mathbb{C}^n$

Since S_k and U(n) act as a dual pair on $\otimes^k \mathbb{C}^n$, Proposition 4.8.2 is applicable. Hence $\otimes^k \mathbb{C}^n$ has a decomposition of the form (4.8.16), with S_k acting irreducibly on V_j and U(n) acting irreducibly on W_j. In this section we give a more explicit description of these factors and representations, though we refer to other sources for proofs.

To start, we recall from §4.4 irreducible representations of U(n) that were produced on subspaces of $\otimes^k \mathbb{C}^n$. Namely we have the representation $\mathcal{D}_\lambda$ of U(n) on the linear subspace of

$$(\Lambda^1 \mathbb{C}^n)^{\otimes j_1} \otimes \cdots \otimes (\Lambda^n \mathbb{C}^n)^{\otimes j_n}, \tag{4.10.1}$$

generated by the weight vector $(u_1)^{j_1} \otimes (u_1 \wedge u_2)^{j_2} \otimes \cdots \otimes (u_1 \wedge \cdots \wedge u_n)^{j_n}$, with highest weight

$$\lambda = j_1\gamma_1 + \cdots + j_n\gamma_n = (r_1, \dots, r_n), \quad r_\nu = j_\nu + \cdots + j_n. \tag{4.10.2}$$

The cases that arise for which (4.10.1) is a subspace of $\otimes^k \mathbb{C}^n$ are those for which

$$r_1 \ge \cdots \ge r_n \ge 0, \quad r_1 + \cdots + r_n = k, \quad r_\nu \in \mathbb{Z}^+. \tag{4.10.3}$$

We denote by F_{nk} the set of n-tuples $\lambda = (r_1, \dots, r_n)$ satisfying (4.10.3). Such λ is called a Young frame on k for U(n). It turns out that precisely these representations $\mathcal{D}_\lambda$ of U(n) occur in the decomposition (4.8.16) when $V = \otimes^k \mathbb{C}^n$.

We now describe the associated representation $\mathcal{S}_\lambda$ of S_k. To the Young frame $\lambda \in F_{nk}$ we associate a Young diagram as follows. The diagram consists of boxes, arranged in columns. Proceeding from left to right, there are j_n columns of length n, then j_{n-1} columns of length $n-1, \dots,$ to j_1 columns of length 1. See Figure 4.10.1. Note that the top row has length $r_1 = j_1 + \cdots + j_n$, the next row has length $r_2 = j_2 + \cdots + j_n$, etc. We number these boxes as follows. The leftmost column is numbered $1, \dots, n$ from the top down (if $j_n \neq 0$). The numbering proceeds to the next column, from top to bottom, etc. With this set-up, we define some special subsets of S_k, as follows. Let $\mathcal{F}_\lambda$ denote the Young diagram just described. We set

$$\begin{aligned} \mathcal{P}_\lambda &= \{\sigma \in S_k : \sigma \text{ preserves each row of } \mathcal{F}_\lambda\}, \\ \mathcal{Q}_\lambda &= \{\sigma \in S_k : \sigma \text{ preserves each column of } \mathcal{F}_\lambda\}. \end{aligned} \tag{4.10.4}$$

We define $p_\lambda, q_\lambda, c_\lambda \in \ell^1(S_k)$ by

$$p_\lambda = \sum_{\sigma \in \mathcal{P}_\lambda} \sigma, \quad q_\lambda = \sum_{\sigma \in \mathcal{Q}_\lambda} (\operatorname{sgn} \sigma)\sigma, \quad c_\lambda = p_\lambda * q_\lambda. \tag{4.10.5}$$

Proposition 4.10.1. *In the convolution algebra $\ell^1(S_k)$,*

$$c_\lambda * c_\lambda = \mu\, c_\lambda, \tag{4.10.6}$$

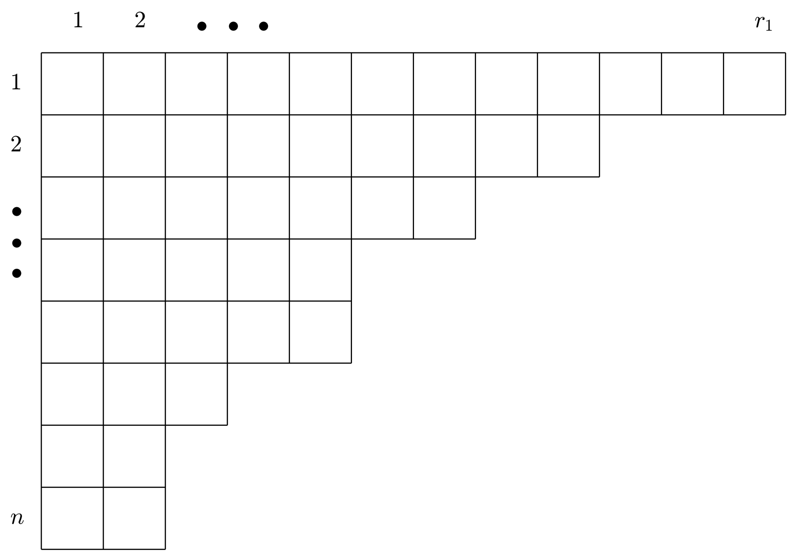

Figure 4.10.1. Young diagram

for some $\mu \in (0,\infty)$.

Thus $\mu^{-1}c_\lambda$ is an idempotent in $\ell^1(S_k)$, yielding a projection C_λ on $\ell^2(S_k)$, via right convolution. The range E_λ of C_λ in $\ell^2(S_k)$ is a linear subspace of $\ell^2(S_k)$ that is invariant under the left regular representation of S_k on $\ell^2(S_k)$. We denote the resulting representation of S_k on E_λ by $\mathcal{S}_\lambda$.

Proposition 4.10.2. *The representation $\mathcal{S}_\lambda$ of S_k on E_λ is irreducible.*

The following result is due to H. Weyl.

Proposition 4.10.3. *For the representations τ of S_k and $\otimes^k$ of* $\mathrm{U}(n)$ *on $\otimes^k\mathbb{C}^n$, we have*

$$\tau\cdot\otimes^k \approx \bigoplus_{\lambda\in F_{nk}} \mathcal{S}_\lambda\otimes\mathcal{D}_\lambda. \tag{4.10.7}$$

Complete proofs of Propositions 4.10.1–4.10.3 can be found in [**19**] and [**34**]. Let us explicitate how much of Proposition 4.10.3 has been proven in these notes. That $\tau\cdot\otimes^k$ has a decomposition of the form (4.10.7), as λ ranges over *some* set of maximal weights for irreducible representations of $\mathrm{U}(n)$, is a consequence of Proposition 4.8.2, combined with Proposition

4.9.1. That each $\lambda \in F_{nk}$ arises in this decomposition follows from our observations about (4.10.1). For a complete proof of Proposition 4.10.3, it remains to establish two things:

(i) There are no other highest weights that should appear in (4.10.7).

(ii) The irreducible representation of S_k that is paired with $\mathcal{D}_\lambda$, whose existence is established in Proposition 4.8.2, is in fact the representation $\mathcal{S}_\lambda$ described above.

Proofs of these results, which can be found on p. 251 if [**34**], rely strictly on the representation theory of S_k.

We obtain some consequences of (4.10.7) for characters. Let us set

$$\chi_\lambda^U(g) = \operatorname{Tr} \mathcal{D}_\lambda(g), \quad \chi_\lambda^S(\sigma) = \operatorname{Tr} \mathcal{S}_\lambda(\sigma). \tag{4.10.8}$$

Then (4.10.7) implies

$$\operatorname{Tr}\big(\tau(\sigma) \cdot \otimes^k g\big) = \sum_{\lambda \in F_{nk}} \chi_\lambda^S(\sigma)\, \chi_\lambda^U(g), \quad \sigma \in S_k,\ g \in \mathrm{U}(n). \tag{4.10.9}$$

In particular, taking $\sigma = e$, the identity element of S_k, gives

$$(\operatorname{Tr} g)^k = \sum_{\lambda \in F_{nk}} f^\lambda\, \chi_\lambda^U(g), \tag{4.10.10}$$

where

$$f^\lambda = \chi_\lambda^S(e) \tag{4.10.11}$$

is the dimension of the representation space for $\mathcal{S}_\lambda$. The Weyl orthogonality relations imply

$$\int_{\mathrm{U}(n)} |\operatorname{Tr} g|^{2k}\, dg = \sum_{\lambda \in F_{nk}} (f^\lambda)^2. \tag{4.10.12}$$

Recall that the left side of (4.10.12) also satisfies (4.9.26). In other words, (4.10.12) is equal to

$$\dim E_{nk} = k! - \dim \operatorname{Ker} \tau_{nk}^{\#}, \tag{4.10.13}$$

where $\tau_{nk}^{\#} : \ell^1(S_k) \to \operatorname{End}(\otimes^k \mathbb{C}^n)$ is as in (4.9.19).

We illustrate the decomposition (4.10.7) in the case $k = n = 3$. The three Young diagrams that arise in F_{33} are pictured in Figure 4.10.2. They correspond, respectively, to

$$\lambda = (1,1,1), \quad \lambda = (2,1,0), \quad \lambda = (3,0,0). \tag{4.10.14}$$

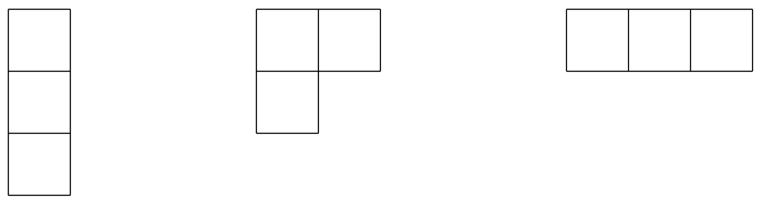

Figure 4.10.2. Young diagrams in F_{33}

The representations of S_3 so obtained are

$$\mathcal{S}_{(1,1,1)} = \mathrm{sgn}, \quad \mathcal{S}_{(2,1,0)} = \pi_S^3, \quad \mathcal{S}_{(3,0,0)} = 1. \tag{4.10.15}$$

The representation π_S^3, discussed in Lemma 2.6.1, represents S_3 as the group of symmetries of an equilateral triangle. Hence (4.10.7) leads to

$$\otimes^3 \mathbb{C}^3 \approx \Lambda^3 \mathbb{C}^3 \oplus V_8 \oplus V_8 \oplus S^3 \mathbb{C}^3, \tag{4.10.16}$$

where U(3) acts on V_8 as $\mathcal{D}_{(2,1,0)}$. As for dimensions, clearly $\dim \otimes^3 \mathbb{C}^3 = 27$ and $\dim \Lambda^3 \mathbb{C}^3 = 1$. We also have

$$\dim S^3 \mathbb{C}^3 = 10, \tag{4.10.17}$$

as a special case of the general result

$$\dim S^k \mathbb{C}^{n+1} = \binom{n+k}{n}, \tag{4.10.18}$$

as shown in §4.3. Hence

$$\dim V_8 = 8. \tag{4.10.19}$$

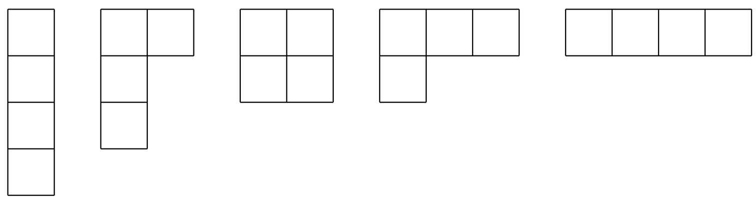

Figure 4.10.3. Young diagrams in F_{44}

In fact, more can be said about V_8. The adjoint representation of U(3) on M(3, $\mathbb{C}$) is

$$S^1 \otimes \overline{S}^1 \approx \mathcal{D}_{(0,0,0)} \oplus \mathcal{D}_{(1,0,-1)}, \tag{4.10.20}$$

by (4.7.13). Here $\mathcal{D}_{(0,0,0)}$ is a one-dimensional representation and $\mathcal{D}_{(1,0,-1)}$ is an 8-dimensional representation, acting on

$$\{A \in \mathrm{M}(3, \mathbb{C}) : \mathrm{Tr}\, A = 0\}. \tag{4.10.21}$$

On the other hand, by (4.4.5),

$$\mathcal{D}_{(2,1,0)}(g) = (\det g)\mathcal{D}_{(1,0,-1)}(g). \tag{4.10.22}$$

Let us turn to $\otimes^4\mathbb{C}^3$. The five Young diagrams that arise in F_{44} are pictured in Figure 4.10.3, but the first one does not belong to F_{34}, though the others do. They correspond, respectively, to

$$\lambda = (2, 1, 1), \quad \lambda = (2, 2, 0), \quad \lambda = (3, 1, 0), \quad \lambda = (4, 0, 0). \tag{4.10.23}$$

Recall the representations of S_4 as described in §2.6:

$$1, \quad \mathrm{sgn}, \quad \pi_S^4, \quad \pi_Q^4, \quad \pi_S^3 \circ \beta, \tag{4.10.24}$$

where $\beta : S_4 \to S_3$ is as in (2.6.20). Of these, of course

$$\mathcal{S}_{(4,0,0)} = 1. \tag{4.10.25}$$

As we have noted, the representation sgn of S_4 does not arise in $\otimes^4\mathbb{C}^3$. We claim that

$$\mathcal{S}_{(2,2,0)} \approx \pi_S^3 \circ \beta, \tag{4.10.26}$$

and that

$$\mathcal{D}_{(2,2,0)} \quad \text{acts on} \quad S^2(\Lambda^2\mathbb{C}^3). \tag{4.10.27}$$

Note that

$$\mathcal{D}_{(2,2,0)}(g) = (\det g)^2\, \overline{S}^2(g). \tag{4.10.28}$$

We ask the reader to pair π_S^4 and π_Q^4 with the two remaining weights listed in (4.10.23), and to work out explicit descriptions (or at least dimension counts) for the representation spaces for $\mathcal{D}_\lambda$ in these two cases.

Look at the formula (4.10.12) for $k = 4, n = 3$. The right side involves all the representations of S_4 but sgn, which is one dimensional, so we have

$$\int\limits_{\mathrm{U}(3)} |\operatorname{Tr} g|^8\, dg = 23. \tag{4.10.29}$$

One has a parallel treatment of $\otimes^4\mathbb{C}^n$ for $n \geq 4$. One significant difference is that the representation sgn of S_4 now appears, too. Another is that $S^2(\Lambda^2\mathbb{C}^k)$ is not irreducible, when $k \geq 4$.

4.11. The Weyl integration formula

Say G is a compact, connected Lie group, $T \subset G$ a maximal torus. The following is Weyl's integration formula:

$$\int_G f(x)\,dx = \frac{1}{w}\int_T \Bigl(\int_G f(g^{-1}kg)\,dg\Bigr)|\det(I - \operatorname{Ad} k)_{\mathfrak{g}/\mathfrak{h}}|\,dk. \tag{4.11.1}$$

Here $\mathfrak{h}$ is the Lie algebra of T, and $w = w(G)$ is a constant, which we will specify below. For most of this section we work in the context of a general compact, connected Lie group, but right at the point of specifying w we will need to refer the reader to other sources for details when G is not $\mathrm{U}(n)$.

The formula (4.11.1) arises from a study of

$$F : G \times T \longrightarrow G, \quad F(g,h) = ghg^{-1}, \tag{4.11.2}$$

and its induced action

$$\widetilde{F} : (G/T) \times T \longrightarrow G. \tag{4.11.3}$$

Since there are natural volume elements on $(G/T) \times T$ and on G, we need to compute $\det D\widetilde{F}$. Note that $DF(g,h) : T_gG \oplus T_hT \to T_{ghg^{-1}}G$; it is convenient to produce a linear map that takes $T_eG \oplus T_eT \to T_eG$. That would be

$$DL_{gh^{-1}g^{-1}}(ghg^{-1}) \circ DF(g,h) \circ (DL_g(e) \times DL_h(e)), \tag{4.11.4}$$

where $L_g(x) = gx$. Note that (4.11.4) is equal to $DG(e,e)$, where

$$\begin{aligned} G(x,z) &= L_{gh^{-1}g^{-1}} \circ F \circ (L_g \times L_h)(x,z) \\ &= gh^{-1}xhzx^{-1}g^{-1}. \end{aligned} \tag{4.11.5}$$

Note that $G(e,e) = e$; we compute

$$DG(e,e) : \mathfrak{g} \oplus \mathfrak{h} \longrightarrow \mathfrak{g}. \tag{4.11.6}$$

First, with $Z \in \mathfrak{h}$, $z(t)$ a curve in T such that $z(0) = e$, $z'(0) = Z$, we have

$$\begin{aligned} D_2G(e,e)Z &= \frac{d}{dt}\, gz(t)g^{-1}\big|_{t=0} \\ &= \operatorname{Ad} g(Z), \end{aligned} \tag{4.11.7}$$

the last identity following from (3.4.2), or alternatively from (3.4.9). Next, with $X \in \mathfrak{g}$, $x(t)$ a curve in G such that $x(0) = e$, $x'(0) = X$, we have

$$\begin{aligned} D_1G(e,e)X &= \frac{d}{dt}\, gh^{-1}x(t)hx(t)^{-1}g^{-1}\big|_{t=0} \\ &= \operatorname{Ad} g\, DK(e)X, \end{aligned} \tag{4.11.8}$$

where

$$K(x) = h^{-1}xhx^{-1}, \tag{4.11.9}$$

so

$$
\begin{aligned}
DK(e)X &= \frac{d}{dt}\, h^{-1}x(t)hx(t)^{-1}\big|_{t=0} \\
&= h^{-1}Xh - X \\
&= (\operatorname{Ad} h^{-1} - I)X.
\end{aligned}
\tag{4.11.10}
$$

(Here we take $G \subset \operatorname{End}(\mathbb{C}^m)$, to simplify the calculation.) Putting together (4.11.7), (4.11.8), and (4.11.10), we have

$$
DG(e,e)(X,Z) = \operatorname{Ad} g(\operatorname{Ad} h^{-1} - I)X + \operatorname{Ad} g\, Z. \tag{4.11.11}
$$

Now we can take $X \in \mathfrak{g}/\mathfrak{h}$. Thus we have

$$
\det D\widetilde{F}(g,h) = \det(\operatorname{Ad} h^{-1} - I)_{\mathfrak{g}/\mathfrak{h}} = \det(I - \operatorname{Ad} h)_{\mathfrak{g}/\mathfrak{h}}. \tag{4.11.12}
$$

The formula (4.11.1) is now a consequence of the following assertion:

Lemma 4.11.1. *The map $\widetilde{F}$ in (4.11.3) is onto, and there is an integer $w = w(G)$ and an open dense set $\mathcal{O} \subset G$, whose complement has measure zero, such that*

$$
\widetilde{F}^{-1}(g) \subset (G/T) \times T \quad \textit{has } w \textit{ elements}, \quad \forall\, g \in \mathcal{O}.
$$

In the case $G = \mathrm{U}(n)$, we take T to be the set $\mathbb{T}$ of diagonal matrices, with diagonal entries in $S^1 \subset \mathbb{C}$, as in (4.2.4). The surjectivity of $\widetilde{F}$ is equivalent to the statement that every unitary matrix yields an orthonormal basis of eigenvectors. If $g \in \mathrm{U}(n)$ has distinct eigenvalues, then the eigenspaces are all 1-dimensional, and the diagonalized form is determined up to ordering of the eigenvalues, so such a matrix has $n!$ pre-images in $(G/\mathbb{T}) \times \mathbb{T}$.

The reader can verify Lemma 4.11.1 and determine $w(G)$ when $G = \mathrm{SO}(n)$. For general compact, connected G, $w(G)$ is the order of a finite group called the Weyl group. See [**16**], [**34**], or another source for a treatment of the general case.

We give an explicit formula for the right side of (4.11.12) when $G = \mathrm{U}(n)$. In such a case, $\mathfrak{g}_{\mathbb{C}} = \operatorname{End}(\mathbb{C}^n)$. As in §4.2, let e_{jk} be the matrix with 1 at row j, column k, 0 elsewhere, and set $e_j = ie_{jj}$. Then $\mathfrak{h}$ is the real linear span of $\{e_j : 1 \le j \le n\}$, and

$$
H = \sum t_j e_j \Longrightarrow [H, e_{jk}] = i(t_j - t_k)e_{jk}. \tag{4.11.13}
$$

Using this, we have that, when $G = \mathrm{U}(n)$, $h = \operatorname{diag}(e^{i\theta_1}, \dots, e^{i\theta_n}) \in T$,

$$
\operatorname{Ad} h(e_{jk}) = e^{i(\theta_j - \theta_k)}\, e_{jk}. \tag{4.11.14}
$$

Thus

$$\det(I - \operatorname{Ad} h)_{\mathfrak{g}/\mathfrak{h}} = \det(I - \operatorname{Ad} h)_{\mathfrak{g}_{\mathbb{C}}/\mathfrak{h}_{\mathbb{C}}} = \prod_{j \neq k} (1 - e^{i(\theta_j - \theta_k)}) = \prod_{j \neq k} e^{-i\theta_k}(e^{i\theta_k} - e^{i\theta_j}), \tag{4.11.15}$$

and hence

$$|\det(I - \operatorname{Ad} h)_{\mathfrak{g}/\mathfrak{h}}| = \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2. \tag{4.11.16}$$

We explicitly specialize (4.11.1) to the case where $G = \mathrm{U}(n)$ and f is a central function, i.e., $f(g^{-1}hg) = f(h)$ for all $g, h \in \mathrm{U}(n)$.

Corollary 4.11.2. *If $f : \mathrm{U}(n) \to \mathbb{C}$ is a central function, then*

$$\int_{\mathrm{U}(n)} f(g)\, dg = \frac{1}{(2\pi)^n n!} \int_{\mathbb{T}^n} f(D(\theta)) J(\theta)\, d\theta_1 \cdots d\theta_n, \tag{4.11.17}$$

where $D(\theta)$ is the diagonal matrix with diagonal entries $e^{i\theta_1}, \dots, e^{i\theta_n}$, and

$$J(\theta) = \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2. \tag{4.11.18}$$

Here we take $\mathbb{T}^n = (\mathbb{R}/2\pi\mathbb{Z})^n$. We mention another way of writing $J(\theta)$, namely as

$$J(\theta) = A(\theta)\overline{A(\theta)}, \quad A(\theta) = \prod_{j<k} \bigl(1 - e^{-i\omega_{jk}(\theta)}\bigr), \tag{4.11.19}$$

where we regard $\theta \in \mathbb{R}^n \approx \mathfrak{h}$, and we take $\omega_{jk} \in \mathfrak{h}'$ as in (19.12).

Here is another way of writing $J(\theta)$, which is useful. Set $e^{i\theta_j} = \zeta_j$. Then

$$J(\theta) = V(\zeta)V(\overline{\zeta}), \quad V(\zeta) = \prod_{j<k} (\zeta_k - \zeta_j). \tag{4.11.20}$$

Now $V(\zeta)$ is a Vandermonde determinant:

$$V(\zeta) = \det \begin{pmatrix} 1 & \cdots & 1 \\ \zeta_1 & \cdots & \zeta_n \\ \vdots & & \vdots \\ \zeta_1^{n-1} & \cdots & \zeta_n^{n-1} \end{pmatrix}. \tag{4.11.21}$$

Hence

$$V(\zeta) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, \zeta_1^{\sigma(1)-1} \cdots \zeta_n^{\sigma(n)-1}. \tag{4.11.22}$$

Now $\overline{\zeta}_j = \zeta_j^{-1}$ for $\zeta_j \in S^1$, so

$$J(\theta) = \sum_{\sigma,\tau \in S_n} (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)\, \zeta_1^{\sigma(1)-\tau(1)} \cdots \zeta_n^{\sigma(n)-\tau(n)}. \tag{4.11.23}$$

Note that

$$\begin{aligned} (2\pi)^{-n} \int_{\mathbb{T}^n} J(\theta)\, d\theta &= \text{constant term in } (4.11.23) \\ &= \sum \bigl\{ (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau) : \sigma = \tau \in S_n \bigr\} \\ &= n!, \end{aligned} \tag{4.11.24}$$

which gives a check on the coefficient on the right side of (4.11.17).

4.12. The character formula

Here we calculate the character χ_λ of the irreducible unitary representation $\mathcal{D}_\lambda$ of $\mathrm{U}(n)$ with highest weight λ. We know χ_λ is a central function, so it suffices to calculate $\chi_\lambda(h)$ for $h \in \mathbb{T}$, the group of diagonal matrices in $U(n)$. Say $h = D(\theta) = \mathrm{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$. Recall from §4.2 that the representation space V of $\mathcal{D}_\lambda$ has a decomposition

$$V = \bigoplus_\mu V_\mu \tag{4.12.1}$$

into spaces of weight vectors. It follows from (4.2.10) that

$$\chi_\lambda(D(\theta)) = \sum m_\mu \, e^{i\mu(\theta)}, \quad m_\mu = \dim V_\mu. \tag{4.12.2}$$

Our goal is to get a more explicit formula for this object. To be sure, (4.12.2) as it stands will be a useful tool.

To begin, recall the role of S_n as a group of automorphisms of $\mathrm{U}(n)$, as described in Lemma 4.4.2 and its proof; $\sigma \in S_n$ acts on $\mathrm{U}(n)$ via conjugation by E_σ. This action leaves $\mathbb{T}$ invariant, so S_n acts on $\mathfrak{h}$. We denote its adjoint action on $\mathfrak{h}'$ by $\mu \mapsto \sigma \cdot \mu$; with $\mathfrak{h}' \approx \mathbb{R}^n$ via the usual basis,

$$\sigma \cdot \mu = (\mu_{\sigma(1)}, \dots, \mu_{\sigma(n)}). \tag{4.12.3}$$

In view of (4.4.4), we have the identity

$$m_{\sigma \cdot \mu} = m_\mu, \quad \forall\, \sigma \in S_n. \tag{4.12.4}$$

Another identity arises by rewriting the identity

$$\int\limits_{\mathrm{U}(n)} \chi_\lambda(g) \, \overline{\chi_\lambda(g)} \, dg = 1, \tag{4.12.5}$$

using the Weyl integration formula (4.11.17):

$$(2\pi)^{-n} \int\limits_{\mathbb{T}^n} A(\theta) \chi_\lambda(D(\theta)) \, \overline{A(\theta)} \, \overline{\chi_\lambda(D(\theta))} \, d\theta = n! \tag{4.12.6}$$

To exploit this, we consider

$$\varphi(\theta) = A(\theta) \, \chi_\lambda(D(\theta)). \tag{4.12.7}$$

Recall from (4.11.19) that

$$A(\theta) = \prod_{j<k} \bigl(1 - e^{-i\omega_{jk}(\theta)}\bigr). \tag{4.12.8}$$

In particular, we have a finite sum

$$\varphi(\theta) = \sum c_\mu \, e^{i\mu(\theta)}, \quad c_\mu \in \mathbb{Z}, \tag{4.12.9}$$

and the identity (4.12.6) implies

$$\sum c_\mu^2 = n! \tag{4.12.10}$$

As we will see, this will help place strong constraints on the coefficients c_μ, particularly in concert with the following observation.

Lemma 4.12.1. *For the highest weight λ of $\mathcal{D}_\lambda$, we have*

$$c_\lambda = m_\lambda = 1. \tag{4.12.11}$$

Proof. Since the elements $\omega_{jk} \in \mathfrak{h}'$ are all positive for $j < k$ (with respect to the ordering defined in §4.2), it is clear from (4.12.7)–(4.12.8) that $c_\lambda = m_\lambda$. That $m_\lambda = 1$ follows from Proposition 4.2.4. $\square$

Further progress in understanding $\varphi(\theta)$ comes from looking at how it behaves under the S_n action on $\mathbb{T}^n$. Clearly $\chi_\lambda(D(\theta))$ is invariant under this action, so we need to see how $A(\theta)$ behaves under the S_n action. We have

$$A(\sigma^t \cdot \theta) = \prod_{j<k} \big(1 - e^{-i\omega_{\sigma(j)\sigma(k)}(\theta)}\big). \tag{4.12.12}$$

We can break up this product into two products, one over $\{(j,k) : j < k \text{ and } \sigma(j) < \sigma(k)\}$ and the second over $\{(j,k) : j < k \text{ and } \sigma(j) > \sigma(k)\}$. Write the factors in the second product as

$$1 - e^{-i\omega_{\sigma(j)\sigma(k)}(\theta)} = -e^{-i\omega_{\sigma(j)\sigma(k)}(\theta)}\big(1 - e^{-i\omega_{\sigma(k)\sigma(j)}(\theta)}\big). \tag{4.12.13}$$

Recombining the two products gives

$$A(\sigma^t \cdot \theta) = \alpha\, e^{-i\beta}\, A(\theta), \tag{4.12.14}$$

with

$$\alpha = \prod_{\{(j,k):j<k,\sigma(j)>\sigma(k)\}} (-1), \quad \beta = \sum_{\{(j,k):j<k,\sigma(j)>\sigma(k)\}} \omega_{\sigma(j)\sigma(k)}(\theta). \tag{4.12.15}$$

A calculation gives $\alpha = \operatorname{sgn}\sigma$. Also, if we set

$$\rho = \frac{1}{2}\sum_{j<k} \omega_{jk} \in \mathfrak{h}', \tag{4.12.16}$$

then

$$\beta = \frac{1}{2}\sum_{j<k} \big[\omega_{\sigma(j)\sigma(k)} - \omega_{jk}\big] = \sigma \cdot \rho - \rho. \tag{4.12.17}$$

Hence (4.12.14) becomes

$$A(\sigma^t \cdot \theta) = (\operatorname{sgn}\sigma)\, e^{i(\rho(\theta) - \sigma\cdot\rho(\theta))}\, A(\theta). \tag{4.12.18}$$

In view of the conjugation invariance of χ_λ and (4.4.2), this gives

$$\varphi(\sigma^t \cdot \theta) = (\operatorname{sgn} \sigma)\, e^{i(\rho(\theta) - \sigma \cdot \rho(\theta))}\, \varphi(\theta). \tag{4.12.19}$$

Equivalently, the coefficients c_μ in (4.12.9) satisfy

$$c_{\sigma\cdot\mu+\sigma\cdot\rho-\rho} = (\operatorname{sgn}\sigma) c_\mu. \tag{4.12.20}$$

We are led to consider a "shifted" action of S_n on $\mathfrak{h}'$:

$$\tilde{\sigma}(\mu) = \sigma \cdot \mu + \sigma \cdot \rho - \rho, \tag{4.12.21}$$

where $\sigma \cdot \mu = (\mu_{\sigma(1)}, \dots, \mu_{\sigma(n)})$. A calculation shows that this is a group action, i.e., $\tilde{\sigma} \circ \tilde{\tau}(\mu) = \widetilde{\sigma\tau}(\mu)$ for $\sigma, \tau \in S_n$. (It is not a linear action, but rather an action by affine transformations.) The following result will reveal a great deal about the coefficients c_μ.

Lemma 4.12.2. *The orbit of the highest weight λ under the action of S_n given by* (4.12.21) *has $n!$ elements.*

Proof. Consider $\sigma \in S_n$ such that $\tilde{\sigma}(\lambda) = \lambda$. This implies

$$\sigma \cdot \lambda + \sigma \cdot \rho = \lambda + \rho. \tag{4.12.22}$$

Now $\lambda = (\lambda_1, \dots, \lambda_n)$ with $\lambda_1 \geq \cdots \geq \lambda_n$ and

$$\rho = (\rho_1, \dots, \rho_n) = \frac{1}{2}(n-1, n-3, \dots, 3-n, 1-n), \tag{4.12.23}$$

so $\rho_1 > \cdots > \rho_n$, and hence $\lambda + \rho = (\xi_1, \dots, \xi_n)$ with $\xi_1 > \cdots > \xi_n$. Thus (4.12.2) can hold only if σ is the identity element of S_n. $\square$

Thus taking $\mu = \lambda$ in (4.12.20) gives $n!$ coefficients c_μ that are equal to ± 1. In view of (4.12.10), these are all the nonzero coefficients in (4.12.9). We have established

$$A(\theta)\chi_\lambda(D(\theta)) = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma)\, e^{i(\sigma\cdot\lambda(\theta) + \sigma\cdot\rho(\theta) - \rho(\theta))}. \tag{4.12.24}$$

In view of the formula (4.12.8) for $A(\theta)$, this gives a rather explicit formula for $\chi_\lambda(D(\theta))$.

Note that if we take the trivial representation, with highest weight $\lambda = 0$, the character is $\equiv 1$, and (4.12.24) yields

$$A(\theta) = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma)\, e^{i(\sigma\cdot\rho(\theta) - \rho(\theta))}, \tag{4.12.25}$$

which one might also try to derive directly from (4.12.8). This suggests writing the character formula in terms of

$$A_\mu(\theta) = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma)\, e^{i\sigma\cdot\mu(\theta)}. \tag{4.12.26}$$

Proposition 4.12.3. *The irreducible representation of* $\mathrm{U}(n)$ *with highest weight* λ *has character satisfying*

$$\chi_\lambda(D(\theta)) = \frac{A_{\lambda+\rho}(\theta)}{A_\rho(\theta)}. \tag{4.12.27}$$

Proof. Multiplying both sides of (28.34) by $e^{i\rho(\theta)}$, one obtains $A_{\lambda+\rho}(\theta)$ on the right and $A(\theta)$ is turned into $A_\rho(\theta)$ on the left. □

Remark. The entries of $\rho = (\rho_1, \dots, \rho_n)$ might be half-integers, rather than integers (if n is odd), so then neither the numerator nor the denominator in the right side of (4.12.27) is periodic of period $2\pi\mathbb{Z}$, but the quotient is. In any case the numerator and denominator have period $4\pi\mathbb{Z}$ in θ.

We can represent $A_\mu(\theta)$ as a product in some special cases. First note that since $A_\rho(\theta)$ is obtained by multiplying (4.12.25) by $e^{i\rho(\theta)}$, the product (4.12.8) for $A(\theta)$ yields

$$A_\rho(\theta) = \prod_{j<k} \bigl(e^{i\omega_{jk}(\theta)/2} - e^{-i\omega_{jk}(\theta)/2}\bigr). \tag{4.12.28}$$

To proceed, note that our choice of basis for $\mathfrak{h}$ gives $\mathfrak{h} \approx \mathbb{R}^n$ and also $\mathfrak{h}' \approx \mathbb{R}^n$, and hence $\mathfrak{h}' \approx \mathfrak{h}$. If we so identify $\mathfrak{h}$ and $\mathfrak{h}'$, we see from (4.12.26) that $A_\mu(\xi) = A_\xi(\mu)$, for $\mu, \xi \in \mathbb{R}^n$, and furthermore $A_\mu(t\xi) = A_\xi(t\mu)$. Hence

$$A_\mu(t\rho) = A_\rho(t\mu) = \prod_{j<k} \bigl(e^{it\langle\omega_{jk},\mu\rangle/2} - e^{-it\langle\omega_{jk},\mu\rangle/2}\bigr). \tag{4.12.29}$$

Here we have used (4.12.28) and replaced the pairing $\mu(\xi)$ of $\mathfrak{h}'$ and $\mathfrak{h}$ by $\langle\mu, \xi\rangle$, the standard inner product in $\mathbb{R}^n$. Using (4.12.29) we can prove the following dimension formula.

Proposition 4.12.4. *The irreducible representation of* $\mathrm{U}(n)$ *with highest weight* λ *acts on a space* $V(\lambda)$ *whose dimension is*

$$\dim V(\lambda) = d_\lambda = \prod_{j<k} \frac{\langle\omega_{jk}, \lambda+\rho\rangle}{\langle\omega_{jk}, \rho\rangle}. \tag{4.12.30}$$

Proof. Clearly $d_\lambda = \chi_\lambda(I)$, and hence

$$d_\lambda = \lim_{t\to 0} \chi_\lambda(D(t\rho)) = \lim_{t\to 0} \frac{A_{\lambda+\rho}(t\rho)}{A_\rho(t\rho)}, \tag{4.12.31}$$

by (4.12.27). Now we can apply (4.12.29) to obtain

$$d_\lambda = \lim_{t\to 0} \prod_{j<k} \frac{\sin t\langle\omega_{jk}, \lambda+\rho\rangle/2}{\sin t\langle\omega_{jk}, \rho\rangle/2}, \tag{4.12.32}$$

which yields (4.12.30), granted that $\prod_{j<k} \langle \omega_{jk}, \rho \rangle \neq 0$. In fact,

$$j < k \Longrightarrow \langle \omega_{jk}, \rho \rangle = \rho_j - \rho_k = k - j, \tag{4.12.33}$$

by (4.12.23), which leads to the following explicit formula for the denominator that arises in (4.12.30):

$$\prod_{j<k} \langle \omega_{jk}, \rho \rangle = \prod_{1 \leq j < k \leq n} (k - j) = \prod_{\ell=1}^{n-1} \ell!. \tag{4.12.34}$$

$\square$

4.13. Examples of characters

Let us first specialize the character formula of §4.12 to the case of U(2). We consider $\chi_\lambda(D(\theta))$ with $\lambda = (\lambda_1, \lambda_2),\ \lambda_1 \ge \lambda_2,\ \lambda_\nu \in \mathbb{Z}$. In this case, $\omega_{12} = (1, -1),\ \rho = (1/2, -1/2)$, and hence $A_\mu(\theta)$, defined by (4.12.25), takes the form

$$A_\mu(\theta) = e^{i(\mu_1\theta_1 + \mu_2\theta_2)} - e^{i(\mu_2\theta_1 + \mu_1\theta_2)}. \tag{4.13.1}$$

In particular, the denominator $A_\rho(\theta)$ in (4.12.27) becomes

$$A_\rho(\theta) = e^{i(\theta_1 - \theta_2)/2} - e^{i(\theta_2 - \theta_1)/2} = 2i\, \sin \frac{\theta_1 - \theta_2}{2}. \tag{4.13.2}$$

To evaluate the numerator in (4.12.27), we take $\mu = \lambda + \rho = (\lambda_1 + 1/2, \lambda_2 - 1/2)$ in (4.13.1). For simplicity, let us take

$$\lambda = (k, 0). \tag{4.13.3}$$

Then

$$A_{\lambda+\rho}(\theta) = e^{i(k+1/2)\theta_1 - i\theta_2/2} - e^{i(k+1/2)\theta_2 - i\theta_1/2}. \tag{4.13.4}$$

If we also take $\theta_2 = -\theta_1$, so $D(\theta) \in \mathrm{SU}(2)$, then

$$A_{\lambda+\rho}(\theta) = e^{i(k+1)\theta_1} - e^{-i(k+1)\theta_1}, \tag{4.13.5}$$

and hence the character formula (4.12.27) gives

$$\chi_{(k,0)}(D(\theta, -\theta)) = \frac{\sin(k + 1/2)\theta}{\sin \theta}. \tag{4.13.6}$$

Taking the limit $\theta \to 0$ gives the dimension formula

$$d_{(k,0)} = k + 1, \tag{4.13.7}$$

familiar from our previous discussion of the representations of SU(2).

Note that it is a direct consequence of (4.1.9), (4.1.14), (4.1.21), and (4.1.22) that

$$\chi_{(k,0)}(D(\theta, -\theta)) = \sum_{j=-k}^{k} e^{ij\theta}, \tag{4.13.8}$$

which sums to the right side of (4.13.6).

Let us generalize these calculations to the representations $S^k = \mathcal{D}_{(k,0,\dots,0)}$ of $\mathrm{U}(n)$. Using (4.3.1)–(4.3.8), we see that

$$\mathrm{Tr}\, S^k(D(\theta)) = \sum_{|\alpha|=k} e^{i\langle \alpha, \theta \rangle}, \tag{4.13.9}$$

where in this sum $\alpha = (\alpha_1, \dots, \alpha_n)$ with $\alpha_\nu \in \mathbb{Z}^+$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$. The character formula (4.12.27) then asserts that

$$A_\rho(\theta)\ \mathrm{Tr}\, S^k(D(\theta)) = \sum_{|\alpha|=k} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, e^{i\langle \sigma\cdot\rho+\alpha,\theta\rangle} \tag{4.13.10}$$

is equal to

$$A_{(k,0,\dots,0)+\rho}(\theta) = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, e^{i\langle \sigma\cdot\rho+\sigma\cdot(k,0,\dots,0),\theta\rangle}. \tag{4.13.11}$$

To check this for $n = 2$, note that (4.13.10) becomes

$$\sum_{\alpha_1+\alpha_2=k} \Bigl(e^{i(\alpha_1+1/2)\theta_1+i(\alpha_2-1/2)\theta_2} - e^{i(\alpha_1-1/2)\theta_1+i(\alpha_2+1/2)\theta_2} \Bigr). \tag{4.13.12}$$

Note also that $(\alpha_1 + 1/2) + (\alpha_2 - 1/2) = k$ and $(\alpha_1 - 1/2) + (\alpha_2 + 1/2) = k$. Hence we get cancellation of all terms except the first part of the sum at $\alpha = (k, 0)$ and the last part of the sum at $\alpha = (0, k)$. Thus (4.13.12) telescopes to the right side of (4.13.4), verifying identity of (4.13.10) and (4.13.11) when $n = 2$.

Note that if we reverse the order of summation in (4.13.10) and sum over $\mathcal{E} = \{(\sigma, \alpha) : \alpha = \sigma \cdot (k, 0, \dots, 0)\}$, we get (4.13.11). Also note that all the frequencies that arise in (4.13.10) with $\alpha = \sigma \cdot (k, 0, \dots, 0)$ are distinct from all the frequencies that arise with $\alpha \neq \sigma \cdot (k, 0, \dots, 0)$. Now we can deduce from (4.12.10), an analogue of which also holds for $A_\rho(\theta)\chi_\lambda(\theta)$, that the rest of the sum in (4.13.10) vanishes, due to cancellations. The reader is invited to find a more direct demonstration of this vanishing.

Next consider the representation Λ^ℓ of U(n) on $\Lambda^\ell\mathbb{C}^n$, discussed in (4.3.9)–(4.3.14). We see that $\Lambda^\ell D(\theta)$ has eigenvalues $e^{i(\theta_{j_1}+\cdots+\theta_{j_\ell})}$, for general $j_1 < \cdots < j_\ell$, with j_ν running from 1 to n. Hence

$$\mathrm{Tr}\, \Lambda^\ell D(\theta) = \sigma_\ell(e^{i\theta_1}, \dots, e^{i\theta_n}), \tag{4.13.13}$$

where σ_ℓ is the ℓth elementary symmetric polynomial. We can write this as

$$\begin{aligned}\mathrm{Tr}\, \Lambda^\ell(D(\theta)) &= \frac{1}{\ell!(n-\ell)!} \sum_{\sigma\in S_n} e^{i(\theta_{\sigma(1)}+\cdots+\theta_{\sigma(\ell)})} \\ &= \frac{1}{n!}\binom{n}{\ell} \sum_{\sigma\in S_n} e^{i\langle \sigma\cdot\gamma_\ell,\theta\rangle},\end{aligned} \tag{4.13.14}$$

where $\gamma_\ell = (1, \dots, 1, 0, \dots, 0)$, with ℓ ones, is the highest weight of Λ^ℓ. Recall that $\Lambda^\ell = \mathcal{D}_{(1,\dots,1,0,\dots,0)}$. According to the character formula (4.12.27), the quantity

$$A_\rho(\theta)\ \mathrm{Tr}\,\Lambda^\ell(D(\theta)) = \frac{1}{n!}\binom{n}{\ell}\sum_{\sigma,\tau\in S_n}(\operatorname{sgn}\tau)e^{i\langle\tau\cdot\rho+\sigma\cdot\gamma_\ell,\theta\rangle} \tag{4.13.15}$$

is equal to

$$A_{\gamma_\ell+\rho}(\theta) = \sum_{\tau\in S_n}(\operatorname{sgn}\tau)e^{i\langle\tau\cdot(\rho+\gamma_\ell),\theta\rangle}. \tag{4.13.16}$$

Note that by taking $\sigma\mapsto\tau\sigma$, we can rewrite the right side of (4.13.15) as

$$\frac{1}{n!}\binom{n}{\ell}\sum_{\sigma,\tau\in S_n}(\operatorname{sgn}\tau)e^{i\langle\tau\cdot(\rho+\sigma\cdot\gamma_\ell),\theta\rangle}. \tag{4.13.17}$$

The sum in (4.13.17) over $\{(\sigma,\tau):\sigma \text{ fixes } \gamma_\ell\}$ is equal to (4.13.16). An argument involving (4.12.10), similar to that made above comparing (4.13.10) and (4.13.11), can be used to show that all the other terms in (4.13.17) must cancel out. Again the reader is invited to find a direct demonstration of this cancellation.

4.14. Duality and the Frobenius character formula

We take a further look at the decomposition of the action of $S_k \times \mathrm{U}(n)$ on $\otimes^k \mathbb{C}^n$ given by (4.10.7), i.e.,

$$\tau \cdot \otimes^k = \bigoplus_{\lambda \in F_{nk}} \mathcal{S}_\lambda \otimes \mathcal{D}_\lambda, \tag{4.14.1}$$

and its implication for characters,

$$\mathrm{Tr}(\tau(\sigma) \cdot \otimes^k g) = \sum_{\lambda \in F_{nk}} \chi_\lambda^S(\sigma)\, \chi_\lambda(g), \quad \sigma \in S_k,\ g \in \mathrm{U}(n). \tag{4.14.2}$$

Here $\chi_\lambda(g) = \mathrm{Tr}\, \mathcal{D}_\lambda$ is the character for which (4.12.27) furnishes a formula. Our goal here is to produce a formula for $\chi_\lambda^S(\sigma)$. To begin, we have the following explicit formula for the left side of (4.14.2).

Lemma 4.14.1. *Suppose $\sigma \in S_k$ has cycles of length $\ell_1, \dots, \ell_r$ (so $\ell_1 + \cdots + \ell_r = k$). Then*

$$\mathrm{Tr}(\tau(\sigma) \cdot \otimes^k g) = \prod_{\nu=1}^r \mathrm{Tr}(g^{\ell_\nu}). \tag{4.14.3}$$

Proof. Since the left side of (4.14.3) is invariant under conjugacy of σ and of g, it suffices to treat the case when

$$\sigma = (1 \cdots \ell_1)(\ell_1 + 1 \cdots \ell_1 + \ell_2) \cdots (k - \ell_r + 1 \cdots k),$$

and when g acts on the standard basis $\{u_1, \dots, u_n\}$ of $\mathbb{C}^n$ by $g u_i = \zeta_i u_i$. Then the left side of (4.14.3) is given by

$$\begin{aligned} &\sum_{1 \le j_1, \dots, j_k \le n} \langle \tau(\sigma) \cdot \otimes^k g(u_{j_1} \otimes \cdots \otimes u_{j_k}), u_{j_1} \otimes \cdots \otimes u_{j_k} \rangle \\ &\qquad = \sum \langle g\, u_{j_{\sigma(1)}}, u_{j_1} \rangle \cdots \langle g\, u_{j_{\sigma(k)}}, u_{j_k} \rangle \\ &\qquad = \sum \zeta_{j_1} \cdots \zeta_{j_k}\, \delta_{j_1 j_{\sigma(1)}} \cdots \delta_{j_k j_{\sigma(k)}}. \end{aligned} \tag{4.14.4}$$

Meanwhile the right side of (4.14.3) is equal to

$$\sum_{j_1, \dots, j_r} \zeta_{j_1}^{\ell_1} \cdots \zeta_{j_r}^{\ell_r}. \tag{4.14.5}$$

Now under our stated hypotheses on σ, the nonzero terms in the last sum in (4.14.4) are indexed by $(j_1, \dots, j_k)$ for which

$$j_1 = \cdots = j_{\ell_1},\ j_{\ell_1+1} = \cdots = j_{\ell_1+\ell_2}, \cdots, j_{k-\ell_r+1} = \cdots = j_k,$$

so (4.14.4) does indeed coincide with (4.14.5). $\square$

We will denote the quantity (4.14.3) by $\Xi(\sigma, g)$, so (4.14.2) reads

$$\Xi(\sigma, g) = \sum_{\lambda \in F_{nk}} \chi_\lambda^S(\sigma)\chi_\lambda(g). \tag{4.14.6}$$

Given $\mu \in F_{nk}$, we can multiply both sides of (4.14.6) by $\overline{\chi}_\mu^S(\sigma)$ and average over $\sigma \in S_k$, obtaining (upon switching notation from μ to λ)

$$\chi_\lambda(g) = \frac{1}{k!} \sum_{\sigma \in S_k} \Xi(\sigma, g)\overline{\chi}_\lambda^S(\sigma). \tag{4.14.7}$$

Similarly,

$$\chi_\lambda^S(\sigma) = \int\limits_{\mathrm{U}(n)} \Xi(\sigma, g)\overline{\chi}_\lambda(g)\, dg. \tag{4.14.8}$$

Another way to write $\Xi(\sigma, g)$ is as follows. Set $P_j(\zeta) = \zeta_1^j + \cdots + \zeta_n^j$ and for $\sigma \in S_k$ set

$$P_\sigma(\zeta) = P_{\ell_1}(\zeta) \cdots P_{\ell_r}(\zeta) \tag{4.14.9}$$

if σ consists of cycles of length $\ell_1, \dots, \ell_r$ (so $\ell_1 + \cdots + \ell_r = k$). Then

$$\Xi(\sigma, g) = P_\sigma(\zeta) \tag{4.14.10}$$

provided the eigenvalues of g are $\zeta_1, \dots, \zeta_n$.

If we insert the character formula (4.12.27) into (4.14.8), we can derive the Frobenius character formula for $\chi_\lambda^S(\sigma)$. Let us proceed. Fix $k \in \mathbb{Z}^+$ and consider $\lambda = (\lambda_1, \dots, \lambda_n)$ with $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, $\lambda_\nu \in \mathbb{Z}^+$, and $\lambda_1 + \cdots + \lambda_n = k$, i.e., $\lambda \in F_{nk}$. We will write (4.14.8) as an integral over $\mathbb{T}^n$, using the Weyl formula (4.11.17). Note that, in place of $J(\theta) = A(\theta)\overline{A(\theta)}$ as in (4.11.19), we can write

$$J(\theta) = A_\rho(\theta)\overline{A_\rho(\theta)}, \tag{4.14.11}$$

with $A_\rho(\theta)$ the denominator in (4.12.27). Hence (4.14.8) yields

$$\chi_\lambda^S(\sigma) = \frac{1}{(2\pi)^n n!} \int\limits_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\, \overline{A_{\lambda+\rho}(\theta)}\, A_\rho(\theta)\, d\theta, \tag{4.14.12}$$

Next, we write

$$A_\rho(\theta) = e^{i\langle \rho, \theta\rangle} \Bigl(\prod_{j<k} e^{-i\theta_j}\Bigr) \Delta(\theta), \quad \Delta(\theta) = \prod_{j<k} (e^{i\theta_j} - e^{i\theta_k}), \tag{4.14.13}$$

and note that

$$\sum_{j<k} \theta_j = (n-1)\theta_1 + (n-2)\theta_2 + \cdots + \theta_{n-1} = \langle \Gamma, \theta\rangle, \tag{4.14.14}$$

with

$$\Gamma = (n-1, n-2, \dots, 1, 0). \tag{4.14.15}$$

Hence

$$A_\rho(\theta) = e^{i\langle \rho-\Gamma,\theta\rangle}\,\Delta(\theta) \qquad (4.14.16)$$
$$= e^{i(n-1)(\theta_1+\cdots+\theta_n)/2}\,\Delta(\theta).$$

We have (via (4.12.26) with $\mu = \lambda+\rho$)

$$\chi^S_\lambda(\sigma) = \frac{1}{n!}\sum_{\tau\in S_n} (\operatorname{sgn}\tau) I^\tau_\lambda(\sigma), \qquad (4.14.17)$$

with

$$I^\tau_\lambda(\sigma) = \frac{1}{(2\pi)^n}\int_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\Delta(\theta)\, e^{i\langle\rho-\Gamma,\theta\rangle} e^{-i\langle\tau\cdot(\lambda+\rho),\theta\rangle}\, d\theta. \qquad (4.14.18)$$

Now $\Xi(\sigma, D(\theta))$ and $e^{i\langle\rho-\Gamma,\theta\rangle}$ are symmetric in $(\theta_1,\dots,\theta_n)$ (the latter by (4.14.16)), while applying a permutation τ to θ multiplies $\Delta(\theta)$ by $\operatorname{sgn}\tau$. Hence we have

$$I^\tau_\lambda(\sigma) = \frac{\operatorname{sgn}\tau}{(2\pi)^n}\int_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\Delta(\theta)\, e^{i\langle\rho-\Gamma,\theta\rangle} e^{-i\langle\lambda+\rho,\theta\rangle}\, d\theta \qquad (4.14.19)$$
$$= \frac{\operatorname{sgn}\tau}{(2\pi)^n}\int_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\Delta(\theta) e^{-i\langle\lambda+\Gamma,\theta\rangle}\, d\theta.$$

Plugging this into (4.14.17), we have Schur's formula:

Proposition 4.14.2. *Given $\lambda \in F_{nk}$, the associated representation $\mathcal{S}_\lambda$ of S_k has character*

$$\chi^S_\lambda(\sigma) = (2\pi)^{-n}\int_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\Delta(\theta) e^{-i\langle\lambda+\Gamma,\theta\rangle}\, d\theta. \qquad (4.14.20)$$

Equivalently, $\chi^S_\lambda(\sigma)$ is equal to the coefficient of $\zeta_1^{\ell_1}\cdots\zeta_n^{\ell_n}$ in

$$P_\sigma(\zeta)\prod_{j<k}(\zeta_j-\zeta_k), \qquad (4.14.21)$$

where

$$\ell_1 = \lambda_1+n-1,\ \ell_2 = \lambda_2+n-2,\dots,\ell_n = \lambda_n. \qquad (4.14.22)$$

The dimension of the representation space of $\mathcal{S}_\lambda$ is $d^S_\lambda = \chi^S_\lambda(e)$, where e is the identity element of S_k. By (4.14.9),

$$P_e(\zeta) = (\zeta_1+\cdots+\zeta_n)^k = \sum_{|\beta|=k}\frac{k!}{\beta_1!\cdots\beta_n!}\,\zeta_1^{\beta_1}\cdots\zeta_n^{\beta_n}. \qquad (4.14.23)$$

Using the Vandermonde determinant, as in (4.11.22), we have

$$\prod_{i<j}(\zeta_i-\zeta_j) = (-1)^{n(n-1)/2}\sum_{\sigma\in S_n}(\operatorname{sgn}\sigma)\,\zeta_1^{\sigma(1)-1}\cdots\zeta_n^{\sigma(n)-1}. \qquad (4.14.24)$$

After some computation, there results the dimension formula

$$d_\lambda^S = \frac{k!}{\ell_1! \cdots \ell_n!} \prod_{i<j} (\ell_i - \ell_j), \tag{4.14.25}$$

for the representation $\mathcal{S}_\lambda$ of S_k, with $\ell_1, \dots, \ell_n$ given by (4.14.22). For details, see [**19**], pp. 49–50.

4.15. Integral of $|\operatorname{Tr} g^k|^2$ and variants

Integrals of the form

$$\mathcal{I}(\sigma_1,\sigma_2) = \int\limits_{\mathrm{U}(n)} \Xi(\sigma_1,g)\overline{\Xi(\sigma_2,g)}\,dg \tag{4.15.1}$$

are of interest in random matrix theory (cf., e.g., [**14**], [**13**]). Here $\Xi(\sigma,g)$ is as in (4.14.6)–(4.14.10). Note that

$$\sigma\in S_k,\ \vartheta\in\mathbb{R}\Longrightarrow \Xi(\sigma,e^{i\vartheta}g) = e^{ik\vartheta}\Xi(\sigma,g), \tag{4.15.2}$$

and since $g\mapsto e^{i\vartheta}g$ is a measure preserving map on $\mathrm{U}(n)$ it easily follows that

$$\sigma_\nu\in S_{k_\nu},\ k_1\neq k_2\Longrightarrow \mathcal{I}(\sigma_1,\sigma_2)=0. \tag{4.15.3}$$

on the other hand, if $\sigma_1,\sigma_2\in S_k$, one can use (4.14.6) to write

$$\mathcal{I}(\sigma_1,\sigma_2) = \sum_{\lambda\in F_{nk}} \chi^S_\lambda(\sigma_1)\,\overline{\chi}^S_\lambda(\sigma_2). \tag{4.15.4}$$

Such an identity is applied to random matrix theory in [**13**].

Cases of (4.15.4) where $\sigma_1=\sigma_2=\sigma$ are of particular interest. One example, which has already been mentioned in (4.10.12), arises from $\sigma=e$, the identity element of S_k:

$$\int\limits_{\mathrm{U}(n)} |\operatorname{Tr} g|^{2k}\,dg = \sum_{\lambda\in F_{nk}} (f^\lambda)^2, \tag{4.15.5}$$

where f^λ is the dimension of the representation space of $\mathcal{S}_\lambda$. Another interesting example is

$$\int\limits_{\mathrm{U}(n)} |\operatorname{Tr} g^k|^2\,dg = \sum_{\lambda\in F_{nk}} |\chi^S_\lambda(c_k)|^2,\quad c_k=(12\cdots k)\in S_k. \tag{4.15.6}$$

See [**13**] for a direct evaluation of the right side of (4.15.6), using results on the symmetric group. Here we will make a direct calculation of the left side of (4.15.6), using the Weyl integration formula.

We have

$$I_{nk} = \int\limits_{\mathrm{U}(n)} |\operatorname{Tr} g^k|^2\,dM = \frac{1}{(2\pi)^n}\int\limits_{\mathbb{T}^n} |e^{ik\theta_1}+\cdots+e^{ik\theta_n}|^2 J(\theta)\,d\theta. \tag{4.15.7}$$

We re-state this as follows. Set $\zeta_j=e^{i\theta_j}$, so

$$|e^{ik\theta_1}+\cdots+e^{ik\theta_n}|^2 = |\zeta_1^k+\cdots+\zeta_n^k|^2 = \sum_{\mu,\nu}\zeta_\mu^k\zeta_\nu^{-k}. \tag{4.15.8}$$

Also, we use (4.11.23) for $J(\theta)$.

Thus I_{nk} is equal to the constant term in

$$\frac{1}{n!}\sum_{\mu,\nu,\sigma,\tau}(\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)\zeta_\mu^k\zeta_\nu^{-k}\zeta_1^{\sigma(1)-\tau(1)}\cdots\zeta_n^{\sigma(n)-\tau(n)}, \tag{4.15.9}$$

which we write as

$$\frac{1}{n!}(S_1+S_2), \tag{4.15.10}$$

where S_1 arises from the sum over $\mu=\nu$ and S_2 arises from the sum over $\mu\neq\nu$. It is straightforward to obtain

$$S_1 = n\cdot n!. \tag{4.15.11}$$

It remains to consider S_2. We see that, for a given $\mu\neq\nu$, a pair $\sigma,\tau\in S_n$ contributes to S_2 in the sum (4.15.9) if and only if $\sigma(j)=\tau(j)$ for all but two values of $j\in\{1,\dots,n\}$, namely $j=\mu$ and ν, and

$$\begin{aligned}\sigma(\mu)-\tau(\mu)&=-k,\\ \sigma(\nu)-\tau(\nu)&=k.\end{aligned} \tag{4.15.12}$$

Equivalently, we require $\tau=\psi\sigma$ where $\psi\in S_n$ has the property that $\psi(j)=j$ except for two values of $j\in\{1,\dots,n\}$, namely $j_1=\sigma(\mu)$ and $j_2=\sigma(\nu)$, and

$$\psi(j_1)=j_1+k,\quad \psi(j_2)=j_2-k. \tag{4.15.13}$$

This requires $\psi(j_1)=j_2,\ \psi(j_2)=j_1$, with

$$j_1=j_2-k. \tag{4.15.14}$$

Then

$$S_2=\sum(\operatorname{sgn}\sigma)(\operatorname{sgn}\psi\sigma), \tag{4.15.15}$$

the sum running over such allowable (μ,ν,σ,ψ). Note that (4.15.14) constrains j_1; we require $k+1\le j_1\le n$. Thus if $k\ge n$ the sum in (4.15.15) is empty and $S_2=0$. If $1\le k\le n-1$, then there are $(n-k)\cdot n!$ terms in the sum (4.15.15). In fact, if we pick $\sigma\in S_n$ and then pick one of the $n-k$ permutations $\psi\in S_n$ for which (4.15.13) holds, then for each such (σ,ψ), the pair (μ,ν) is uniquely determined. Furthermore, each term in (4.15.15) is equal to $\operatorname{sgn}\psi=-1$. Hence

$$S_2=-(n-k)\cdot n!,\quad 1\le k\le n-1. \tag{4.15.16}$$

Putting together these computations, we have, for integers $k\ge 1$,

$$\int_{\mathrm{U}(n)}|\operatorname{Tr} g^k|^2\,dg = k\wedge n. \tag{4.15.17}$$

The formula (4.15.17) is useful for evaluating inner products of trace functions on $\mathrm{U}(n)$, which arise as follows. If $f:S^1\to\mathbb{C}$ is a bounded

Borel function, define $f(g)$ by the spectral representation of $g \in \mathrm{U}(n)$. Set $X_f(g) = \operatorname{Tr} f(g)$. Using (4.15.17), one can show that

$$\int_{\mathrm{U}(n)} X_u(g) X_v(g)\, dg = \sum_{k=-\infty}^{\infty} a_{nk}\, \hat{u}(k)\hat{v}(-k), \tag{4.15.18}$$

where $\hat{u}(k)$ are the Fourier coefficients of u, $a_{n0} = n^2$, and $a_{nk} = (|k| \wedge n)$ for $k \neq 0$. See §E.5 for details.

To compare the derivation (4.15.7)–(4.15.17) with a treatment via (4.15.6), note that, by Proposition 4.14.2, $\chi_\lambda^S(c_k)$ is the coefficient of $\zeta_1^{\ell_1} \cdots \zeta_n^{\ell_n}$ in

$$\begin{aligned}(\zeta_1^k + \cdots + \zeta_n^k) \prod_{i<j} (\zeta_i - \zeta_j) \\ = (-1)^{n(n-1)/2} \sum_{j=1}^{n} \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \zeta_j^k \zeta_1^{\sigma(1)-1} \cdots \zeta_n^{\sigma(n)-1}.\end{aligned} \tag{4.15.19}$$

Using this one can show that $\chi_\lambda^S(c_k)$ is either 0 or ± 1. Computing the right side of (4.15.6) then apparently involves calculations somewhat similar to those done in (4.15.7)–(4.15.17).

Chapter 5

Some analysis on $\mathrm{U}(n)$

In this chapter we deal with some aspects of analysis on $\mathrm{U}(n)$ connected with the Laplace operator and the heat equation. In §5.1 we give two characterizations of the Laplace operator, first

$$\Delta = \sum X_j^2, \tag{5.0.1}$$

where $\{X_j\}$ is an orthonormal basis of the Lie algebra $\mathfrak{g}$ of $\mathrm{U}(n)$, consisting of vector fields. So defined, Δ is shown to be independent of the choice of such a basis. Second, Δ is the Laplace-Beltrami operator on $\mathrm{U}(n)$, endowed with a bi-invariant metric tensor that defines the same inner product on the tangent space to $\mathrm{U}(n)$ at I as we have on $\mathfrak{g}$. We have that Δ is a bi-invariant differential operator, so it acts as a scalar on each irreducible representation space. We derive the formula

$$d\mathcal{D}_\lambda(\Delta) = -(|\lambda+\rho|^2 - |\rho|^2)I, \tag{5.0.2}$$

where ρ is half the sum of the positive roots of $\mathfrak{g}$.

In §5.2 we look at the heat equation

$$\frac{\partial u}{\partial t} = \Delta u, \quad u(0,x) = f(x), \tag{5.0.3}$$

with solution

$$u(t,x) = H_t * f(x), \tag{5.0.4}$$

and produce a formula for H_t, using (5.0.2), the Weyl character formula, and the dimension formula. H_t is a central function, and we relate $H_t(D(\theta))$ to the heat kernel on $\mathbb{R}^+ \times \mathbb{T}^n$, i.e., to theta functions.

In §5.3 we look at the integral

$$\mathcal{H}(s,X,Y)=\int\limits_{\mathrm{U}(n)} e^{s\,\mathrm{Tr}(gXg^{-1}Y)}\,dg, \tag{5.0.5}$$

for $X,Y\in M(n,\mathbb{C})$, which is of great interest in random matrix theory. In one approach, we concentrate on $X=x,\ Y=y\in \mathrm{U}(n)$, and consider formulas for $\mathcal{H}(s,x,y)$, making use of results of §4.10. In another approach, we take $X,Y\in\mathfrak{u}(n)$ and make contact with the heat kernel, from §5.2.

5.1. The Laplace operator on U(n)

If $\{X_j\}$ is an orthonormal basis of the Lie algebra $\mathfrak{g} = \mathfrak{u}(n)$, regarded as an algebra of left-invariant vector fields, then the Laplace operator on $\mathrm{U}(n)$ is the second order differential operator

$$\Delta = \sum X_j^2. \tag{5.1.1}$$

The Δ is independent of the choice of orthonormal basis. To see this, let $\{Y_j\}$ be another orthonormal basis of $\mathfrak{g}$. Then

$$Y_j = \sum_k a_{kj} X_k, \quad (a_{kj}) = A \in \mathrm{O}(m), \quad m = \dim \mathrm{U}(n).$$

Hence

$$\begin{aligned}\sum_j Y_j^2 &= \sum_j \Big(\sum_k a_{kj} X_k\Big)\Big(\sum_\ell a_{\ell j} X_\ell\Big)\\ &= \sum_{j,k,\ell} a_{kj} a_{\ell j} X_k X_\ell\\ &= \sum_{k,\ell} \delta_{k\ell} X_k X_\ell\\ &= \sum_k X_k^2,\end{aligned}$$

as desired (the third identity by $AA^t = I$).

We will give several explicit formulas for Δ and establish some important basic properties. One is that Δ lies in the center of $\mathfrak{U}(\mathfrak{g})$. Hence its image under an irreducible representation is a scalar, which we will compute.

Recall that $\mathbb{C}\mathfrak{g} = \mathrm{M}(n, \mathbb{C})$, with basis $\{e_{jk} : 1 \le j, k \le n\}$. We have $e_j = ie_{jj} \in \mathfrak{h}$. If we also take $x_{jk}, y_{jk} \in \mathfrak{g}$, for $j < k$, as

$$x_{jk} = \frac{1}{\sqrt{2}}(e_{jk} - e_{kj}), \quad y_{jk} = -\frac{i}{\sqrt{2}}(e_{jk} + e_{kj}), \tag{5.1.2}$$

then we have an orthonormal basis of $\mathfrak{g}$, so

$$\Delta = \sum_j e_j^2 + \sum_{j<k} (x_{jk}^2 + y_{jk}^2). \tag{5.1.3}$$

Proposition 5.1.1. *For all $X \in \mathfrak{g}$, $[\Delta, X] = 0$.*

Proof. It suffices to show that, for all j, k,

$$[\Delta, e_{jk}] = 0, \tag{5.1.4}$$

This is a straightforward computation using (4.2.11), i.e.,

$$[e_{ij}, e_{k\ell}] = \delta_{jk} e_{i\ell} - \delta_{i\ell} e_{kj}. \tag{5.1.5}$$

Details are an exercise. $\square$

We next bring in a geometrical characterization of Δ.

Proposition 5.1.2. *Let* $\mathrm{U}(n)$ *be endowed with a bi-invariant Riemannian metric tensor, coinciding with the inner product on* $\mathfrak{g} = \mathfrak{u}(n)$ *at the identity element of* $\mathrm{U}(n)$. *Then* Δ *is equal to the Laplace-Beltrami operator on* $\mathrm{U}(n)$.

Proof. Let L denote the Laplace-Beltrami operator. Then Δ and L are both second-order, self-adjoint, bi-invariant differential operators on $\mathrm{U}(n)$, whose expressions in local coordinates have the same leading order terms at I, hence on all of $\mathrm{U}(n)$. Thus $L - \Delta$ is a first-order differential operator with real coefficients, so

$$L - \Delta = X + V,$$

where X is a real vector field and V is multiplication by a function $V(x)$. Since $L1 = \Delta 1 = 0$, we have $V = 0$, so $L - \Delta = X$. But X is a real, bi-invariant vector field, hence is skew-adjoint, so also $X = 0$. □

We produce some more useful formulas for Δ. Note that

$$x_{jk} + iy_{jk} = \sqrt{2}\, e_{jk}, \quad x_{jk} - iy_{jk} = -\sqrt{2}\, e_{kj}. \tag{5.1.6}$$

Also, using (5.1.5), we have

$$[x_{jk}, y_{jk}] = -(e_j - e_k). \tag{5.1.7}$$

Hence we can rewrite (5.1.3) as

$$\begin{aligned} \Delta &= \sum_j e_j^2 + \sum_{j<k} \big\{ (x_{jk} - iy_{jk})(x_{jk} + iy_{jk}) - i[x_{jk}, y_{jk}] \big\} \\ &= \sum_j e_j^2 + i \sum_{j<k} (e_j - e_k) - 2 \sum_{j<k} e_{kj} e_{jk}. \end{aligned} \tag{5.1.8}$$

The significance for representation theory is highlighted by the following:

Proposition 5.1.3. *For the irreducible representation* $\mathcal{D}_\lambda$ *of* $\mathrm{U}(n)$ *on* $V(\lambda)$, *we have*

$$d\mathcal{D}_\lambda(\Delta) = -(|\lambda + \rho|^2 - |\rho|^2) I. \tag{5.1.9}$$

Proof. It follows from Proposition 3.7.1 that $d\mathcal{D}_\lambda(\Delta)$ is scalar on $V(\lambda)$. Thus it suffices to evaluate $d\mathcal{D}_\lambda(\Delta)v$ when v is a highest weight vector. In such a case $d\mathcal{D}_\lambda(e_{jk})v = 0$ when $j < k$, so

$$d\mathcal{D}_\lambda(\Delta)v = \sum_j d\mathcal{D}_\lambda(e_j)^2 v + i \sum_{j<k} d\mathcal{D}_\lambda(e_j - e_k)v. \tag{5.1.10}$$

Now $d\mathcal{D}_\lambda(e_j)v = i\lambda(e_j)v$, so

$$\begin{aligned} d\mathcal{D}_\lambda(\Delta)v &= -\sum_j \lambda(e_j)^2 v - \sum_{j<k} \lambda(e_j - e_k)v \\ &= -(|\lambda|^2 + 2\langle \lambda, \rho \rangle)v, \end{aligned} \tag{5.1.11}$$

where we recall from (4.12.16) that

$$\rho = \frac{1}{2} \sum_{j<k} \omega_{jk} \in \mathfrak{h}'.$$

This gives (5.1.8). $\square$

These results generalize to any compact Lie group G. Give G a bi-invariant Riemannian metric tensor and let Δ be the Laplace-Beltrami operator. Then left and right translations are isometries, so Δ commutes with $L(g)$ and $R(g)$ for all $g \in G$. If π^α is an irreducible unitary representation of G, onto a space of dimension d_α, and with matrix form (π^α_{jk}), let

$$\mathcal{V}_\alpha = \text{span of } \{\pi^\alpha_{jk} : 1 \le j, k \le d_\alpha\}, \tag{5.1.12}$$

and let $\mathcal{P}_\alpha$ denote the orthogonal projection of $L^2(G)$ onto $\mathcal{V}_\alpha$, as in Proposition 11.5. We have

$$\mathcal{P}_\alpha u(x) = d_\alpha \sum_{j,k} \pi^\alpha_{jk}(x) \int_G u(y) \overline{\pi}^\alpha_{jk}(y)\, dy. \tag{5.1.13}$$

Writing $\sum \pi^\alpha_{jk}(x) \overline{\pi}^\alpha_{jk}(y) = \operatorname{Tr}(\pi^\alpha(x)\pi^\alpha(y)^*) = \chi_\alpha(xy^{-1})$, we see that

$$\mathcal{P}_\alpha u(x) = d_\alpha\, \chi_\alpha * u(x). \tag{5.1.14}$$

It follows that $\mathcal{P}_\alpha$ commutes with Δ, so

$$\Delta : \mathcal{V}_\alpha \longrightarrow \mathcal{V}_\alpha.$$

Now $G \times G$ acts on $\mathcal{V}_\alpha$ via

$$\Gamma_\alpha(g,h)u(x) = u(g^{-1}xh), \quad u \in \mathcal{V}_\alpha. \tag{5.1.15}$$

A brief calculation shows that

$$\Gamma_\alpha(g,h)\pi^\alpha_{jk}(x) = \sum_{\ell,m} \overline{\pi}^\alpha_{\ell j}(g) \pi^\alpha_{mk}(h) \pi^\alpha_{\ell m}(x). \tag{5.1.16}$$

It follows readily that

$$\operatorname{Tr} \Gamma_\alpha(g,h) = \overline{\chi}_\alpha(g) \chi_\alpha(h), \tag{5.1.17}$$

and in particular

$$\int_{G\times G} |\operatorname{Tr} \Gamma_\alpha(g,h)|^2\, dg\, dh = \int_G |\overline{\chi}_\alpha(g)|^2\, dg \int_G |\chi_\alpha(h)|^2\, dh = 1, \tag{5.1.18}$$

so Γ_α is an irreducible representation of $G \times G$. Since the Laplace operator Δ commutes with Γ_α, it must be a scalar on $\mathcal{V}_\alpha$:

$$\Delta \pi^\alpha_{jk}(x) = c(\alpha)\, \pi^\alpha_{jk}(x). \tag{5.1.19}$$

The formula (5.1.1) also holds for Δ in the more general setting of a compact Lie group G with a bi-invariant metric, and $d\pi(\Delta)$ is defined for a finite-dimensional representation π of G. Since $R|_{\mathcal{V}_\alpha}$ acts as a sum of copies of π^α, by (5.1.16), we see that

$$d\pi^\alpha(\Delta) = c(\alpha)I. \tag{5.1.20}$$

We mention that (5.1.9) generalizes from $\mathrm{U}(n)$ to a general compact G; see, e.g., [**38**], pp. 123–124, for a derivation.

Let us return to $G = \mathrm{U}(n)$, with irreducible representations $\mathcal{D}_\lambda$. Specializing the fact that identical factors of $c(\alpha)$ appear in (5.1.19) and (5.1.20), we see that (5.1.9) gives

$$\Delta\pi^\lambda_{jk}(x) = -(|\lambda+\rho|^2 - |\rho|^2)\pi^\lambda_{jk}(x), \quad 1, \le j, k \le d_\lambda, \tag{5.1.21}$$

if (π^λ_{jk}) denotes a matrix form of $\mathcal{D}_\lambda$.

5.2. The heat equation on $\mathrm{U}(n)$

Before specializing to $G = \mathrm{U}(n)$, we begin with an arbitrary compact Lie group G, with bi-invariant Riemannian metric and Laplace operator Δ, as discussed in §5.1. We consider the heat equation for $u(t,x)$ on $\mathbb{R}^+ \times G$:

$$\frac{\partial u}{\partial t} = \Delta u, \quad u(0,x) = f(x). \tag{5.2.1}$$

As seen in §2.3, we can write

$$f(x) = \sum_{\alpha\in\mathcal{I}} d_\alpha \sum_{j,k} \hat{f}_{jk}(\alpha)\, \pi^\alpha_{jk}(x), \tag{5.2.2}$$

where $\{\pi^\alpha : \alpha \in \mathcal{I}\}$ is a complete set of irreducible unitary representations of G and

$$\hat{f}_{jk}(\alpha) = \overline{\pi}^\alpha_{jk}(f) = \int\limits_G f(y)\overline{\pi}^\alpha_{jk}(y)\, dy. \tag{5.2.3}$$

In view of (5.1.19), we then have

$$u(t,x) = \sum_{\alpha\in\mathcal{I}} d_\alpha e^{c(\alpha)t} \sum_{j,k} \hat{f}_{jk}(\alpha)\pi^\alpha_{jk}(x). \tag{5.2.4}$$

We can write

$$u(t,x) = H_t * f(x), \tag{5.2.5}$$

with $H_t(x)$, known as the heat kernel, given as follows. By (2.7.5), (5.2.5) is equivalent to

$$\widehat{H}_t(\alpha) = e^{c(\alpha)t}\, I, \tag{5.2.6}$$

so, parallel to (5.2.2),

$$H_t(x) = \sum_{\alpha\in\mathcal{I}} d_\alpha e^{c(\alpha)t}\, \chi_\alpha(x). \tag{5.2.7}$$

Specializing to $\mathrm{U}(n)$, with irreducible representations $\mathcal{D}_\lambda$, parametrized by $\mathcal{P}_+ = \{\lambda \in \mathbb{Z}^n : \lambda_1 \geq \cdots \geq \lambda_n\}$, we have

$$H_t(x) = \sum_{\lambda\in\mathcal{P}_+} d_\lambda e^{-(|\lambda+\rho|^2 - |\rho|^2)t}\, \chi_\lambda(x). \tag{5.2.8}$$

Note that H_t is a central function, uniquely determined by its values at $D(\theta)$, $\theta \in \mathbb{R}^n$. We bring in the Weyl character formula and the dimension formula to write

$$H_t(D(\theta)) = \frac{e^{|\rho|^2 t}}{MA_\rho(\theta)} \sum_{\lambda\in\mathcal{P}_+} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\Delta(\lambda+\rho) e^{i\langle\sigma\cdot(\lambda+\rho),\theta\rangle} e^{-|\lambda+\rho|^2 t}. \tag{5.2.9}$$

Here $M = \prod_{\ell=1}^{n-1} \ell!$ is the denominator calculated in (4.12.34) and

$$\Delta(\lambda) = \prod_{j<k} \langle \omega_{jk}, \lambda \rangle = \prod_{j<k} (\lambda_j - \lambda_k). \tag{5.2.10}$$

We can rewrite (5.2.9) as

$$H_t(D(\theta)) = \frac{e^{|\rho|^2 t}}{MA_\rho(\theta)} \sum_{\lambda \in \widetilde{\mathcal{P}}_+} \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \Delta(\lambda) e^{i\langle \sigma \cdot \lambda, \theta \rangle} e^{-|\lambda|^2 t}, \tag{5.2.11}$$

where

$$\widetilde{\mathcal{P}}_+ = \{\lambda \in (\mathbb{Z} + \gamma)^n : \lambda_1 > \cdots > \lambda_n\}, \tag{5.2.12}$$

and where

$$\gamma = 0 \quad \text{for} \quad n \quad \text{odd}, \;\; 1/2 \quad \text{for} \quad n \quad \text{even}. \tag{5.2.13}$$

Note that

$$\Delta(\sigma \cdot \lambda) = (\operatorname{sgn} \sigma) \Delta(\lambda), \tag{5.2.14}$$

and that $\Delta(\lambda) = 0$ whenever $\lambda_\mu = \lambda_\nu$ for some $\mu \neq \nu$. Hence

$$H_t(D(\theta)) = \frac{e^{|\rho|^2 t}}{MA_\rho(\theta)} \sum_{\lambda \in (\mathbb{Z}+\gamma)^n} \Delta(\lambda) e^{i\langle \lambda, \theta \rangle} e^{-|\lambda|^2 t}. \tag{5.2.15}$$

Let us set

$$E_\gamma(t, \theta) = \sum_{\lambda \in (\mathbb{Z}+\gamma)^n} e^{i\langle \lambda, \theta \rangle} e^{-|\lambda|^2 t}. \tag{5.2.16}$$

Then (5.2.15) yields

$$H_t(D(\theta)) = \frac{e^{|\rho|^2 t}}{MA_\rho(\theta)} Q(D) E_\gamma(t, \theta), \tag{5.2.17}$$

where $Q(D)$ is the differential operator on functions of θ:

$$Q(D)v(\theta) = \prod_{j<k} \frac{1}{i} \Big(\frac{\partial}{\partial \theta_j} - \frac{\partial}{\partial \theta_k} \Big). \tag{5.2.18}$$

Note that $E_\gamma(t, \theta)$ satisfies the heat equation on $\mathbb{R}^+ \times \mathbb{R}^n$:

$$\frac{\partial E_\gamma}{\partial t} = \sum_{j=1}^n \frac{\partial^2 E_\gamma}{\partial \theta_j^2}. \tag{5.2.19}$$

Also, $E_\gamma(t,\theta)$ is periodic in each variable θ_j, of period 2π if $\gamma = 0$ and of period 4π if $\gamma = 1/2$. In fact,

$$\begin{aligned} E_0(0,\theta) &= (2\pi)^n \sum_{\nu\in\mathbb{Z}^n} \delta_{2\pi\nu}(\theta), \\ E_{1/2}(0,\theta) &= (2\pi)^n \sum_{\nu\in\mathbb{Z}^n} (-1)^{\nu_1+\cdots+\nu_n}\, \delta_{2\pi\nu}(\theta). \end{aligned} \tag{5.2.20}$$

Hence, for $t > 0$,

$$E_0(t,\theta) = \left(\frac{\pi}{t}\right)^{n/2} \sum_{\nu\in\mathbb{Z}^n} e^{-|\theta-2\pi\nu|^2/4t}, \tag{5.2.21}$$

and

$$E_{1/2}(t,\theta) = \left(\frac{\pi}{t}\right)^{n/2} \sum_{\nu\in\mathbb{Z}^n} (-1)^{\nu_1+\cdots+\nu_n}\, e^{-|\theta-2\pi\nu|^2/4t}. \tag{5.2.22}$$

Both functions have the same asymptotic behavior for $|\theta| \le \pi$ as $t \searrow 0$:

$$E_\gamma(t,\theta) \sim \left(\frac{\pi}{t}\right)^{n/2} e^{-|\theta|^2/4t}. \tag{5.2.23}$$

Regarding the heat kernel $H_t(x)$, we also have

$$H_t(x) \sim (4\pi t)^{-n^2/2}\, e^{-d(x)^2/4t}, \quad t \searrow 0, \tag{5.2.24}$$

where $d(x)$ denotes the distance from x to the identity element in the Riemannian metric on $\mathrm{U}(n)$. In particular, for $X \in \mathfrak{u}(n)$, $|X| \le \pi/2$,

$$H_t(e^X) \sim (4\pi t)^{-n^2/2}\, e^{-|X|^2/4t}, \quad t \searrow 0. \tag{5.2.25}$$

The n^2 in the exponent of t arises as the dimension of $\mathrm{U}(n)$. This is a special case of a general analysis of the heat kernel on a Riemannian manifold; see [**39**], Chapter 7 for a proof, and Chapter 10 for important geometrical applications. The reader might try to obtain (5.2.25) from (5.2.17) and (5.2.23).

5.3. The Harish-Chandra/Itzykson-Zuber integral

The integral

$$\int_{\mathrm{U}(n)} e^{s\,\mathrm{Tr}(gXg^{-1}Y)}\,dg = \mathcal{H}(s,X,Y), \quad X,Y\in \mathrm{M}(n,\mathbb{C}), \tag{5.3.1}$$

is of great interest in random matrix theory. Here we give several formulas for this, with arguments adapted from [**2**] and [**25**]. Note that $\mathcal{H}(s,X,Y)$ is holomorphic in its arguments, so it is uniquely determined from its values on various subsets. Let us take $X=x, Y=y\in \mathrm{U}(n)$, and write

$$\begin{aligned}\mathcal{H}(s,x,y) &= \sum_{k=0}^{\infty}\frac{s^k}{k!}\int_{\mathrm{U}(n)}\bigl(\mathrm{Tr}(gxg^{-1}y)\bigr)^k\,dg\\ &= \sum_{k=0}^{\infty}\frac{s^k}{k!}\int_{\mathrm{U}(n)}\mathrm{Tr}\otimes^k(gxg^{-1}y)\,dg.\end{aligned}\tag{5.3.2}$$

To proceed, we use the following.

Lemma 5.3.1. *Let π,π' be irreducible representations of a compact Lie group G, with characters $\chi_\pi,\chi_{\pi'}$. Then*

$$\int_G \chi_\pi(xy)\overline{\chi}_{\pi'}(y)\,dy = d_\pi^{-1}\chi_\pi(x)\,\delta_{\pi\pi'}, \tag{5.3.3}$$

where $\delta_{\pi\pi'}=1$ if $\pi\approx\pi'$, 0 otherwise. Furthermore,

$$\int_G \chi_\pi(gxg^{-1}y)\,dg = d_\pi^{-1}\chi_\pi(x)\chi_\pi(y). \tag{5.3.4}$$

Proof. If we write the integrand in the left side of (5.3.3) as $\mathrm{Tr}(\pi(x)\pi(y))\overline{\chi}_{\pi'}(y)$ and apply Proposition 2.4.2 to $\int_G \pi(y)\overline{\chi}_{\pi'}(y)\,dy$, we get the identity (5.3.3). As for (5.3.4), one easily shows the left side is invariant under $y\mapsto h^{-1}yh,\ h\in G$. Hence

$$\int_G \chi_\pi(gxg^{-1}y)\,dg = \sum_{\alpha\in\mathcal{I}}\psi_\alpha(x)\,\chi_\alpha(y), \tag{5.3.5}$$

with

$$\psi_\alpha(x) = \int_G\int_G \chi_\pi(gxg^{-1}y)\overline{\chi}_\alpha(y)\,dy\,dg = d_\pi^{-1}\,\delta_{\pi\pi_\alpha}, \tag{5.3.6}$$

the last identity by (5.3.3). ∎

To apply (5.3.4) to (5.3.2), we break up $\otimes^k$ into irreducibles, using Proposition 4.10.3. We get

$$\int_{\mathrm{U}(n)} \operatorname{Tr} \otimes^k (gxg^{-1}y)\, dy = \sum_{\lambda \in F_{nk}} \frac{f^\lambda}{d_\lambda}\, \chi_\lambda(x)\chi_\lambda(y), \tag{5.3.7}$$

where f^λ is the dimension of the representation space for $\mathcal{S}_\lambda$. Hence

$$\mathcal{H}(s, x, y) = \sum_{k=0}^{\infty} \frac{s^k}{k!} \sum_{\lambda \in F_{nk}} \frac{f^\lambda}{d_\lambda}\, \chi_\lambda(x)\chi_\lambda(y). \tag{5.3.8}$$

For a second approach, we apply (5.3.4) to each term in the series (5.2.8) for the heat kernel $H_t(x)$, obtaining

$$\int_{\mathrm{U}(n)} H_t(gxg^{-1}y)\, dg = \sum_{\lambda \in \mathcal{P}_+} e^{-(|\lambda+\rho|^2 - |\rho|^2)t}\, \chi_\lambda(x)\chi_\lambda(y). \tag{5.3.9}$$

We analyze this in a fashion parallel to (5.2.9)–(5.2.15). Denoting the quantity (5.3.9) by $K_t(x, y)$, we have

(5.3.10)

$$\begin{aligned} &K_t(D(\theta), D(\varphi)) \\ &= \frac{e^{|\rho|^2 t}}{A_\rho(\theta)A_\rho(\varphi)} \sum_{\lambda \in \mathcal{P}_+} \sum_{\sigma,\tau} (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau) e^{i\langle \sigma\cdot(\lambda+\rho), \theta\rangle} e^{i\langle \tau\cdot(\lambda+\rho), \varphi\rangle} e^{-|\lambda+\rho|^2 t} \\ &= \frac{e^{|\rho|^2 t}}{A_\rho(\theta)A_\rho(\varphi)} \sum_{\lambda \in \widetilde{\mathcal{P}}_+} \sum_{\sigma,\tau} (\operatorname{sgn}\tau) e^{i\langle \sigma\cdot\lambda, \theta + \tau^t\varphi\rangle} e^{-|\lambda|^2 t}, \end{aligned}$$

where we take $\tau \mapsto \tau\sigma$ to produce the last identity, and we define $\widetilde{\mathcal{P}}_+$ as in (5.2.16). To proceed further, we note that

$$B_\lambda(\theta + \tau^t\varphi) = \sum_{\sigma,\tau} (\operatorname{sgn}\tau) e^{i\langle \sigma\cdot\lambda, \theta+\tau^t\varphi\rangle} = A_\lambda(\theta)A_\lambda(\varphi), \tag{5.3.11}$$

and

$$A_{\sigma\cdot\lambda}(\theta) = (\operatorname{sgn}\sigma)A_\lambda(\theta). \tag{5.3.12}$$

Hence $B_\lambda(\theta + \tau^t\varphi)$ vanishes whenever there exists $\sigma \neq e$ such that $\sigma\cdot\lambda = \lambda$ and $\operatorname{sgn}\sigma = -1$, hence whenever $\lambda_\mu = \lambda_\nu$ for some $\mu \neq \nu$. Thus we can rewrite (5.3.10) as

(5.3.13)

$$K_t(D(\theta), D(\varphi)) = \frac{e^{|\rho|^2 t}}{A_\rho(\theta)A_\rho(\varphi)} \sum_{\lambda \in (Z+\gamma)^n} \sum_{\tau \in S_n} (\operatorname{sgn}\tau) e^{i\langle \lambda, \theta+\tau^t\varphi\rangle} e^{-|\lambda|^2 t}.$$

Equivalently, with $E_\gamma(t,\theta)$ as in (33.16),

$$K_t(D(\theta),D(\varphi)) = \frac{e^{|\rho|^2 t}}{A_\rho(\theta)A_\rho(\varphi)} \sum_{\tau\in S_n} (\operatorname{sgn}\tau) E_\gamma(t,\theta+\tau^t\varphi). \tag{5.3.14}$$

The relevance of (5.3.14) to a calculation of (5.3.1) arises from the heat kernel asymptotics

$$H_t(g) \sim (4\pi t)^{-n^2/2}\, e^{-d(g)^2/4t}, \quad t\searrow 0, \tag{5.3.15}$$

discussed in §5.2. (Actually, an extra factor A_n enters, which we will discuss at the end of this section.) Let us take

$$x = e^{\eta X}, \quad y = e^{\eta Y}, \quad X,Y\in \mathfrak{u}(n), \tag{5.3.16}$$

with $|\eta|$ small. Then

$$gxg^{-1}y = e^{\eta\, \mathrm{Ad}(g)X} e^{\eta Y} = e^{\eta(\mathrm{Ad}(g)X+Y)} + O(\eta^2). \tag{5.3.17}$$

We take

$$\eta = 2\sqrt{t}. \tag{5.3.18}$$

Then

$$\begin{aligned} H_t(gxg^{-1}y) &\sim (4\pi t)^{-n^2/2}\, e^{-\eta^2|\mathrm{Ad}(g)X+Y|^2/4t} \\ &= (4\pi t)^{-n^2/2}\, e^{-(|X|^2+|Y|^2+2\langle \mathrm{Ad}(g)X,Y\rangle)}. \end{aligned} \tag{5.3.19}$$

Hence, with (5.3.16) and (5.3.18) in effect,

$$\begin{aligned} K_t(x,y) &\sim (4\pi t)^{-n^2/2} e^{-(|X|^2+|Y|^2)} \int\limits_{\mathrm{U}(n)} e^{-2\langle \mathrm{Ad}(g)X,Y\rangle}\, dg \\ &= (4\pi t)^{-n^2/2} e^{-(|X|^2+|Y|^2)} \mathcal{H}(2,X,Y), \end{aligned} \tag{5.3.20}$$

since $\langle X,Y\rangle = -\operatorname{Tr}(XY)$ in this case, so, for $X,Y\in\mathfrak{u}(n)$,

$$\mathcal{H}(2,X,Y) = e^{|X|^2+|Y|^2} \lim_{t\searrow 0} (4\pi t)^{n^2/2}\, K_t(e^{2\sqrt{t}X}, e^{2\sqrt{t}Y}). \tag{5.3.21}$$

Say

$$e^{\eta X} \sim D(\eta\theta), \quad e^{\eta Y}\sim D(\eta\varphi), \tag{5.3.22}$$

i.e., these matrices are similar. Then (5.3.14) gives

$$\begin{aligned} &K_t(e^{2\sqrt{t}X}, e^{2\sqrt{t}Y}) \\ &\quad = \frac{e^{|\rho|^2 t}}{A_\rho(2\sqrt{t}\theta)A_\rho(2\sqrt{t}\varphi)} \sum_{\tau\in S_n} (\operatorname{sgn}\tau) E_\gamma(t, 2\sqrt{t}(\theta+\tau^t\varphi)). \end{aligned} \tag{5.3.23}$$

Now, as $t\searrow 0$,

$$A_\rho(2\sqrt{t}\theta) \sim (2i\sqrt{t})^{n(n-1)/2}\, \Delta(\theta), \tag{5.3.24}$$

and, by (5.2.22),

$$E_\gamma(t, 2\sqrt{t}\theta) \sim \left(\frac{\pi}{t}\right)^{n/2} e^{-|\theta|^2}. \tag{5.3.25}$$

Thus, as $t \searrow 0$,

(5.3.26)

$$K_t(e^{2\sqrt{t}X}, e^{2\sqrt{t}Y}) \sim t^{-n/2} \frac{\pi^{n/2}}{(2i)^{n(n-1)}} \frac{1}{\Delta(\theta)\Delta(\varphi)} \sum_{\tau \in S_n} (\operatorname{sgn} \tau) e^{-|\theta + \tau^t \varphi|^2}.$$

Hence, by (5.3.21),

$$\mathcal{H}(2, X, Y) = \frac{C_n}{\Delta(\theta)\Delta(\varphi)} \sum_{\tau \in S_n} (\operatorname{sgn} \tau) e^{-2\langle \tau \cdot \theta, \varphi \rangle}, \tag{5.3.27}$$

where X and Y are related to θ and φ by (5.3.22). Since clearly $\mathcal{H}(s, X, Y) = \mathcal{H}(1, sX, Y) = \mathcal{H}(2, (s/2)X, Y)$, we have

(5.3.28)

$$\mathcal{H}(s, X, Y) = s^{-n(n-1)/2} \frac{C_n'}{\Delta(\theta)\Delta(\varphi)} \sum_{\tau \in S_n} (\operatorname{sgn} \tau) e^{-s\langle \tau \cdot \theta, \varphi \rangle}, \quad X, Y \in \mathfrak{u}(n).$$

Our renaming of the constants C_n, C_n' derives from the fact that (5.3.15) holds when $\mathrm{U}(n)$ has the Riemannian metric in which the norm on $T_I \, \mathrm{U}(n) = \mathfrak{u}(n)$ is the Hilbert-Schmidt norm. However, in normalizing the Haar measure on $\mathrm{U}(n)$ to have mass one, we scale the metric. One way to evaluate C_n' in (5.3.28) is to consider the $s \to 0$ limit, using $\mathcal{H}(0, X, Y) = 1$. In fact,

$$C_n' = \prod_{\ell=1}^{n-1} \ell!. \tag{5.3.29}$$

We mention that if instead of X and Y being skew-adjoint, as in (5.3.28), we take X and Y self-adjoint, with eigenvalues θ_j and φ_j, respectively, then (5.3.28) holds with $-s$ changed to s in the exponent. This is the form in which the identity commonly appears. See [**18**] for an application of such an identity.

Chapter 6

Representations of general compact Lie groups

Given a compact, connected Lie group G, with Lie algebra $\mathfrak{g}$, we want to present results on its representation theory parallel to those discussed for $G = \mathrm{U}(n)$ and $\mathrm{SU}(n)$ in Chapter 4.

To start, we can find a torus $\mathbb{T} \subset G$, of maximal dimension, whose Lie algebra $\mathfrak{h}$ is a commutative Lie subalgebra of the Lie algebra $\mathfrak{g}$ of maximal dimension. Given a continuous unitary representation π of G on a finite dimensional V, we simultaneously diagonalize it on $\mathbb{T}$, forming a decomposition

$$V = \bigoplus_{\lambda \in \mathfrak{h}'} V_\lambda, \quad V_\lambda = \{v \in V : d\pi(X)v = i\lambda(X)v,\ X \in \mathfrak{h}\}. \tag{6.0.1}$$

When $V_\lambda \neq 0$, we say λ is a weight for π, and a nonzero $v \in V_\lambda$ is called a weight vector. In the special case where π is the adjoint representation, on $V = \mathfrak{g}_{\mathbb{C}}$, we have

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus \bigoplus_{\alpha} \mathfrak{g}_\alpha, \tag{6.0.2}$$

where α runs over the roots of $\mathfrak{g}$.

We put an order on $\mathfrak{h}'$, which induces an order on the roots and more generally on the weights. In particular, we refine (6.0.2) to

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus \mathfrak{n}_+ \oplus \mathfrak{n}_-, \quad \mathfrak{n}_\pm = \bigoplus_{\pm\alpha > 0} \mathfrak{g}_\alpha. \tag{6.0.3}$$

It follows that each finite dimensional representation π of G has a highest weight. We show that if π is an irreducible unitary representation of G on V, then there is exactly one weight λ for which V_λ is annihilated by all the "raising operators," i.e., the operators $d\pi(e_\alpha)$ for all $e_\alpha \in \mathfrak{g}_\alpha,\ \alpha > 0$. This λ is the highest weight, and $\dim V_\lambda = 1$. Furthermore, if π_1 and π_2 are irreducible representations with the same highest weight, then they are unitarily equivalent. As in the case of $\mathrm{SU}(n)$, this presents the problem of identifying just which elements of $\mathfrak{h}'$ arise as highest weights of irreducible representations.

In case $\mathfrak{g}$ has trivial center and G is simply connected, the answer to this problem is given by the Theorem of the Highest Weight, which says that an element $\lambda \in \mathfrak{h}'$ is the highest weight of an irreducible representation of G if and only if

$$2\frac{\langle \lambda, \alpha\rangle}{\langle \alpha, \alpha\rangle} \in \mathbb{Z}^+, \tag{6.0.4}$$

for each positive root α of $\mathfrak{g}$. Here the inner product on $\mathfrak{h}'$ is derived from that on $\mathfrak{g}$, namely the Ad-invariant inner product determined by the Killing form. An element $\lambda \in \mathfrak{h}'$ satisfying (6.0.4) for each positive root α is called a *dominant integral weight*. When this result is specialized to $G = \mathrm{SU}(n)$, we get the weights described in Chapter 4.

Considering the adjoint representation of G on $\mathfrak{g}_C$, we show that

$$n_{\alpha\beta} = 2\frac{\langle \alpha, \beta\rangle}{\langle \beta, \beta\rangle} \in \mathbb{Z}, \tag{6.0.5}$$

whenever α and β are roots. The integers $n_{\alpha\beta}$ are called Cartan integers. Further analysis shows that

$$n_{\alpha\beta} \in \{0, \pm 1, \pm 2, \pm 3\}. \tag{6.0.6}$$

This leads to a number of restrictions on the structure of roots, obtained in §6.2. An important source of results obtained in this section is the injection $\mathfrak{su}(2) \hookrightarrow \mathfrak{g}$ associated to each root pair $\{\pm\alpha\}$.

In §6.3 we discuss the Weyl group, which acts as a group of automorphisms of $\mathbb{T}$. In Chapter 4, for $G = \mathrm{U}(n)$, the Weyl group was S_n. Its significance is that each central function $f \in C(G)$ is uniquely determined by its restriction to $\mathbb{T}$, on which it is invariant under the Weyl group action. In §6.4 we examine a certain family of functions φ_λ, which have an intriguing relation to the characters χ_λ of irreducible representations. Unlike in Chapter 4, we do not derive the character formula in general here. For this, one can consult [**34**].

As seen above, we make use of the complexification $\mathfrak{g}_{\mathbb{C}}$ of the Lie algebra $\mathfrak{g}$. We also make use of the complexification of the Lie group G. Section 6.5 gives a construction of this complexification.

In §6.6 we consider a subset Σ of the set of positive roots, called simple roots. The matrix $(n_{\alpha\beta})$ of Cartan integers formed as α, β range over Σ is called the Cartan matrix of $\mathfrak{g}$. Using it, one constructs a certain graph, whose vertices are the elements of Σ, called the Dynkin diagram. We discuss the Cartan matrices and Dynkin diagrams of various Lie algebras, including those of $\mathrm{U}(n), \mathrm{SO}(2k), \mathrm{SO}(2k+1)$, and G_2.

Section 6.7 presents some results on representations of compact Lie groups that are not connected, such as $O(n)$, which has two connected components, one being $SO(n)$. We see that the representation theory of $O(n)$ is a simple variant of that of $SO(n)$ if n is odd, but entails further complications if n is even. A preview of this for $n = 2$ was given in §2.5. Here we take a detailed look at the case $n = 4$.

Appendix 6.A has additional material on maximal tori. If G is a compact Lie group, we say a torus $\mathbb{T} \subset G$ is a conjugating torus for G if each $x \in G$ is conjugate to an element of $\mathbb{T}$. We record standard examples of conjugating tori for $U(n), SU(n)$, and $SO(n)$. We also describe conjugating tori for $Sp(n)$, referring to Chapter 10 for a proof. We show that if $\mathbb{T}$ is a conjugating torus for G and $\mathbb{T}'$ is another torus, then there exists $g \in G$ such that $\mathbb{T}' \subset g^{-1}\mathbb{T}g$. It follows that such $\mathbb{T}$ is a maximal torus, and moreover each maximal torus in G is conjugate to $\mathbb{T}$.

6.1. Roots and weights for general compact Lie groups

The notions of roots and weights, described for $\mathrm{U}(n)$ in §4.2, have natural counterparts for a general compact, connected Lie group G. Take such a group, denote its Lie algebra by $\mathfrak{g}$, and endow G with a bi-invariant Riemannian metric, so $\mathfrak{g}$ has an inner product $\langle\ ,\ \rangle$ with the property that for each $g \in G$, $\operatorname{Ad} g$ is an orthogonal transformation on $\mathfrak{g}$, and hence for each $X \in \mathfrak{g}$, $\operatorname{ad} X$ is a skew-adjoint operator on $\mathfrak{g}$.

Let $\mathfrak{h}$ be a commutative subalgebra of $\mathfrak{g}$ of maximal dimension, and denote the Lie group it generates by $\mathbb{T}$. This group is commutative, and it is closed, hence compact. Indeed, otherwise its closure $\overline{\mathbb{T}}$ would be a commutative Lie subgroup of G of larger dimension. Note that the exponential map $\mathfrak{h} \to \mathbb{T}$ is a group homomorphism. Hence $\mathbb{T}$ is a compact quotient of a Euclidean space by a discrete subgroup; hence $\mathbb{T}$ is a torus. It is called a *maximal torus* in G. The dimension (say n) of $\mathbb{T}$, or equivalently of $\mathfrak{h}$, is called the *rank* of G.

If $\{h_1, \dots, h_n\}$ is a basis of $\mathfrak{h}$, we can simultaneously put the skew-adjoint operators $\operatorname{ad} h_j$ on $\mathfrak{g}$ in normal form. In fact, for almost all choices $r_j \in \mathbb{R}$, $h^b = \sum r_j h_j$ separates out the spectra of $\operatorname{ad} h_j$ and it suffices to put $\operatorname{ad} h^b$ in normal form. Hence there is a set of elements $x_1, \dots, x_k, y_1, \dots, y_k \in \mathfrak{g}$ such that

$$\{h_1, \dots, h_n, x_1, \dots, x_k, y_1, \dots, y_k\}$$

is a basis of $\mathfrak{g}$ with the property that

$$\operatorname{ad} h(x_j \pm i y_j) = \pm i \alpha_j(h)(x_j \pm i y_j), \quad \forall\, h \in \mathfrak{h}, \tag{6.1.1}$$

for certain $\alpha_j \in \mathfrak{h}'$. Hence we can decompose the complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$ as

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus \bigoplus_{\alpha} \mathfrak{g}_\alpha, \tag{6.1.2}$$

where, given $\alpha \in \mathfrak{h}'$,

$$\mathfrak{g}_\alpha = \{z \in \mathfrak{g}_{\mathbb{C}} : [h, z] = i\alpha(h) z,\ \forall\, h \in \mathfrak{h}\}. \tag{6.1.3}$$

If $\mathfrak{g}_\alpha \neq 0$, we call α a *root*, and nonzero elements of $\mathfrak{g}_\alpha$ are called *root vectors*, provided $\alpha \neq 0$. Note that $\mathfrak{g}_0 = \mathfrak{h}_{\mathbb{C}}$. From the Jacobi identity, in the form

$$\operatorname{ad} h([z_\alpha, z_\beta]) = [\operatorname{ad} h(z_\alpha), z_\beta] + [z_\alpha, \operatorname{ad} h(z_\beta)], \tag{6.1.4}$$

it follows that

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}. \tag{6.1.5}$$

Note from (6.1.1) that if α is a root, so is $-\alpha$.

The choice of ordered basis $\{h_j : j = 1, \dots, n\}$ of $\mathfrak{h}$ induces an ordering of $\mathfrak{h}'$ as follows. Given $\alpha, \beta \in \mathfrak{h}'$, we say $\alpha > \beta$ provided the first nonzero number $(\alpha - \beta)(h_j)$ is positive. As in §4.2, the root vectors corresponding to positive roots will play the role of raising operators in the representation theory of G. We first consider as a special case the adjoint representation of G on $\mathfrak{g}$. This will give some valuable information on the structure of $\mathfrak{g}$.

To begin, associate to each root $\alpha \in \mathfrak{h}'$ an element H_α uniquely determined by

$$H_\alpha \in \mathfrak{h}, \quad \alpha(h) = \langle H_\alpha, h\rangle, \quad \forall\, h \in \mathfrak{h}. \tag{6.1.6}$$

Here $\langle\ ,\ \rangle$ is the Ad-invariant inner product on $\mathfrak{g}$ mentioned above (restricted in (6.1.6) to an inner product on $\mathfrak{h}$). Next, extend $\langle\ ,\ \rangle$ to a symmetric *bilinear* form on $\mathfrak{g}_{\mathbb{C}} \times \mathfrak{g}_{\mathbb{C}}$. We have the following.

Lemma 6.1.1. *If $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, then*

$$[X, Y] = i\langle X, Y\rangle H_\alpha. \tag{6.1.7}$$

Proof. By (6.1.5), $[X, Y] \in \mathfrak{h}_{\mathbb{C}} = \mathfrak{g}_0$. Now, for any $H \in \mathfrak{h}$,

$$\langle [X, Y], H\rangle = \langle Y, [H, X]\rangle = i\alpha(H)\langle Y, X\rangle, \tag{6.1.8}$$

and, by (6.1.6), this identity is equivalent to (6.1.7). $\square$

Note also that

$$\alpha(H_\alpha) = \langle H_\alpha, H_\alpha\rangle > 0. \tag{6.1.9}$$

We are ready for the following key result.

Proposition 6.1.2. *For each root α,*

$$\dim \mathfrak{g}_\alpha = 1. \tag{6.1.10}$$

Proof. Assume $\dim \mathfrak{g}_\alpha \geq 2$. We will show that $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are not orthogonal. Granted this, we can pick $X, Z \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$ such that

$$\langle X, Y\rangle = 1, \quad \langle Z, Y\rangle = 0. \tag{6.1.11}$$

This implies

$$[X, Y] = iH_\alpha, \quad [Z, Y] = 0. \tag{6.1.12}$$

From here an inductive argument shows

$$\operatorname{ad} Y (\operatorname{ad} X)^n Z = \frac{n(n+1)}{2}\,\alpha(H_\alpha)(\operatorname{ad} X)^{n-1} Z. \tag{6.1.13}$$

To do the induction, we first check (6.1.13) for $n = 1$. Indeed,

$$\begin{aligned}\operatorname{ad} Y (\operatorname{ad} X) Z &= (\operatorname{ad} X)(\operatorname{ad} Y) Z - [\operatorname{ad} X, \operatorname{ad} Y] Z\\ &= 0 - i\ \operatorname{ad} H_\alpha(Z)\\ &= \alpha(H_\alpha) Z,\end{aligned}$$

since

$$[\mathrm{ad}\, X, \mathrm{ad}\, Y] = \mathrm{ad}[X, Y] = i\ \mathrm{ad}(H_\alpha), \tag{6.1.14}$$

$$\mathrm{ad}\ H_\alpha = i\beta(H_\alpha)\mathrm{Id}, \quad \text{on} \ \ \mathfrak{g}_\beta, \tag{6.1.15}$$

and $Z \in \mathfrak{g}_\alpha$. Now, given $n \geq 2$, assume (6.1.13) holds with n replaced by $n-1$. Then

(6.1.16)
$$\begin{aligned} \mathrm{ad}\, Y(\mathrm{ad}\, X)^n Z &= \mathrm{ad}\, X(\mathrm{ad}\, Y)(\mathrm{ad}\, X)^{n-1} Z - [\mathrm{ad}\, X, \mathrm{ad}\, Y](\mathrm{ad}\, X)^{n-1} Z \\ &= \frac{(n-1)n}{2}\alpha(H_\alpha)(\mathrm{ad}\, X)^{n-1} - i\ \mathrm{ad}(H_\alpha)(\mathrm{ad}\, X)^{n-1} Z, \end{aligned}$$

the second identity by the inductive hypothesis and (6.1.14). Now $(\mathrm{ad}\, X)^{n-1} Z \in \mathfrak{g}_{n\alpha}$, so taking $\beta = n\alpha$ in (6.1.15) gives

$$-i\ \mathrm{ad}(H_\alpha)(\mathrm{ad}\, X)^{n-1} Z = n\alpha(H_\alpha)(\mathrm{ad}\, X)^{n-1} Z, \tag{6.1.17}$$

so

$$\mathrm{ad}\, Y(\mathrm{ad}\, X)^n Z = \Big[\frac{(n-1)n}{2} + n\Big]\alpha(H_\alpha)(\mathrm{ad}\, X)^{n-1} Z, \tag{6.1.18}$$

yielding (6.1.13).

By (6.1.9) and (6.1.13), it follows that, if $\dim\, \mathfrak{g}_\alpha \geq 2$, then all the elements $(\mathrm{ad}\, X)^n Z \in \mathfrak{g}_{(n+1)\alpha}$ are nonzero. This contradicts the fact that $\mathfrak{g}$ is finite dimensional.

To complete the proof of Proposition 6.1.2, it remains to show that $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are not orthogonal with respect to $\langle\ ,\ \rangle$. Indeed,

(6.1.19)
$$\begin{aligned} x + iy \in \mathfrak{g}_\alpha\ (x, y \in \mathfrak{g}) &\Longrightarrow x - iy \in \mathfrak{g}_{-\alpha} \\ &\Longrightarrow \langle x + iy, x - iy\rangle = \langle x, x\rangle + \langle y, y\rangle, \end{aligned}$$

which is > 0 as long as $x + iy \neq 0$. This completes the proof of Proposition 6.1.2. □

Remark. The endgame of this last proof uses the fact that, if α is a root, then $n\alpha$ cannot be a root for all $n \in \mathbb{N}$. See Proposition 6.2.4 for a much stronger result.

Given Proposition 6.1.2, we can pick nonzero vectors $e_\alpha \in \mathfrak{g}_\alpha$, and arrange that $e_{-\alpha}$ be the complex conjugate of e_α,

$$e_{\pm\alpha} = x_\alpha \pm i y_\alpha, \quad x_\alpha, y_\alpha \in \mathfrak{g}. \tag{6.1.20}$$

Furthermore, we can scale these elements so that $\langle e_\alpha, e_{-\alpha}\rangle = 1$. Thus

$$[e_\alpha, e_{-\alpha}] = iH_\alpha. \tag{6.1.21}$$

These commutation relations, together with $[H_\alpha, e_{\pm\alpha}] = \pm i\alpha(H_\alpha)e_{\pm\alpha}$ are equivalent to
(6.1.22)
$$[x_\alpha, y_\alpha] = -\frac{1}{2}H_\alpha, \quad [H_\alpha, x_\alpha] = -\alpha(H_\alpha)y_\alpha, \quad [H_\alpha, y_\alpha] = \alpha(H_\alpha)x_\alpha.$$

The following result bears on the size of the linear span of the set of roots α in $\mathfrak{h}'$.

Proposition 6.1.3. *We have*

$$\bigcap_{\alpha\ \mathit{root}} \ker\alpha = \mathfrak{z}, \tag{6.1.23}$$

the center of $\mathfrak{g}$.

Proof. Given $h \in \mathfrak{h}$,

$$\begin{aligned} \alpha(h) = 0\ \forall\, \alpha &\Longleftrightarrow \operatorname{ad} h = 0 \ \text{ on each } \mathfrak{g}_\alpha \\ &\Longleftrightarrow h \in \mathfrak{z}. \end{aligned} \tag{6.1.24}$$

Of course, $\mathfrak{z} \subset \mathfrak{h}$, so this gives (6.1.23). Note that an equivalent statement is that (with the orthogonal complement taken in $\mathfrak{h}$)

$$\bigcap_{\alpha\ \text{root}} (\text{Span } H_\alpha)^\perp = \mathfrak{z}.$$

$\square$

Corollary 6.1.4. *If* $\mathfrak{z} = 0$, *then the set of roots spans* $\mathfrak{h}'$; *hence* $\{H_\alpha : \alpha \text{ is a root}\}$ *spans* $\mathfrak{h}$.

Examples. If $\mathfrak{g} = \mathfrak{u}(n)$, then $\mathfrak{z} = \{iaI : a \in \mathbb{R}\}$. If $\mathfrak{g} = \mathfrak{su}(n)$, then $\mathfrak{z} = 0$. If $\mathfrak{g} = \mathfrak{so}(n)$, then $\mathfrak{z} = 0$.

We turn now to the representation theory of G. Let π be a unitary representation of G on a finite dimensional complex vector space V. This gives rise to a representation $d\pi$ of $\mathfrak{g}$ by skew adjoint operators on V, which extends to a complex linear representation, also denoted $d\pi$ of $\mathfrak{g}_\mathbb{C}$ on V. As in §4.2, we will find it convenient to bring in the complexification $G_\mathbb{C}$ of G, and use the following fact:

Proposition 6.1.5. *The representation* π *of* G *on* V *extends to a holomorphic representation of* $G_\mathbb{C}$ *on* V.

See §6.5 for a description of $G_\mathbb{C}$ and a proof of Proposition 6.1.5.

To pursue our analysis of the representation π, take a maximal torus $\mathbb{T}$ of G as above, with Lie algebra $\mathfrak{h}$, and for $\lambda \in \mathfrak{h}'$ set

$$V_\lambda = \{v \in V : d\pi(h)v = i\lambda(h)v, \ \forall\, h \in \mathfrak{h}\}. \tag{6.1.25}$$

We have

$$V = \bigoplus_\lambda V_\lambda. \tag{6.1.26}$$

If $V_\lambda \neq 0$ we call λ a *weight*, and any nonzero $v \in V_\lambda$ a *weight vector*. The decomposition (6.1.25)–(6.1.26) is called the weight space decomposition of V.

For the root vectors e_α considered above, set

$$E_\alpha = d\pi(e_\alpha). \tag{6.1.27}$$

We call E_α a *raising operator* if $\alpha > 0$ and a lowering operator if $\alpha < 0$. The commutation relations

$$[h, e_\alpha] = i\alpha(h)e_\alpha, \quad \forall\, h \in \mathfrak{h} \tag{6.1.28}$$

imply

$$d\pi(h)E_\alpha = E_\alpha d\pi(h) + i\alpha(h)E_\alpha. \tag{6.1.29}$$

Using this we can prove the following.

Proposition 6.1.6. *For each root α, we have*

$$E_\alpha : V_\lambda \longrightarrow V_{\lambda+\alpha}. \tag{6.1.30}$$

In particular, if λ is a weight and α is a root, then either E_α annihilates V_λ or $\lambda + \alpha$ is a weight.

Proof. If $\xi \in V_\lambda$, we have, for all $h \in \mathfrak{h}$,

$$\begin{aligned} d\pi(h)(E_\alpha\xi) &= E_\alpha d\pi(h)\xi + i\alpha(h)E_\alpha\xi \\ &= i\big(\lambda(h) + \alpha(h)\big)E_\alpha\xi, \end{aligned} \tag{6.1.31}$$

which proves the proposition. $\square$

The ordering we have put on $\mathfrak{h}'$ induces an ordering on the weights. For a given finite dimensional representation π, with respect to this ordering there will be a *highest weight* λ_m, and also a lowest weight λ_s. From Proposition 6.1.6 we see that

$$\begin{aligned} &E_\alpha = 0 \text{ on } V_{\lambda_m}, \text{ for all raising operators } E_\alpha, \\ &E_\alpha = 0 \text{ on } V_{\lambda_s}, \text{ for all lowering operators } E_\alpha. \end{aligned} \tag{6.1.32}$$

In general, call a weight λ nonraisable if V_λ is annihilated by all raising operators and call it nonlowerable if V_λ is annihilated by all lowering operators. Later in this section we will show that if π is irreducible, then the only nonraisable weight is maximal. Here we record our progress up to this point.

Proposition 6.1.7. *If π is a unitary representation of the compact Lie group G in a finite dimensional space V, then there exists a highest weight vector ξ, and in particular there exists a nonzero weight vector $\xi \in V$ annihilated by all raising operators.*

This result gives a tool for showing that certain representations of G are irreducible, namely:

Corollary 6.1.8. *Let π be a unitary representation of G on a finite dimensional space V. Suppose the set of weight vectors $\xi \in V$ annihilated by all raising operators is equal to the set of nonzero multiples of a single element. Then π is irreducible.*

Proof. Otherwise, $V = V_1 \oplus V_2$ with π acting on each factor, and Proposition 6.1.5 produces two linearly independent weight vectors $\xi_j \in V_j$, annihilated by all raising operators. □

We note the following. Set

$$\mathcal{H}(\pi) = \bigcap_{\alpha>0} \operatorname{Ker} E_\alpha. \tag{6.1.33}$$

From (6.1.29) it follows that

$$h \in \mathfrak{h} \Longrightarrow d\pi(h) : \mathcal{H}(\pi) \to \mathcal{H}(\pi), \tag{6.1.34}$$

and of course $\{d\pi(h)|_{\mathcal{H}(\pi)} : h \in \mathfrak{h}\}$ forms a commuting family of skew-adjoint operators, so they are simultaneously diagonalizable on $\mathcal{H}(\pi)$, i.e., $\mathcal{H}(\pi)$ is spanned by weight vectors. Hence the hypothesis in Corollary 6.1.8 is equivalent to the hypothesis that $\dim \mathcal{H}(\pi) = 1$.

We now head for a circle of results that include a converse to Corollary 6.1.8, parallel to Propositions 4.2.4–4.2.5. Let $\overline{\pi}$ denote the representation of G on V' contragredient to π, given by

$$\langle \xi, \overline{\pi}(g)\eta \rangle = \langle \pi(g^{-1})\xi, \eta \rangle, \quad \xi \in V, \ \eta \in V'. \tag{6.1.35}$$

Suppose $\xi_0 \in V$ is a nonraisable weight vector for π, with weight $\lambda \in \mathfrak{h}'$, and suppose $\eta_0 \in V'$ is a nonlowerable weight vector for $\overline{\pi}$, with weight $-\mu \in \mathfrak{h}'$. As in §4.2, we will study the function

$$\varphi(g) = \langle \pi(g)\xi_0, \eta_0 \rangle. \tag{6.1.36}$$

As stated in Proposition 6.1.5, we can extend π to a holomorphic representation of the complexified group $G_{\mathbb{C}}$, which then extends φ to a holomorphic function on $G_{\mathbb{C}}$.

Write the complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$ as

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus \mathfrak{n}_+ \oplus \mathfrak{n}_-, \quad \mathfrak{n}_+ = \bigoplus_{\alpha>0} \mathfrak{g}_\alpha, \quad \mathfrak{n}_- = \bigoplus_{\alpha<0} \mathfrak{g}_\alpha. \tag{6.1.37}$$

Let D, N_+, and N_- denote the Lie subgroups of $G_{\mathbb{C}}$ with Lie algebras $\mathfrak{h}_{\mathbb{C}}$, $\mathfrak{n}_+$, and $\mathfrak{n}_-$, respectively. It follows from the inverse function theorem that

$$N_- D N_+ = G_{\mathrm{reg}} \tag{6.1.38}$$

is a subset of $G_{\mathbb{C}}$ that contains an open neighborhood of the identity element e. Let

$$g = \zeta\delta z, \quad \zeta \in N_-, \ z \in N_+, \quad \delta = \exp(h) = \exp(h_1 + ih_2) \in D, \tag{6.1.39}$$

with $h_j \in \mathfrak{h}$. We see that

$$\begin{aligned} \pi(z)\xi_0 = \xi_0, &\quad \pi(\delta)\xi_0 = e^{i\lambda(h)}\xi_0, \\ \overline{\pi}(\zeta)\eta_0 = \eta_0, &\quad \overline{\pi}(\delta^{-1})\eta_0 = e^{i\mu(h)}\eta_0. \end{aligned} \tag{6.1.40}$$

Consequently,

$$\varphi(\zeta g) = \varphi(g), \quad \varphi(\delta g) = e^{i(\mu(h_1)+i\mu(h_2))}\varphi(g), \tag{6.1.41}$$

and

$$\varphi(gz) = \varphi(g), \quad \varphi(g\delta) = e^{i(\lambda(h_1)+i\lambda(h_2))}\varphi(g). \tag{6.1.42}$$

Hence

$$\begin{aligned} \varphi(\zeta\delta z) = \varphi(\delta) &= e^{i(\lambda(h_1)+i\lambda(h_2))}\varphi(e) \\ &= e^{i(\mu(h_1)+i\mu(h_2))}\varphi(e). \end{aligned} \tag{6.1.43}$$

This identity is very significant, in light of the following result.

Lemma 6.1.9. *Assume π is irreducible. Then the function φ has the property*

$$\varphi(e) \neq 0. \tag{6.1.44}$$

Proof. If $\varphi(e) = 0$, then (6.1.43) implies that $\varphi(g) = 0$ on G_{reg}. Since φ is holomorphic and G_{reg} contains a neighborhood of g, it follows that $\varphi \equiv 0$ on G. However, if π is irreducible and $\xi_0 \neq 0$ then

$$\mathrm{Span}\{\pi(g)\xi_0 : g \in G\} \tag{6.1.45}$$

is invariant, hence all of V, so $\varphi \equiv 0$ cannot hold. This proves the lemma. $\square$

From (6.1.43) and the lemma, we can deduce the following important result.

Theorem 6.1.10. *If π is irreducible on V, the only weight λ that is nonraisable is the highest weight. Furthermore, the highest weight vector is unique, up to a scalar multiple. Finally, if π and π_2 are irreducible representations with the same highest weight, they are unitarily equivalent.*

Proof. The identity $\lambda = \mu$ (a consequence of (6.1.43)–(6.1.44)) proves the uniqueness of λ, and establishes the first assertion. To proceed, note that if we normalize the weight vectors so $\varphi(e) = 1$, the function $\varphi(g)$ is uniquely characterized by the following three properties:

$$\varphi \ \text{ is holomorphic on } \ G_{\mathbb{C}}, \tag{6.1.46}$$

$$\varphi(\zeta g z) = \varphi(g), \quad \forall\, \zeta \in N_-,\ z \in N_+,\ g \in G, \tag{6.1.47}$$

$$\varphi(\delta) = e^{i(\lambda(h_1)+i\lambda(h_2))}, \quad \forall\, \delta = \exp(h_1 + ih_2) \in D. \tag{6.1.48}$$

Thus, if ξ_1 were another highest weight vector, also normalized so $\langle \xi_1, \eta_0 \rangle = 1$, we would have

$$\langle \pi(g)\xi_1, \eta_0 \rangle = \varphi(g), \quad \forall\, g \in G, \tag{6.1.49}$$

so $\langle \pi(g)(\xi_1 - \xi_0), \eta_0 \rangle = 0$ for all g, hence

$$W = \operatorname{Span}\{\pi(g)(\xi_1 - \xi_0) : g \in G\} \perp \eta_0. \tag{6.1.50}$$

Since $\pi(g) : W \to W$ and π is irreducible, this gives $\xi_1 = \xi_0$.

As for the final assertion of Theorem 6.1.10, let π_2 be an irreducible representation on V_2, with the same highest weight λ as π. Pick a maximal weight vector ξ_2 for π_2 and a minimal weight vector η_2 for its contragredient representation $\overline{\pi}_2$, normalized so $\langle \xi_2, \eta_2 \rangle = 1$, and form

$$\varphi_2(g) = \langle \pi_2(g)\xi_2, \eta_2 \rangle. \tag{6.1.51}$$

Then φ_2 also satisfies the conditions (6.1.46)–(6.1.48). Hence $\varphi \equiv \varphi_2$. Hence π_2 must be equivalent to π, since otherwise the Weyl orthogonality relations would imply that φ and φ_2 are orthogonal in $L^2(G)$. $\square$

6.2. Roots and weights for compact G, II: injections $\mathfrak{su}(2) \hookrightarrow \mathfrak{g}$

Recall from (6.1.20)–(6.1.22) the construction of $e_{\pm\alpha} = x_\alpha \pm iy_\alpha$, spanning $\mathfrak{g}_{\pm\alpha}$, satisfying

$$[x_\alpha, y_\alpha] = -\frac{1}{2}H_\alpha, \quad [H_\alpha, x_\alpha] = -\alpha(H_\alpha)y_\alpha, \quad [H_\alpha, y_\alpha] = \alpha(H_\alpha)x_\alpha, \tag{6.2.1}$$

with $H_\alpha \in \mathfrak{h}$ given by (6.1.6). This holds for each root α. Recall that $\alpha(H_\alpha) = \langle H_\alpha, H_\alpha\rangle > 0$. If we take the inner product on $\mathfrak{h}'$ induced by that on $\mathfrak{h}$, we also have

$$\alpha(H_\alpha) = \langle \alpha, \alpha\rangle, \tag{6.2.2}$$

and more generally $\lambda(H_\alpha) = \langle \lambda, \alpha\rangle$ for each $\lambda \in \mathfrak{h}'$. Let us set

$$X_1^\alpha = \frac{1}{\langle \alpha,\alpha\rangle}H_\alpha, \quad X_2^\alpha = \sqrt{\frac{2}{\langle \alpha,\alpha\rangle}}\,y_\alpha, \quad X_3^\alpha = \sqrt{\frac{2}{\langle \alpha,\alpha\rangle}}\,x_\alpha. \tag{6.2.3}$$

Then the commutation relations (6.2.1) are equivalent to

$$[X_1^\alpha, X_2^\alpha] = X_3^\alpha, \quad [X_2^\alpha, X_3^\alpha] = X_1^\alpha, \quad [X_3^\alpha, X_1^\alpha] = X_2^\alpha. \tag{6.2.4}$$

Now these commutation relations are identical to those in (4.1.2). Hence each root α gives rise to an injective Lie algebra homomorphism $\mathfrak{su}(2) \hookrightarrow \mathfrak{g}$, which in turn, since SU(2) is simply connected, exponentiates to a Lie group homomorphism

$$\gamma^\alpha : \mathrm{SU}(2) \longrightarrow G, \tag{6.2.5}$$

defined for each root α. Since $d\gamma^\alpha$ is injective, either γ^α is injective or $\operatorname{Ker}\gamma^\alpha = \{\pm I\}$, the only proper normal subgroup of SU(2).

The homomorphisms (6.2.5) have implications for the behavior of a unitary representation π of G (say on V). In fact, given such π, the composition $\pi^\alpha = \pi \circ \gamma^\alpha$ is a unitary representation of SU(2), and the material of §4.1 applies. Suppose λ is a weight of π, with weight space $V_\lambda \subset V$. Then

$$\begin{aligned} v \in V_\lambda &\Longrightarrow \pi(\operatorname{Exp} tX_1^\alpha)v = e^{it\lambda(X_1^\alpha)}v \\ &\Longrightarrow \pi \circ \gamma^\alpha(e^{tX_1})v = e^{it\lambda(X_1^\alpha)}v \\ &\Longrightarrow d\pi^\alpha(X_1)v = i\lambda(X_1^\alpha)v. \end{aligned} \tag{6.2.6}$$

Results of §4.1, analyzing (4.1.1), imply that $\lambda(X_1^\alpha) = n/2$ for some $n \in \mathbb{Z}$, hence

$$\frac{\lambda(H_\alpha)}{\langle \alpha,\alpha\rangle} = \frac{\langle \lambda,\alpha\rangle}{\langle \alpha,\alpha\rangle} = \frac{n}{2}, \quad \text{for some} \quad n \in \mathbb{Z}. \tag{6.2.7}$$

Note that

$$d\gamma^\alpha(X_2 \mp iX_3) = \mp i\sqrt{\frac{2}{\langle\alpha,\alpha\rangle}}(x_\alpha \pm iy_\alpha) = \mp i\sqrt{\frac{2}{\langle\alpha,\alpha\rangle}}e_{\pm\alpha}, \tag{6.2.8}$$

hence, taking into account (4.1.10), we see that if V_λ is annihilated by all raising operators for the representation π of G, then it is annihilated by the raising operator for the representation $\pi \circ \gamma^\alpha$ of SU(2), for each $\alpha > 0$. This forces $n \geq 0$ in (6.2.7), for $\alpha > 0$. We record the result.

Proposition 6.2.1. *Let π be a unitary representation of G on V. Then for each root α of $\mathfrak{g}$ and each weight λ of π,*

$$2\frac{\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} \tag{6.2.9}$$

is an integer. If $\alpha > 0$ and V_λ is annihilated by all raising operators (e.g., if λ is a highest weight), then (6.2.9) *is a non-negative integer.*

Example 1. Take $G = \mathrm{U}(n)$, take the basis $\{e_j : 1 \leq j \leq n\}$ of $\mathfrak{h}$ given by (4.2.9), which is orthonormal with respect to the Ad-invariant Hilbert-Schmidt inner product on $\mathfrak{g} = \mathfrak{u}(n)$. This then defines an order on $\mathfrak{h}$, and an order and an inner product on $\mathfrak{h}'$. The roots are ω_{jk}, given by (4.2.14), which are positive provided $j < k$. An element $\lambda \in \mathfrak{h}'$ is given by $\lambda = (d_1, \dots, d_n)$. We have $\langle\omega_{jk}, \omega_{jk}\rangle = 2$ and

$$2\frac{\langle\lambda,\omega_{jk}\rangle}{\langle\omega_{jk},\omega_{jk}\rangle} = d_j - d_k. \tag{6.2.10}$$

This is non-negative for all positive roots if and only if $d_1 \geq \cdots \geq d_n$. For the right side to be an integer for all $j \neq k$, it is sufficient (but not necessary) that all d_j be integers. Compare the characterization of highest weights in Theorem 4.4.1.

Example 2. Take $G = \mathrm{SU}(n)$. If $\mathfrak{u}(n)_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus (\oplus_{j\neq k}\mathfrak{g}_{\omega_{jk}})$, then $\mathfrak{su}(n)_{\mathbb{C}}$ has the same form, with $\mathfrak{h}$ replaced by $\tilde{\mathfrak{h}}$, the codimension one subspace of $\mathfrak{h}$ defined by

$$\tilde{\mathfrak{h}} = \{h \in \mathfrak{h} : \mathrm{Tr}\, h = 0\}. \tag{6.2.11}$$

It is natural to take the order on $\tilde{\mathfrak{h}}$ induced from that on $\mathfrak{h}$ via inclusion. The roots for $\mathfrak{su}(n)$ are the restrictions to $\tilde{\mathfrak{h}}$ of the elements ω_{jk}, and the root spaces are still the one-dimensional spans of the elements e_{jk}, for each $j \neq k$. We have

$$\tilde{\mathfrak{h}}' = \mathfrak{h}'/\{(r, \dots, r) : r \in \mathbb{R}\}. \tag{6.2.12}$$

The weights are equivalence classes $(d_1, \dots, d_n) \sim (d_1 + r, \dots, d_n + r)$, and (6.2.10) holds in this context; note that $d_j - d_k = (d_j + r) - (d_k + r)$. Again the condition that (6.2.10) be ≥ 0 whenever $j < k$ becomes $d_1 \geq \cdots \geq d_n$. If we pick $r = -d_n$, then the representative of λ is $(d_1, \dots, d_{n-1}, 0)$, satisfying $d_j \in \mathbb{Z}^+,\ d_1 \geq \cdots \geq d_{n-1}$. Compare the description of the highest weights for the irreducible representations of $\mathrm{SU}(n)$ in Proposition 4.5.2.

REMARK. In Example 2 we see that the necessary condition given in Proposition 6.2.1 for an element $\lambda \in \mathfrak{h}'$ to be the highest weight for some irreducible representation of $\mathrm{SU}(n)$ is also sufficient. By contrast, in Example 1 the necessary condition given in Proposition 6.2.1 is not quite sufficient, since these conditions do not imply that the entries d_j be integers (only that their differences be integers). It turns out that what is behind this dichotomy is that the Lie algebra of $\mathrm{SU}(n)$ has a trivial center, while the center of the Lie algebra of $\mathrm{U}(n)$ is $\{iaI : a \in \mathbb{R}\}$, which is nontrivial. The following result completes Proposition 6.2.1.

Theorem of the Highest Weight. If G is a compact, simply connected Lie group whose Lie algebra $\mathfrak{g}$ has a trivial center, then the condition that (6.2.9) be a non-negative integer for each positive root α is necessary and sufficient for a given $\lambda \in \mathfrak{h}'$ to be the highest weight of some irreducible representation of G. One calls such λ a *dominant integral weight.*

A proof can be found in Chapter 4 of [**47**]. We outline an approach to obtaining such a proof. Namely, one produces a certain finite set $\{\lambda_1, \dots, \lambda_K\} \subset \mathfrak{h}'$ of dominant integral weights (called "fundamental weights"), with the property that each dominant integral weight λ has the form

$$\lambda = n_1\lambda_1 + \cdots + n_K\lambda_K, \quad n_j \in \mathbb{Z}^+.$$

Then one exhibits irreducible unitary representations π_j of G with highest weight $\lambda_j,\ 1 \leq j \leq K$. Once one has this, the Theorem is a consequence of the following.

Proposition 6.2.2. *Suppose π_j is a unitary representation of G on V_j with highest weight μ_j (with highest weight vector $v_j \in V_j$). Then the representation*

$$\pi_1 \otimes \pi_2 \quad on \quad V_1 \otimes V_2$$

has highest weight $\mu_1 + \mu_2$ (with highest weight vector $v_1 \otimes v_2$).

Proof. Same as for Proposition 4.4.4. $\square$

Recall that this was the program used in §4.4 to classify the irreducible representations of $\mathrm{U}(n)$.

We turn our attention to the adjoint representation of G on $\mathfrak{g}_{\mathbb{C}}$. If we apply Proposition 6.2.1 to the adjoint representation, we get:

Corollary 6.2.3. *If α and β are roots, then*

$$n_{\alpha\beta} = 2\frac{\langle \alpha, \beta\rangle}{\langle \beta, \beta\rangle} \in \mathbb{Z}. \tag{6.2.13}$$

The integers $n_{\alpha\beta}$ are called the Cartan integers.

Remark 1. The orthogonal projection of β onto the linear span of α in $\mathfrak{h}'$ is given by

$$P_\alpha \beta = \frac{\langle \alpha, \beta\rangle}{\langle \alpha, \alpha\rangle}\alpha. \tag{6.2.14}$$

Hence Corollary 6.2.3 impacts the geometry of the roots, as a subset of $\mathfrak{h}'$. We look further at this impact.

Remark 2. Of course, one can reverse the roles of α and β in (6.2.13). Comparing the results implies the following. If $\theta_{\alpha\beta}$ denotes the angle between α and β in $\mathfrak{h}'$, then

$$\cos^2 \theta_{\alpha\beta} = \frac{n_{\alpha\beta} n_{\beta\alpha}}{4}, \tag{6.2.15}$$

and hence, since the numerator must be an integer,

$$\cos^2 \theta_{\alpha\beta} \in \Big\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\Big\}. \tag{6.2.16}$$

It also follows that

$$n_{\alpha\beta} \in \{0, \pm 1, \pm 2, \pm 3, \pm 4\}. \tag{6.2.17}$$

More precisely, we have the following. Assume

$$\langle \alpha, \beta\rangle \neq 0, \quad \langle \alpha, \alpha\rangle \geq \langle \beta, \beta\rangle \quad (\text{so } n_{\alpha\beta} \geq n_{\beta\alpha}). \tag{6.2.18}$$

Then, with $\sigma = \pm 1$,

$$\begin{aligned} \cos^2 \theta_{\alpha\beta} = \frac{1}{4} &\Longleftrightarrow n_{\alpha\beta} = n_{\beta\alpha} = \sigma \\ &\Longrightarrow \langle \alpha, \alpha\rangle = \langle \beta, \beta\rangle, \end{aligned} \tag{6.2.19}$$

$$\begin{aligned} \cos^2 \theta_{\alpha\beta} = \frac{1}{2} &\Longleftrightarrow n_{\alpha\beta} = 2\sigma, \ n_{\beta\alpha} = \sigma \\ &\Longrightarrow \langle \alpha, \alpha\rangle = 2\langle \beta, \beta\rangle, \end{aligned} \tag{6.2.20}$$

$$\begin{aligned} \cos^2 \theta_{\alpha\beta} = \frac{3}{4} &\Longleftrightarrow n_{\alpha\beta} = 3\sigma, \ n_{\beta\alpha} = \sigma \\ &\Longrightarrow \langle \alpha, \alpha\rangle = 3\langle \beta, \beta\rangle. \end{aligned} \tag{6.2.21}$$

Here is another restriction on the set of roots.

Proposition 6.2.4. *If α is a root and also $\beta = s\alpha$ is a root, for some $s \in \mathbb{R} \setminus 0$, then $s = \pm 1$.*

Proof. Interchanging the roles of α and β and changing the sign of β if necessary, we see it suffices to show that if α is a root and $0 < s < 1$, then $s\alpha$ is not a root. If such $s\alpha$ were a root, (6.2.13) would imply $2s \in \mathbb{Z}$. This forces $s = 1/2$, i.e., $\beta = (1/2)\alpha$, or $\alpha = 2\beta$. Thus it suffices to show:

$$\text{If } \beta \text{ is a root, then } 2\beta \text{ is not a root.} \tag{6.2.22}$$

To see this, consider

$$W_\beta = \mathbb{C}\text{-Span}(H_\beta) \oplus \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{k\beta}. \tag{6.2.23}$$

Then W_β is invariant under $\operatorname{Ad} \circ \gamma^\beta$, i.e., we have a representation π^β of $\mathrm{SU}(2)$ on W_β. This representation splits into an orthogonal direct sum of irreducible pieces, each isomorphic to a representation of the form $D_{\ell/2}$, given in Proposition 4.1.2, having weight space decomposition

$$V_{-\ell/2} \oplus V_{-\ell/2+1} \oplus \cdots \oplus V_{\ell/2}, \quad dD_{\ell/2}(X_1) = i\mu \ \text{ on } \ V_\mu. \tag{6.2.24}$$

Now if we take $\pi = \operatorname{Ad}$ in (6.2.6) (and replace α by β, and λ by $k\beta$), we get

$$\begin{aligned} \pi^\beta(g) = \operatorname{Ad}(\gamma^\beta(g)) \Rightarrow d\pi^\beta(X_1) &= \operatorname{ad} d\gamma^\beta(X_1) = \operatorname{ad} X_1^\beta \\ &= \frac{1}{\langle \beta, \beta \rangle} \operatorname{ad} H_\beta \quad (\text{by (6.2.3)}) \\ \Rightarrow d\pi^\beta(X_1) &= ik \ \text{ on } \ \mathfrak{g}_{k\beta}, \end{aligned} \tag{6.2.25}$$

since

$$\operatorname{ad} H_\beta\big|_{\mathfrak{g}_{k\beta}} = ik\beta(H_\beta)\mathrm{Id} = ik\langle \beta, \beta \rangle \mathrm{Id}.$$

Hence the only representations $D_{\ell/2}$ of $\mathrm{SU}(2)$ that occur in the decomposition of π^β are those for which ℓ is even. Each one of these has a copy of V_0, on which $d\pi^\beta(X_1) = 0$. However, comparing (6.2.23) and (6.2.24), we see that this occurs only as the one-dimensional space $\mathrm{Span}(H_\beta)$. Hence π^β is irreducible. Since $\mathrm{Span}(H_\beta) \oplus \mathfrak{g}_\beta \oplus \mathfrak{g}_{-\beta}$ is invariant under $\pi^\beta = \operatorname{Ad} \circ \gamma^\beta$, we must have (6.2.23) equal to $\mathrm{Span}(H_\beta) \oplus \mathfrak{g}_\beta \oplus \mathfrak{g}_{-\beta}$. This establishes (6.2.22) and completes the proof of Proposition 6.2.4. ∎

In light of Proposition 6.2.4, we can complement (6.2.19)–(6.2.21) with

$$\begin{aligned} \cos^2 \theta_{\alpha\beta} = 1 \Longleftrightarrow \alpha &= \sigma\beta \\ \Longrightarrow n_{\alpha\beta} &= n_{\beta\alpha} = 2\sigma, \end{aligned} \tag{6.2.26}$$

where again $\sigma = \pm 1$. Also, we can sharpen (6.2.17) to

$$n_{\alpha\beta} \in \{0, \pm 1, \pm 2, \pm 3\}. \tag{6.2.27}$$

To see this, note that $n_{\alpha\beta} = \pm 4 \Rightarrow n_{\beta\alpha} = \pm 1 \Rightarrow \langle \beta, \beta \rangle = 4\langle \alpha, \alpha \rangle$ and $\cos^2 \theta_{\alpha\beta} = 1$, but these last two identities contradict (6.2.26).

We also note that, given $|\alpha| \geq |\beta|$ and $\beta \neq \pm\alpha$, then, by (6.2.19)–(6.2.21),

$$\begin{aligned} n_{\alpha\beta} = 1 \Leftrightarrow n_{\beta\alpha} = 1, &\quad n_{\alpha\beta} = -1 \Leftrightarrow n_{\beta\alpha} = -1, \\ n_{\alpha\beta} = 2 \Leftrightarrow n_{\beta\alpha} = 1, &\quad n_{\alpha\beta} = -2 \Leftrightarrow n_{\beta\alpha} = -1, \\ n_{\alpha\beta} = 3 \Leftrightarrow n_{\beta\alpha} = 1, &\quad n_{\alpha\beta} = -3 \Leftrightarrow n_{\beta\alpha} = -1, \end{aligned} \tag{6.2.28}$$

Here is another perspective on the numbers in (6.2.9). Assume λ is a weight for the unitary representation π of G on V, take $\varphi_0 \in V_\lambda$, let α be a root of $\mathfrak{g}$, and, with $E_{\pm\alpha} = d\pi(e_{\pm\alpha})$, assume

$$\begin{aligned} (E_\alpha)^j \varphi_0 = \varphi_j \neq 0 \ \text{ in } \ V_{\lambda + j\alpha}, &\quad \text{for } \ 0 \leq j \leq p, \\ (E_{-\alpha})^j \varphi_0 = \varphi_{-j} \neq 0 \ \text{ in } \ V_{\lambda - j\alpha}, &\quad \text{for } \ 0 \leq j \leq m, \end{aligned} \tag{6.2.29}$$

while

$$E_\alpha \varphi_p = 0 = E_{-\alpha} \varphi_{-m}. \tag{6.2.30}$$

We call $\{\varphi_j : -m \leq j \leq p\}$ an α-string.

For such a string, we set

$$\psi_0 = (E_\alpha)^p \varphi_0 \in V_{\lambda^*}, \quad \lambda^* = \lambda + p\alpha, \tag{6.2.31}$$

and we have

$$\begin{aligned} &(E_{-\alpha})^k \psi_0 = \psi_k \neq 0 \ \text{ in } \ V_{\lambda^* - k\alpha}, \quad \text{for } \ 0 \leq k \leq q = m + p, \\ &\quad E_{-\alpha} \psi_q = 0, \quad E_\alpha \psi_0 = 0. \end{aligned} \tag{6.2.32}$$

We have the following important identity.

Proposition 6.2.5. *Given an α-string, as described above,*

$$m + p = 2\, \frac{\langle \lambda + p\alpha, \alpha \rangle}{\langle \alpha, \alpha \rangle}, \tag{6.2.33}$$

hence

$$m - p = 2\, \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle}. \tag{6.2.34}$$

Proof. Equivalently, in the setup (6.2.31)–(6.2.32), we claim

$$q = 2 \frac{\langle \lambda^*, \alpha \rangle}{\langle \alpha, \alpha \rangle}, \tag{6.2.35}$$

with $q = m + p$. Clearly $\{\psi_k : 0 \leq k \leq q\}$ are linearly independent vectors in V, spanning a $(q+1)$-dimensional space W, on which the complex Lie algebra spanned by $\{E_\alpha, E_{-\alpha}, H_\alpha\}$ acts irreducibly. Thus we have the representation $D_{q/2}$ of SU(2) on W. The calculation (6.2.6)–(6.2.7), with V_λ replaced by V_{λ^*}, and n by q, exactly gives (6.2.35). $\square$

When Proposition 6.2.5 is specialized to the adjoint representation of G on $\mathfrak{g}_{\mathbb{C}}$, we get some useful results on the roots of $\mathfrak{g}$. Here is one.

Proposition 6.2.6. *If α and β are roots of $\mathfrak{g}$ and $[e_\alpha, e_\beta] = 0$, then $\mathfrak{g}_{\alpha+\beta} = 0$.*

Proof. In this case, we have an α-string generated by e_β, of the form $\{\mathfrak{g}_{\beta+\mu\alpha} : -M \le \mu \le 0\}$, for some $M \in \mathbb{Z}^+$. By (6.2.34), with $m = M, p = 0$,

$$2\frac{\langle \beta, \alpha\rangle}{\langle \alpha, \alpha\rangle} = M. \tag{6.2.36}$$

Now assume $\xi \in \mathfrak{g}_{\alpha+\beta}$ and $\xi \neq 0$. Then we have an α-string generated by ξ, of the form $\{\mathfrak{g}_{\alpha+\beta+\nu\alpha} : -m \le \nu \le p\}$, with $m, p \in \mathbb{Z}^+$. By (6.2.34),

$$m - p = 2\frac{\langle \alpha+\beta, \alpha\rangle}{\langle \alpha, \alpha\rangle} = 2 + M, \tag{6.2.37}$$

hence

$$m = 2 + M + p \ge 2 + M \ge 2. \tag{6.2.38}$$

Also

$$\begin{aligned} \operatorname{ad} e_{-\alpha}(\xi) &= A e_\beta \quad (\text{since } \dim \mathfrak{g}_\beta = 1) \\ &\neq 0 \qquad (\text{since } m \ge 2). \end{aligned} \tag{6.2.39}$$

Thus the bottom of this α-string is $\mathfrak{g}_{\beta - M\alpha} = \mathfrak{g}_{\beta+\alpha-(M+1)\alpha}$, hence

$$m = M + 1. \tag{6.2.40}$$

However, this contradicts (6.2.38), so there can be no such ξ. $\square$

Corollary 6.2.7. *If α and β are roots of $\mathfrak{g}$, then*

$$\mathfrak{g}_{\alpha+\beta} = \operatorname{ad} e_\alpha(\mathfrak{g}_\beta). \tag{6.2.41}$$

Note that Proposition 6.2.6 applies with $\alpha = \beta$ to reprove (6.2.22), again proving Proposition 6.2.4.

6.3. The Weyl group

In §4.4 we found it useful to know that conjugation by a permutation matrix E_σ, defined on the standard basis $\{u_1, \dots, u_n\}$ of $\mathbb{C}^n$ by

$$E_\sigma u_k = u_{\sigma(k)}, \tag{6.3.1}$$

preserves the maximal torus $\mathbb{T} \subset \mathrm{U}(n)$, consisting of diagonal unitary matrices, and permutes their entries:

$$E_\sigma^{-1} \mathrm{diag}(c_1, \dots, c_n) E_\sigma = \mathrm{diag}(c_{\sigma(1)}, \dots, c_{\sigma(n)}). \tag{6.3.2}$$

It followed that if π is a unitary representation of $\mathrm{U}(n)$ on V, then applications of $\pi(E_\sigma)$ permute the weight spaces; cf. (4.4.4).

Here we study an analogous structure on a general compact, connected Lie group G. The role of the symmetric group S_n for $G = \mathrm{U}(n)$ is taken by the *Weyl group* $W(G)$, defined as

$$W(G) = N(\mathbb{T})/\mathbb{T}, \tag{6.3.3}$$

where $\mathbb{T}$ is a maximal torus of G and $N(\mathbb{T})$ is the normalizer of $\mathbb{T}$:

$$N(\mathbb{T}) = \{g \in G : g^{-1} x g \in \mathbb{T},\ \forall\, x \in \mathbb{T}\}. \tag{6.3.4}$$

Note that

$$g \in N(\mathbb{T}) \Longrightarrow \mathrm{Ad}\, g : \mathfrak{h} \to \mathfrak{h}. \tag{6.3.5}$$

We define the representation $\mathcal{W}$ of $N(\mathbb{T})$ on $\mathfrak{h}$ by

$$\mathcal{W}(g) = \mathrm{Ad}(g)\big|_{\mathfrak{h}}, \quad \text{for} \quad g \in N(\mathbb{T}). \tag{6.3.6}$$

Then $N(\mathbb{T})$ has the contragredient representation $\overline{\mathcal{W}}$ on $\mathfrak{h}'$:

$$\langle \mathcal{W}(g^{-1}) H, \lambda \rangle = \langle H, \overline{\mathcal{W}}(g) \lambda \rangle, \quad g \in N(\mathbb{T}),\ H \in \mathfrak{h},\ \lambda \in \mathfrak{h}'. \tag{6.3.7}$$

Clearly $g \in \mathbb{T} \Rightarrow \mathcal{W}(g) = \mathrm{Ad}(g)|_{\mathfrak{h}} = I$, so we get representations of $W(G)$ on $\mathfrak{h}$ and $\mathfrak{h}'$, which we also denote $\mathcal{W}$ and $\overline{\mathcal{W}}$. We put an Ad-invariant inner product on $\mathfrak{g}$, inducing an inner product on $\mathfrak{h}$ invariant under $\mathcal{W}(g)$ for each $g \in N(\mathbb{T})$, and this induces an inner product on $\mathfrak{h}'$, invariant under $\overline{\mathcal{W}}(g)$ for each $g \in N(\mathbb{T})$. Since the representation $\mathcal{W}$ is real, $\mathcal{W}$ and $\overline{\mathcal{W}}$ are equivalent representations, intertwined by the isomorphism $\mathfrak{h} \approx \mathfrak{h}'$ induced by the inner product on $\mathfrak{h}$ just mentioned.

The following result generalizes (4.4.4).

Proposition 6.3.1. *Let π be a unitary representation of G on V, with weight space decomposition $V = \oplus V_\lambda$. Then*

$$g \in N(\mathbb{T}) \Longrightarrow \pi(g) : V_\lambda \to V_{\overline{\mathcal{W}}(g)\lambda}. \tag{6.3.8}$$

Proof. Recall that

$$V_\lambda = \{v \in V : d\pi(h)v = i\lambda(h)v,\ \forall\, h \in \mathfrak{h}\} \\ = \{v \in V : \pi(\operatorname{Exp} h)v = e^{i\lambda(h)}v,\ \forall\, h \in \mathfrak{h}\}. \tag{6.3.9}$$

Now

$$g \in N(\mathbb{T}),\ v \in V_\lambda \Longrightarrow \pi(g^{-1})\pi(\operatorname{Exp} h)\pi(g)v = \pi(\operatorname{Exp} \operatorname{Ad} g^{-1}\, h)v \\ = e^{i\lambda(\operatorname{Ad} g^{-1}\, h)}v, \tag{6.3.10}$$

and

$$\lambda(\operatorname{Ad} g^{-1}\, h) = \langle \mathcal{W}(g^{-1})h, \lambda\rangle = \langle h, \overline{\mathcal{W}}(g)\lambda\rangle, \tag{6.3.11}$$

so

$$g \in N(\mathbb{T}),\ v \in V_\lambda \Longrightarrow \pi(\operatorname{Exp} h)\pi(g)v = e^{i(\overline{\mathcal{W}}(g)\lambda)(h)}\pi(g)v \\ \Longrightarrow \pi(g)v \in V_{\overline{\mathcal{W}}(g)\lambda}, \tag{6.3.12}$$

as stated in (6.3.8). □

In particular, the Weyl group permutes the roots of $\mathfrak{g}$. That is, if α is a root, so is $\overline{\mathcal{W}}(g)\alpha$, for each $g \in N(\mathbb{T})$. The following result gives valuable information on how $W(G)$ permutes the roots, and implies that $W(G)$ has lots of elements.

Proposition 6.3.2. *For each root α, there exists $g_\alpha \in N(\mathbb{T})$ such that*

$$\mathcal{W}(g_\alpha)H = H - 2\frac{\langle H_\alpha, H\rangle}{\langle H_\alpha, H_\alpha\rangle}H_\alpha, \quad \forall\, H \in \mathfrak{h}, \tag{6.3.13}$$

i.e., $\mathcal{W}(g_\alpha)$ is reflection across the hyperplane in $\mathfrak{h}$ orthogonal to H_α. Hence

$$\overline{\mathcal{W}}(g_\alpha)\lambda = \lambda - 2\frac{\langle \alpha, \lambda\rangle}{\langle \alpha, \alpha\rangle}\alpha, \quad \forall\, \lambda \in \mathfrak{h}'. \tag{6.3.14}$$

This is reflection across the hyperplane in $\mathfrak{h}'$ orthogonal to α.

Proof. We will show that (6.3.13) holds with $g_\alpha = A_\alpha(\pi)$, where

$$A_\alpha(t) = \operatorname{Exp} tX_3^\alpha. \tag{6.3.15}$$

Here X_3^α is as in (6.2.3) and $\pi = 3.14159\cdots$. To begin, note that

$$\operatorname{Ad}(A_\alpha(t))H = e^{t\, \operatorname{ad} X_3^\alpha}H. \tag{6.3.16}$$

Now

$$H \in \ker \alpha \subset \mathfrak{h} \Longrightarrow \operatorname{ad} e_{\pm\alpha}(H) = 0 \quad (\text{since } [H, e_{\pm\alpha}] = \pm i\alpha(H)e_{\pm\alpha}) \\ \Longrightarrow \operatorname{ad} X_3^\alpha(H) = 0 \\ \Longrightarrow e^{t\, \operatorname{ad} X_3^\alpha}(H) = H. \tag{6.3.17}$$

We note parenthetically that by the same reasoning, $\alpha(H) = 0 \Rightarrow \operatorname{ad} X_2^\alpha(H) = 0$, and of course $\operatorname{ad}(H_\alpha)H = 0$; that is to say, more generally than (6.3.17), we have

$$H \in \ker \alpha \Longrightarrow \operatorname{Ad} \circ \gamma^\alpha(g) H = H, \quad \forall\, g \in SU(2), \tag{6.3.18}$$

with γ^α defined by (6.2.5).

Since $\alpha(H) = \langle H, H_\alpha \rangle$, we have $\ker \alpha = (H_\alpha)^\perp$. The result (6.3.13) (and the containment $A_\alpha(\pi) \in N(\mathbb{T})$) will hence follow from (6.3.17) together with the result

$$\operatorname{Ad}(A_\alpha(\pi)) H_\alpha = -H_\alpha. \tag{6.3.19}$$

To establish (6.3.19), we analyze the action of $\operatorname{Ad}(A_\alpha(t))$ on X_1^α (which by (6.2.3) is parallel to H_α). The commutation relations (6.2.4) give

$$\operatorname{ad} X_3^\alpha (X_1^\alpha \pm i X_2^\alpha) = \mp i (X_1^\alpha \pm i X_2^\alpha), \tag{6.3.20}$$

hence

$$e^{t \operatorname{ad} X_3^\alpha} (X_1^\alpha \pm i X_2^\alpha) = e^{\mp it} (X_1^\alpha \pm i X_2^\alpha), \tag{6.3.21}$$

hence

$$\begin{aligned} \operatorname{Ad}(A_\alpha(t)) X_1^\alpha &= e^{t \operatorname{ad} X_3^\alpha} X_1^\alpha \\ &= \frac{1}{2} e^{t \operatorname{ad} X_3^\alpha} [(X_1^\alpha + i X_2^\alpha) + (X_1^\alpha - i X_2^\alpha)] \\ &= \frac{1}{2} [e^{-it} (X_1^\alpha + i X_2^\alpha) + e^{it} (X_1^\alpha - i X_2^\alpha)], \end{aligned} \tag{6.3.22}$$

hence

$$\operatorname{Ad}(A_\alpha(\pi)) X_1^\alpha = -X_1^\alpha, \tag{6.3.23}$$

which gives (6.3.19). This proves (6.3.13), and (6.3.14) follows. $\square$

Remark. If $\alpha, \beta \in \mathfrak{h}'$ are both roots, then (6.3.14) together with (6.2.13) gives

$$\overline{\mathcal{W}}(g_\alpha)\beta = \beta - n_{\beta\alpha}\alpha, \tag{6.3.24}$$

where $n_{\alpha\beta}$ are the Cartan integers. We use the notation

$$S_\alpha \beta = \overline{\mathcal{W}}(g_\alpha)\beta. \tag{6.3.25}$$

The following result implies we can identify $W(G)$ with its image under $\mathcal{W}$ in $\mathrm{Gl}(\mathfrak{h})$, or under $\overline{\mathcal{W}}$ in $\mathrm{Gl}(\mathfrak{h}')$.

Proposition 6.3.3. *If $g \in G$ and $g^{-1}xg = x$ for each $x \in \mathbb{T}$, then $g \in \mathbb{T}$. Hence if $g \in N(\mathbb{T})$ and $\mathcal{W}(g) = I$ on $\mathfrak{h}$, then $g \in \mathbb{T}$.*

For a proof valid for general compact, connected G, see [**34**], p. 167. In the special case $G = \mathrm{U}(n)$ (or $G = \mathrm{SU}(n)$) we can see the result as follows. Take $g \in \mathrm{U}(n)$ and $x = \mathrm{diag}(c_1, \dots, c_n) \in \mathbb{T}$. Then forming gx multiplies the jth column of g by c_j and forming xg multiplies the jth row of g by c_j. From this it is apparent that if $gx = xg$ for all such x, then g must be a diagonal matrix.

Here is another proof, for $G = \mathrm{U}(n)$. If $g \in \mathrm{U}(n)$ commutes with each $x \in \mathbb{T}$, then g commutes with superpositions of such elements. In particular, one can take $x_j = (1, \dots, -1, \dots, 1)$, with all ones except -1 in the jth position, and see that g commutes with $I - x_j$, hence with the orthogonal projection onto $\mathrm{Span}\, e_j$, for each j. This requires g to be diagonal, and forces $g \in \mathbb{T}$.

The reader is invited to produce a similar argument for $G = \mathrm{SO}(n)$.

Regarding the image of $W(G)$ under $\mathcal{W}$, of course each element $\mathcal{W}(g)$ ($g \in N(\mathbb{T})$) acts trivially on the center $\mathfrak{z}$ of $\mathfrak{g}$ ($\mathfrak{z} \subset \mathfrak{h}$). By Proposition 6.1.3 and (6.1.6), we have

$$\mathfrak{h} = \mathfrak{z} \oplus \mathrm{Span}\,\{H_\alpha : \alpha \text{ root}\}. \tag{6.3.26}$$

Consequently,

$$\begin{aligned} g \in N(\mathbb{T}),\ \mathcal{W}(g)H_\alpha = H_\alpha,\ \forall\, \alpha& \\ &\Longrightarrow \mathcal{W}(g) = I \quad \text{on} \quad \mathfrak{h} \\ &\Longrightarrow [g] = [e] \quad \text{in} \quad W(G), \end{aligned} \tag{6.3.27}$$

where $[g]$ denotes the image of g under $N(\mathbb{T}) \to N(\mathbb{T})/\mathbb{T} = W(G)$. We deduce that the isomorphic image of $W(G)$ in $\mathrm{Gl}(\mathfrak{h})$ is in turn isomorphic to a subgroup of the group of permutations of the set

$$\Delta = \{\alpha \in \mathfrak{h}' : \alpha \text{ root of } \mathfrak{g}\}. \tag{6.3.28}$$

In particular,

$$\#W(G) \,\big|\, (\#\Delta)!, \tag{6.3.29}$$

where $\#S$ denotes the number of elements of a set S. To reiterate, for $g \in N(\mathbb{T})$, the action of $\overline{\mathcal{W}}(g)$ on $\mathfrak{h}'$ is uniquely determined by the action of $\overline{\mathcal{W}}(g)$ on the roots. The following result complements this assertion.

Proposition 6.3.4. *The image of* $W(G)$ *under* $\overline{\mathcal{W}}$ *in* $\mathrm{Gl}(\mathfrak{h}')$ *is generated by the set of reflections* S_α*, given by* (6.3.24)–(6.3.25).

For a proof of this, see [**34**], Chapter 8. It is easy enough to verify in case $G = \mathrm{U}(n)$. In that case,

$$\Delta = \{\omega_{jk} : j \neq k,\ 1 \leq j, k \leq n\}, \tag{6.3.30}$$

with ω_{jk} as in (4.2.14). Equivalently, with $\{e_j : 1 \le j \le n\}$ the basis of $\mathfrak{h}$ given by (4.2.9) and $\{e'_j : 1 \le j \le n\}$ the dual basis, $\omega_{jk} = e'_j - e'_k$. A calculation gives

$$S_{\omega_{jk}}\omega_{\ell m} = \omega_{\sigma(\ell)\sigma(m)}, \quad \text{where} \ \ \sigma = (j\ k), \tag{6.3.31}$$

i.e., $\sigma \in S_n$ is the transposition that switches j and k and leaves the other elements of $\{1, \dots, n\}$ fixed. It is well known that the set of transpositions generates S_n, so $\{S_{\omega_{jk}}\}$ generates

$$\{S_\sigma : \sigma \in S_n\}, \quad S_\sigma\omega_{\ell m} = \omega_{\sigma(\ell)\sigma(m)}. \tag{6.3.32}$$

That this exhausts $W(\mathrm{U}(n))$ follows from:

Proposition 6.3.5. *Let V be an n-dimensional real inner product space with orthonormal basis $\{e'_j : 1 \le j \le n\}$. Let S be an orthogonal transformation on V such that S fixes $e'_1 + \cdots + e'_n$ and permutes the vectors $\omega_{jk} = e'_j - e'_k$, $j \ne k$. Then S has the form* (6.3.32)*, for some $\sigma \in S_n$.*

Proof. Left to the reader. □

Given Proposition 6.3.5, we have

$$W(\mathrm{U}(n)) \approx S_n. \tag{6.3.33}$$

We also claim

$$W(\mathrm{SU}(n)) \approx S_n. \tag{6.3.34}$$

The action of $\sigma \in S_n$ on the maximal torus in $\mathrm{SU}(n)$ is given by a slight modification of (6.3.1)–(6.3.2), needed because $\det E_\sigma = \operatorname{sgn}\sigma$. More generally than (6.3.1), we can take $\theta = (\theta_1, \dots, \theta_n) \in \{\pm 1\} \times \cdots \times \{\pm 1\}$ and define $E^\theta_\sigma \in \mathrm{O}(n) \subset \mathrm{U}(n)$ by

$$E^\theta_\sigma u_k = \theta_k u_{\sigma(k)}. \tag{6.3.35}$$

Then

$$\det E^\theta_\sigma = \theta_1 \cdots \theta_n \cdot \operatorname{sgn}\sigma, \tag{6.3.36}$$

so for each $\sigma \in S_n$ there exist elements $E^\theta_\sigma \in \mathrm{SO}(n) \subset \mathrm{SU}(n)$. One has the following extension of (6.3.2):

$$(E^\theta_\sigma)^{-1} \operatorname{diag}(c_1, \dots, c_n) E^\theta_\sigma = \operatorname{diag}(c_{\sigma(1)}, \dots, c_{\sigma(n)}). \tag{6.3.37}$$

If also $E^\varphi_\sigma \in \mathrm{SU}(n)$, then $E^\theta_\sigma(E^\varphi_\sigma)^{-1}$ is a diagonal element of $\mathrm{SU}(n)$, so the two elements define the same element of $N(\mathbb{T})/\mathbb{T}$, where $\mathbb{T}$ is the maximal torus consisting of diagonal elements of $\mathrm{SU}(n)$.

6.4. A generating function

Let G be a compact, connected Lie group, with maximal torus $\mathbb{T}$, whose Lie algebra is denoted $\mathfrak{h}$. Let $\lambda \in \mathfrak{h}'$ run over the collection of highest weights for irreducible unitary representations of G. Denote the corresponding representation by π_λ, acting on W_λ. Parallel to (6.1.36), we let $\xi_\lambda \in W_\lambda$ be a highest weight vector, and take $\eta_\lambda \in W'_\lambda$ to be a lowest weight vector for $\overline{\pi}_\lambda$. We know by Lemma 6.1.9 that $\langle \xi_\lambda, \eta_\lambda \rangle \neq 0$; normalize so that $\langle \xi_\lambda, \eta_\lambda \rangle = 1$, and set

$$\varphi_\lambda(g) = \langle \pi_\lambda(g)\xi_\lambda, \eta_\lambda \rangle. \tag{6.4.1}$$

This is just as in (6.1.36), except that here we record the dependence on λ. This family of functions on G has the following important property.

Proposition 6.4.1. *If λ and μ are highest weights, then*

$$\varphi_{\lambda+\mu}(g) = \varphi_\lambda(g)\varphi_\mu(g). \tag{6.4.2}$$

Proof. We know by Proposition 6.2.2 that $\lambda + \mu$ is the highest weight for an irreducible component of $\pi_\lambda \otimes \pi_\mu$ on $W_\lambda \otimes W_\mu$, with weight vector $\xi_\lambda \otimes \xi_\mu$. Similarly, $-\lambda - \mu$ is the lowest weight for an irreducible component of $\overline{\pi}_\lambda \otimes \overline{\pi}_\mu$ on $W'_\lambda \otimes W'_\mu$, with weight vector $\eta_\lambda \otimes \eta_\mu$. Hence, by uniqueness (cf. Theorem 6.1.10),

$$\begin{aligned} \varphi_{\lambda+\mu}(g) &= \langle \pi_\lambda(g) \otimes \pi_\mu(g)(\xi_\lambda \otimes \xi_\mu), \eta_\lambda \otimes \eta_\mu \rangle \\ &= \varphi_\lambda(g)\varphi_\mu(g), \end{aligned} \tag{6.4.3}$$

as asserted. □

Let us recall the conjugate linear map $C : V_\lambda \to V'_\lambda$ from (2.3.13)–(2.3.14), satisfying

$$(u, v) = \langle u, Cv \rangle, \quad \overline{\pi}_\lambda(g) = C\pi_\lambda(g)C^{-1}. \tag{6.4.4}$$

In this setting we have (up to scaling)

$$\eta_\lambda = C\xi_\lambda, \tag{6.4.5}$$

and hence

$$\varphi_\lambda(g) = (\pi_\lambda(g)\xi_\lambda, \xi_\lambda), \tag{6.4.6}$$

using the Hermitian inner product on W_λ rather than the $W_\lambda - W'_\lambda$ duality. We require $(\xi_\lambda, \xi_\lambda) = 1$, so as before

$$\varphi_\lambda(e) = 1. \tag{6.4.7}$$

We now demonstrate a connection between φ_λ and the character $\chi_\lambda(g) = \operatorname{Tr} \pi_\lambda(g)$.

Proposition 6.4.2. *We have*

$$\chi_\lambda(x) = d_\lambda \int\limits_G \varphi_\lambda(g^{-1}xg)\,dg, \tag{6.4.8}$$

where $d_\lambda = \dim W_\lambda$.

Proof. Denote the right side of (6.4.8) by $\psi_\lambda(x)$. We have

$$\psi_\lambda(g^{-1}xg) = \psi_\lambda(x), \quad \forall\, x, g \in G. \tag{6.4.9}$$

That is to say, ψ_λ is central, so by Proposition 2.4.1 it must be a constant multiple of χ_λ. Since $\psi_\lambda(e) = d_\lambda = \chi_\lambda(e)$, we have the identity (6.4.8). $\square$

We do not address here the possible use of Propositions 6.4.1–6.4.2 as a tool for a derivation of the Weyl character formula, nor of the Weyl dimension formula, which in the current setting is

$$d_\lambda = \prod_{\alpha\in\Delta^+} \frac{\langle \lambda+\rho, \alpha\rangle}{\langle \rho, \alpha\rangle}, \tag{6.4.10}$$

where Δ^+ is the set of positive roots,

$$\rho = \frac{1}{2} \sum_{\alpha\in\Delta^+} \alpha, \tag{6.4.11}$$

and $\langle\ ,\ \rangle$ is the inner product on $\mathfrak{h}'$ arising from the Killing form.

6.5. The complexification of a general compact Lie group

Here we construct the complexification $G_{\mathbb{C}}$ of a compact, connected Lie group G and extend Theorem 4.6.1. To begin, take a faithful unitary representation ρ of G on some space $\mathbb{C}^n$. The existence of such a representation is guaranteed by Proposition 2.8.8, and is apparent for the standard examples. Thus we have

$$\rho : G \longrightarrow \mathrm{U}(n) \subset \mathrm{Gl}(n, \mathbb{C}), \tag{6.5.1}$$

taking G isomorphically onto its image G^ρ, with

$$d\rho : \mathfrak{g} \longrightarrow \mathfrak{u}(n) \subset \mathrm{M}(n, \mathbb{C}), \tag{6.5.2}$$

taking $\mathfrak{g}$ isomorphically onto its image, $\mathfrak{g}^\rho$. Define $G^\rho_{\mathbb{C}}$ to be the Lie subgroup of $\mathrm{Gl}(n, \mathbb{C})$ generated by $\mathfrak{g}^\rho_{\mathbb{C}} \subset \mathrm{M}(n, \mathbb{C})$.

The group $G^\rho_{\mathbb{C}}$ is a complex submanifold of $\mathrm{Gl}(n, \mathbb{C})$, with a natural holomorphic structure, such that the group actions are holomorphic maps. To see this, note that

$$\mathrm{Exp} : M(n, \mathbb{C}) \longrightarrow \mathrm{Gl}(n, \mathbb{C}) \tag{6.5.3}$$

is a holomorphic map (being given by a convergent power series), and $D\,\mathrm{Exp}(0) = I$. The inverse function theorem applied to holomorphic maps yields local inverses that are not only smooth but also holomorphic. Hence Exp is a holomorphic diffeomorphism of a neighborhood Ω of $0 \in M(n, \mathbb{C})$ onto a neighborhood $\mathcal{O}$ of I in $\mathrm{Gl}(n, \mathbb{C})$. Thus it restricts to a holomorphic diffeomorhism of $\Omega \cap \mathfrak{g}^\rho_{\mathbb{C}}$ onto its image, defining the complex structure of $G^\rho_{\mathbb{C}}$ in a neighborhood of I. Given any $g_0 \in G^\rho_{\mathbb{C}}$, one has the holomorphic diffeomorphism

$$E_{g_0} : \Omega \longrightarrow g_0\mathcal{O}, \quad E_{g_0}(X) = g_0\,\mathrm{Exp}(X), \tag{6.5.4}$$

mapping $\Omega \cap \mathfrak{g}^\rho_{\mathbb{C}}$ diffeomorphically onto a neighborhood of g_0 in $G^\rho_{\mathbb{C}}$.

Shortly we will show that this complexification of G is independent of the choice of ρ, up to natural isomorphism.

Here is the extension of Theorem 4.6.1.

Theorem 6.5.1. *If π is a representation of G on a finite dimensional complex vector space V, then there is a holomorphic representation π^ρ of $G^\rho_{\mathbb{C}}$ on V such that*

$$\pi^\rho \circ \rho(g) = \pi(g), \quad \forall\, g \in G. \tag{6.5.5}$$

The proof will parallel that of Theorem 4.6.1. To set it up, define the representation $T^{p,q}_\rho$ of G on $T^{p,q}(\mathbb{C}^n) = (\otimes^p\mathbb{C}^n) \otimes (\otimes^q\mathbb{C}^n)$ by

$$\begin{aligned} &T^{p,q}_\rho(g)v_1 \otimes \cdots \otimes v_p \otimes w_1 \otimes \cdots \otimes w_q \\ &= \rho(g)v_1 \otimes \cdots \otimes \rho(g)v_p \otimes \rho(g^{-1})^t w_1 \otimes \cdots \otimes \rho(g^{-1})^t w_q. \end{aligned} \tag{6.5.6}$$

Note that

$$\rho(g) \in \mathrm{U}(n) \Longrightarrow \rho(g^{-1})^t = \overline{\rho(g)}. \tag{6.5.7}$$

Next we define the representation $T_{K,\rho}$ of G on $T_K(\mathbb{C}^n) = \oplus_{p+q\le K} T^{p,q}(\mathbb{C}^n)$ by

$$T_{K,\rho}(g)\Big(\bigoplus_{p+q\le K} v_{pq}\Big) = \bigoplus_{p+q\le K} T_\rho^{p,q}(g) v_{pq}. \tag{6.5.8}$$

Then we have:

Proposition 6.5.2. *If π is a finite dimensional representation of G on V, then there exists $K < \infty$ such that π is contained in $T_{K,\rho}$.*

Proof. Same as that of Proposition 4.6.2. □

The content of Proposition 6.5.2 is that, for some K, there is a linear subspace W of $T_K(\mathbb{C}^n)$, invariant under the action of $T_{K,\rho}$, and a linear isomorphism $J : V \to W$ such that

$$\pi(g) = J^{-1} T_{K,\rho}(g) J, \tag{6.5.9}$$

for all $g \in G$. Note that, with T_K as in (4.6.3)–(4.6.4), we have

$$\pi(g) = J^{-1} T_K(\rho(g)) J, \quad \forall\, g \in G. \tag{6.5.10}$$

Theorem 6.5.1 follows from this, via arguments similar to those used in §4.6. As noted there, T_K extends from $\mathrm{U}(n)$ to $\mathrm{Gl}(n,\mathbb{C})$, holomorphically. We have

$$\pi^\rho(\tilde{g}) = J^{-1} T_K(\tilde{g}) J, \tag{6.5.11}$$

for $\tilde{g} \in G^\rho$. We claim that (6.5.11) extends from $\tilde{g} \in G^\rho$ to $\tilde{g} \in G^\rho_{\mathbb{C}}$. To see this, we start with the fact that

$$T_K(\tilde{g}) : W \longrightarrow W, \tag{6.5.12}$$

for all $\tilde{g} \in G^\rho$. We want to show that (6.5.12) holds for all $\tilde{g} \in G^\rho_{\mathbb{C}}$. Indeed, the validity of (6.5.12) for all $\tilde{g} \in G^\rho$ implies

$$dT_K(X) : W \longrightarrow W, \tag{6.5.13}$$

for all $X \in \mathfrak{g}^\rho$. Since T_K is holomorphic, $dT_K : M(n,\mathbb{C}) \to \mathcal{L}(T_K(\mathbb{C}^n))$ is $\mathbb{C}$-linear, so (6.5.13) holds for all $X \in \mathfrak{g}^\rho_{\mathbb{C}}$. Thus

$$T_K(e^{tX}) = e^{t\, dT_K(X)} : W \longrightarrow W, \tag{6.5.14}$$

for all $X \in \mathfrak{g}^\rho_{\mathbb{C}}$, so (6.5.12) holds for all $\tilde{g}$ in a neighborhood of the identity in $G^\rho_{\mathbb{C}}$, hence for all $\tilde{g} \in G^\rho_{\mathbb{C}}$. This yields the desired extension of (6.5.11) to all $\tilde{g} \in G^\rho_{\mathbb{C}}$, and hence yields Theorem 6.5.1.

We next establish uniqueness:

Proposition 6.5.3. *If σ is another faithful unitary representation of G, on $\mathbb{C}^m$, we have a natural holomorphic isomorphism*

$$G_{\mathbb{C}}^{\rho} \approx G_{\mathbb{C}}^{\sigma}. \tag{6.5.15}$$

Proof. Applying Theorem 6.5.1 to $\pi = \sigma$, we have a holomorphic representation σ^ρ of $G_{\mathbb{C}}^{\rho}$ on $\mathbb{C}^m$, i.e.,

$$\sigma^\rho : G_{\mathbb{C}}^{\rho} \longrightarrow \mathrm{Gl}(m, \mathbb{C}), \tag{6.5.16}$$

such that

$$\sigma^\rho \circ \rho(g) = \sigma(g), \quad \forall\, g \in G. \tag{6.5.17}$$

We see that $d\sigma^\rho$ takes the Lie algebra $\mathfrak{g}^\rho$ isomorphically onto $\mathfrak{g}^\sigma$; hence it extends to an isomorphism of the complexifications of these Lie algebras. This implies

$$\sigma^\rho : G_{\mathbb{C}}^{\rho} \longrightarrow G_{\mathbb{C}}^{\sigma}, \tag{6.5.18}$$

with $d\sigma^\rho : \mathfrak{g}_{\mathbb{C}}^{\rho} \to \mathfrak{g}_{\mathbb{C}}^{\sigma}$, isomorphically. Interchanging the roles of ρ and σ, we have a holomorphic homomorphism

$$\rho^\sigma : G_{\mathbb{C}}^{\sigma} \longrightarrow G_{\mathbb{C}}^{\rho}. \tag{6.5.19}$$

We claim that the maps in (6.5.18) and (6.5.19) are inverse to each other, which will yield (6.5.15).

To see this, consider

$$\kappa = \rho^\sigma \circ \sigma^\rho : G_{\mathbb{C}}^{\rho} \to G_{\mathbb{C}}^{\rho}. \tag{6.5.20}$$

We have

$$\kappa(g) = g, \quad \forall\, g \in G^\rho. \tag{6.5.21}$$

Hence $d\kappa : \mathfrak{g}_{\mathbb{C}}^{\rho} \to \mathfrak{g}_{\mathbb{C}}^{\rho}$ is a $\mathbb{C}$-linear Lie algebra homomorphism that is the identity on $\mathfrak{g}^\rho \subset \mathfrak{g}_{\mathbb{C}}^{\rho}$, hence is the identity on $\mathfrak{g}_{\mathbb{C}}^{\rho}$. It follows that $\kappa : G_{\mathbb{C}}^{\rho} \to G_{\mathbb{C}}^{\rho}$ is the identity. Similarly $\sigma^\rho \circ \rho^\sigma : G_{\mathbb{C}}^{\sigma} \to G_{\mathbb{C}}^{\sigma}$ is the identity, and we are done. $\square$

In light of this uniqueness, we choose any such $G_{\mathbb{C}}^{\rho}$ as constructed above, denote it $G_{\mathbb{C}}$, and call it "the complexification" of G.

Exercises

1. Show that
$$U(1)_{\mathbb{C}} = \{e^{iz} : z \in \mathbb{C}\} = \mathbb{C} \setminus 0 = Gl(1, \mathbb{C}),$$
and
$$SO(2)_{\mathbb{C}} = \{e^{zJ} : z \in \mathbb{C}\}, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$
and verify that $U(1)_{\mathbb{C}} \approx SO(2)_{\mathbb{C}}$.

2. Show that
$$U(n)_{\mathbb{C}} = Gl(n, \mathbb{C}), \quad SU(n)_{\mathbb{C}} = Sl(n, \mathbb{C}).$$

3. Show that
$$SO(n)_{\mathbb{C}} = \{A \in Sl(n, \mathbb{C}) : A^t A = I\},$$
with Lie algebra
$$\mathfrak{so}(n)_{\mathbb{C}} = \{X \in M(n, \mathbb{C}) : X^t = -X\}.$$
Compare the case $n = 2$ with the result of Exercise 1.

6.6. Simple roots, Cartan matrices, and Dynkin diagrams

Let G be a compact, connected Lie group, with Lie algebra $\mathfrak{g}$ and maximal torus $\mathbb{T}$, whose Lie algebra is $\mathfrak{h}$. A choice of ordered basis of $\mathfrak{h}$ produces an order on $\mathfrak{h}'$, hence on the set of roots of $\mathfrak{g}$. We set

$$\Delta = \text{ set of roots}, \quad \Delta^+ = \text{ set of positive roots.} \tag{6.6.1}$$

A root α is called a *simple root* if α is positive and it cannot be written as a sum of two positive roots. We set

$$\Sigma = \text{ set of simple roots.} \tag{6.6.2}$$

For example, if $G = \mathrm{U}(n)$, or $\mathrm{SU}(n)$, then, with ω_{jk} as in (19.13),

$$\Delta = \{\omega_{jk} : j \neq k\}, \quad \Delta^+ = \{\omega_{jk} : j < k\}, \tag{6.6.3}$$

and one has

$$\Sigma = \{\omega_{j,j+1} : 1 \le j \le n-1\}. \tag{6.6.4}$$

Recall from (6.2.13) the quantities

$$n_{\alpha\beta} = 2\,\frac{\langle \alpha, \beta\rangle}{\langle \beta, \beta\rangle} \in \mathbb{Z}, \tag{6.6.5}$$

associated to $\alpha, \beta \in \Delta$, called the Cartan integers. These possess a number of properties, given in (6.2.15)–(6.2.21) and (6.2.26)–(6.2.27). We form the *Cartan matrix* of $\mathfrak{g}$, from (6.6.5) with α and β from

$$\Sigma = \{\alpha_1, \dots, \alpha_m\}. \tag{6.6.6}$$

This is the $m \times m$ matrix A, with entries

$$A_{jk} = 2\,\frac{\langle \alpha_j, \alpha_k\rangle}{\langle \alpha_k, \alpha_k\rangle}. \tag{6.6.7}$$

Clearly $A_{jj} = 2$ for each j. We say more about the off-diagonal entries below.

If $G = \mathrm{U}(n)$ or $\mathrm{SU}(n)$, we take $\alpha_j = \omega_{j,j+1}$, as in (6.6.4). Thus $m = n-1$. Note that

$$\begin{aligned} \langle \omega_{j,j+1}, \omega_{k,k+1}\rangle = -1 &\quad \text{if } \ |j-k| = 1, \\ 2 &\quad \text{if } \ j = k, \\ 0 &\quad \text{otherwise,} \end{aligned} \tag{6.6.8}$$

and we have the $(n-1) \times (n-1)$ matrix

$$A = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & \ddots & & \\ & -1 & \ddots & -1 & \\ & & \ddots & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}. \tag{6.6.9}$$

The *Dynkin diagram* associated to $\mathfrak{g}$ is produced as follows. It is a graph whose vertices are the simple roots $\{\alpha_j\}$. Thus, for each $\alpha_j \in \Sigma$, place a small circle. Connect the dots α_j and α_k by a number of straight lines equal to $A_{jk}A_{kj}$. By (6.2.15),

$$\begin{aligned} A_{jk}A_{kj} &= 4\,\frac{\langle \alpha_j, \alpha_k \rangle^2}{|\alpha_j|^2 |\alpha_k|^2} \\ &= 4\cos^2\theta_{jk}, \end{aligned} \tag{6.6.10}$$

where θ_{jk} is the angle between α_j and α_k in $\mathfrak{h}'$. By (36.16), $4\cos^2\theta_{jk} \in \{0,1,2,3,4\}$. However, $\cos^2\theta_{jk} = 1 \Leftrightarrow \alpha_j$ is a real multiple of α_k, which, by Proposition 6.2.4, forces $\alpha_j = \alpha_k$. Thus

$$j \neq k \Longrightarrow 4\cos^2\theta_{jk} \in \{0,1,2,3\}. \tag{6.6.11}$$

When $\alpha_j \perp \alpha_k$, these vertices are not connected by any lines. Otherwise, the number of lines connecting α_j to α_k is 1, 2, or 3.

The Dynkin diagram for $\mathrm{U}(n)$ or $\mathrm{SU}(n)$ has $n-1$ dots, and is depicted in Figure 6.6.1.

There is one further ingredient in the constructon of a Dynkin diagram. If $\mathfrak{g}$ is a simple Lie algebra, it turns out that there are at most two distinct lengths of the simple roots. In such a case, put a bull's eye in each circle corresponding to the smaller roots. For $\mathfrak{g} = \mathfrak{su}(n)$, all the simple roots $\omega_{j,j+1}$ have length $\sqrt{2}$, so we do not alter any of them.

We record some general results concerning simple roots. To begin, we have

Proposition 6.6.1.

$$\alpha, \beta \in \Sigma \Longrightarrow \alpha - \beta \notin \Delta. \tag{6.6.12}$$

Proof. If $\alpha - \beta \in \Delta$, then either $\alpha - \beta \in \Delta^+$ or $\beta - \alpha \in \Delta^+$. Hence either $\alpha = \beta + (\alpha - \beta)$ or $\beta = \alpha + (\beta - \alpha)$ is a sum of two positive roots. $\square$

Next, we claim

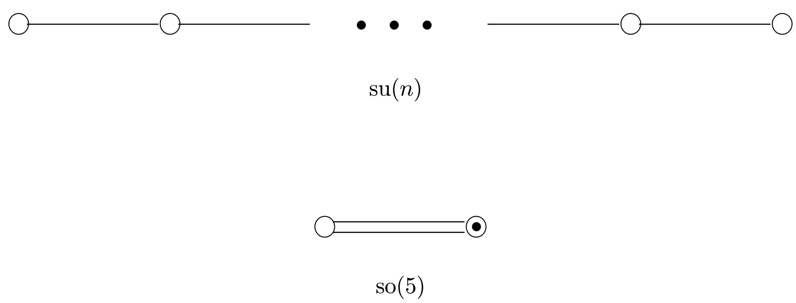

Figure 6.6.1. Dynkin diagrams of su(n) and so(5)

Proposition 6.6.2.

$$\alpha, \beta \in \Sigma, \ \alpha \neq \beta \Longrightarrow \langle \alpha, \beta \rangle \leq 0. \tag{6.6.13}$$

Proof. We apply Proposition 6.2.5, taking π to be the adjoint representation. Given distinct $\alpha, \beta \in \Sigma$, (6.6.12) implies β is at the bottom of an α-string of roots,

$$\beta + j\alpha, \quad 0 \leq j \leq p, \tag{6.6.14}$$

for some $p \in \mathbb{Z}^+$. Then (6.2.34) applies, with $\lambda = \beta, m = 0$:

$$2 \, \frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} = -p. \tag{6.6.15}$$

□

A consequence of (6.6.13) is that each off-diagonal element of the Cartan matrix is either 0 or a negative integer. The example (6.6.9) illustrates this.

From Proposition 6.6.2, we have the following.

Proposition 6.6.3. *The set Σ of simple roots is linearly independent in $\mathfrak{h}'$.*

Proof. If not, there is an identity of the form

$$\sum_{\alpha_j \in \Sigma_0} x_j \alpha_j = \sum_{\alpha_k \in \Sigma_1} y_k \alpha_k, \quad x_j, y_k > 0, \tag{6.6.16}$$

with Σ_0 and Σ_1 disjoint subsets of Σ. Taking the inner product of both sides with the left side yields

$$\Big\langle \sum x_j \alpha_j, \sum y_j \alpha_k \Big\rangle > 0, \tag{6.6.17}$$

contradicting (6.6.13). □

There are enough simple roots to yield the following.

Proposition 6.6.4. *Each $\alpha \in \Delta^+$ can be written as a positive sum of simple roots.*

Proof. If not, let $\beta \in \Delta^+$ be the smallest root for which this is not true. Then β is not simple, so it is a sum $\beta = \beta_1 + \beta_2$ of positive roots β_j. But then each β_j is a positive sum of simple roots. □

Recall from Corollary 6.1.4 that if the center $\mathfrak{z}$ of $\mathfrak{g}$ is 0, then $\{\alpha \in \Delta\}$ spans $\mathfrak{h}'$. Of course, this implies $\{\alpha \in \Delta^+\}$ spans $\mathfrak{h}'$. We then have the following, from Propositions 6.6.3–6.6.4:

Corollary 6.6.5. *If $\mathfrak{z} = 0$, then $\{\alpha \in \Sigma\}$ is a basis of $\mathfrak{h}'$.*

Thus $\dim \mathfrak{h} = \dim \mathfrak{h}' = \#\Sigma$ if $\mathfrak{z} = 0$. We see this for $\mathfrak{g} = \mathfrak{su}(n)$, where $\dim \mathfrak{h} = n - 1 = \#\Sigma$, directly from (6.6.4).

The following is a stronger statement about how plentiful the simple roots are.

Proposition 6.6.6. *The set $\{e_{\alpha_j} : \alpha_j \in \Sigma\}$ generates the Lie algebra $\mathfrak{n}_+ =$ Span$\{e_\alpha : \alpha \in \Delta^+\}$.*

Proof. The Lie algebra generated by $\{e_{\alpha_j} : \alpha_j \in \Sigma\}$ is the span of

$$\operatorname{ad} e_{\alpha_{j_1}} \cdots \operatorname{ad} e_{\alpha_{j_\mu}} (e_{\alpha_{j_\ell}}), \tag{6.6.18}$$

as j_ν range independently over $\{1, \dots, m\}$. This is clearly contained in $\mathfrak{n}_+$. For the converse, we need to show that, whenever $\beta \in \Delta^+$, e_β is a linear combination of elements of the form (6.6.18).

Clearly each element of Σ has this property. If not each element of Δ^+ does, there is a smallest $\gamma \in \Delta^+$ such that e_γ is not a linear combination of elements of the form (6.6.18). Since $\gamma \notin \Sigma$, we have $\gamma = \beta_1 + \beta_2$ with $\beta_j \in \Delta^+$, and each term is smaller than γ. Hence each e_{β_j} is a linear combination of elements of the form (6.6.18). By Proposition 6.2.6, $[e_{\beta_1}, e_{\beta_2}]$ must be a nonzero multiple of e_γ, so by Jacobi's identity e_γ must be of the desired form. This contradiction proves the proposition. □

The following refinement of Proposition 6.6.4 will be useful.

Proposition 6.6.7. *If $\beta \in \Delta^+$, $\beta \notin \Sigma$, then there exists $\alpha_j \in \Sigma$ and $\beta^b \in \Delta^+$ such that $\beta = \beta^b + \alpha_j$.*

Proof. If not, then for each $\alpha_j \in \Sigma$, β is at the bottom of an α_j-string of roots, $\beta + \nu\alpha_j,\ 0 \le \nu \le p_j$, for some $p_j \in \mathbb{Z}^+$. As in (6.6.15), we have

$$2\,\frac{\langle \beta, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} = -p_j \le 0,$$

for each $\alpha_j \in \Sigma$. Then the argument proving Proposition 6.6.3 would imply that $\Sigma \cup \{\beta\}$ is linearly independent, which contradicts Proposition 6.6.4. $\square$

Let us now assume $\mathfrak{g}_{\mathbb{C}}$ is simple (so $\mathfrak{z} = 0$). We will show that the Cartan matrix A, with entries (6.6.7) (for $\Sigma = \{\alpha_1, \dots, \alpha_m\}$) determines all the roots of $\mathfrak{g}$. It suffices to determine all the positive roots. Each $\beta \in \Delta^+$ can be written

$$\beta = \sum_j k_j \alpha_j, \quad k_j \in \mathbb{Z}^+. \tag{6.6.19}$$

Call $\Sigma_j k_j$ the level of the root β, so the simple roots have level 1. Suppose one has determined all the roots up to level n, and wants to determine those at level $n+1$. For each root $\beta \in \Delta^+$ at level n, and each $\alpha_j \in \Sigma$, we need to determine whether $\beta + \alpha_j$ is a root. By Proposition 6.6.7, this will suffice to determine all the roots at level $n+1$.

Given such knowledge of roots up to level n, for each root β of level n, we know how far back the α_j-string of roots extends: $\beta, \beta - \alpha_j, \dots, \beta - m\alpha_j$. From here, we can use (6.2.34) to determine how far forward the string extends: $\beta, \beta + \alpha_j, \dots, \beta + p\alpha_j$. We have

$$\begin{aligned} m - p &= 2\,\frac{\langle \beta, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} \\ &= 2 \sum_\ell k_\ell \frac{\langle \alpha_\ell, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} \\ &= \sum_\ell k_\ell A_{\ell j}. \end{aligned} \tag{6.6.20}$$

Thus, using Proposition 6.2.6, we see that $\beta + \alpha_j$ is a root if and only if

$$p = m - \sum_\ell k_\ell A_{\ell j} > 0. \tag{6.6.21}$$

In case $\mathfrak{g} = \mathfrak{su}(n)$, with the simple roots (6.6.4), we have the following:

$$\begin{aligned} &\text{level } 2: && \{\omega_{j,j+2} : 1 \le j \le n-2\}, \\ &\text{level } 3: && \{\omega_{j,j+3} : 1 \le j \le n-3\}, \\ &\quad\vdots \\ &\text{level } n-1: && \{\omega_{1n}\}. \end{aligned} \tag{6.6.22}$$

This can be seen directly from (6.6.3). However, the point of the discussion above is that it can also be deduced, without a priori knowledge of the set of positive roots of $\mathfrak{g}$, from the Cartan matrix, via (6.6.21).

To summarize, we start with the linear space $\mathfrak{h}$, $\Sigma \subset \mathfrak{h}'$, and the Cartan matrix A. We assume there exists a simple Lie algebra $\mathfrak{g}$, of which $\mathfrak{h}$ is a maximal commutative Lie subalgebra, for which Σ is the set of simple roots. Given that the center $\mathfrak{z}$ of $\mathfrak{g}$ is 0, we know by Corollary 6.6.5 that Σ is a basis of $\mathfrak{h}'$. We assume that A is given by (6.6.7), where $\mathfrak{g}$ has an Ad-invariant inner product, inducing an inner product on $\mathfrak{h}$ and hence on $\mathfrak{h}'$. Now we have constructed $\Delta^+ \subset \mathfrak{h}'$. These roots and their negatives form $\Delta \subset \mathfrak{h}'$.

To see how much of the structure of $\mathfrak{g}$ is revealed by the data $\mathfrak{h}$, Σ, and A, we will find the following result useful.

Proposition 6.6.8. *Assume the Lie algebra $\mathfrak{g}$ of the compact Lie group G has the property that $\mathfrak{g}_{\mathbb{C}}$ is simple. Then Σ cannot be partitioned into two disjoint nonempty sets $\Sigma_0 \cup \Sigma_1$ such that $\alpha_j \perp \alpha_k$ whenever $\alpha_j \in \Sigma_0$ and $\alpha_k \in \Sigma_1$. Consequently the Dynkin diagram of $\mathfrak{g}$ is connected. Furthermore, the Cartan matrix of $\mathfrak{g}$ determines the inner product on $\mathfrak{h}$ uniquely, up to a constant factor.*

Proof. It follows from (6.2.34) that

$$\alpha_j, \alpha_k \in \Sigma,\ \langle \alpha_j, \alpha_k \rangle = 0 \Longrightarrow [e_{\alpha_j}, e_{\alpha_k}] = 0, \tag{6.6.23}$$

since (by (6.6.13)) in this setting $m = 0$ in (6.2.34), hence $p = 0$. Thus if there were such a partition $\Sigma = \Sigma_0 \cup \Sigma_1$, we would have

$$\alpha_j \in \Sigma_0,\ \alpha_k \in \Sigma_1 \Longrightarrow [e_{\alpha_j}, e_{\alpha_k}] = 0. \tag{6.6.24}$$

It follows that, if $\mathfrak{n}_+^0$ is the Lie algebra generated by $\{e_{\alpha_j} : \alpha_j \in \Sigma_0\}$ and $\mathfrak{n}_+^1$ that generated by $\{e_{\alpha_k} : \alpha_k \in \Sigma_1\}$, then

$$X_0 \in \mathfrak{n}_+^0,\ X_1 \in \mathfrak{n}_+^1 \Longrightarrow [X_0, X_1] = 0. \tag{6.6.25}$$

Furthermore, Proposition 6.6.6 and its proof imply

$$\mathfrak{n}_+ = \mathfrak{n}_+^0 \oplus \mathfrak{n}_+^1. \tag{6.6.26}$$

Similarly, if $\mathfrak{n}_-^0$ is the Lie algebra generated by $\{e_{-\alpha_j} : \alpha_j \in \Sigma_0\}$ and $\mathfrak{n}_-^1$ that generated by $\{e_{-\alpha_k} : \alpha_k \in \Sigma_1\}$, then

$$Y_0 \in \mathfrak{n}_-^0,\ Y_1 \in \mathfrak{n}_-^1 \Longrightarrow [Y_0, Y_1] = 0, \tag{6.6.27}$$

and

$$\mathfrak{n}_- = \mathfrak{n}_-^0 \oplus \mathfrak{n}_-^1. \tag{6.6.28}$$

Furthermore, by Proposition 6.6.1,

$$\alpha_j \in \Sigma_0,\ \alpha_k \in \Sigma_1 \Longrightarrow [e_{\alpha_j}, e_{-\alpha_k}] = 0 = [e_{-\alpha_j}, e_{\alpha_k}], \tag{6.6.29}$$

so

$$[\mathfrak{n}_+^0, \mathfrak{n}_-^1] = 0 = [\mathfrak{n}_+^1, \mathfrak{n}_-^0]. \tag{6.6.30}$$

Now, $\mathfrak{n}_+$ and $\mathfrak{n}_-$ generate $\mathfrak{g}_{\mathbb{C}}$, and, by (6.6.25)–(6.6.30), if $\mathfrak{n}_+^0$ and $\mathfrak{n}_-^0$ generate $\mathfrak{G}^0$ and $\mathfrak{n}_+^1$ and $\mathfrak{n}_-^1$ generate $\mathfrak{G}^1$, then

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{G}^0 \oplus \mathfrak{G}^1, \quad [\mathfrak{G}^0, \mathfrak{G}^1] = 0. \tag{6.6.31}$$

This implies $\mathfrak{g}_{\mathbb{C}}$ is not simple. This proves the first assertion of Proposition 6.6.8. The connectedness of the Dynkin diagram is an immediate consequence.

We now address the degree to which the formula (6.6.7) for the Cartan matrix determines the inner product on $\mathfrak{h}'$, hence on $\mathfrak{h}$. Setting $|\alpha|^2 = \langle \alpha, \alpha \rangle$, we have

$$\alpha_j, \alpha_k \in \Sigma,\ \langle \alpha_j, \alpha_k \rangle \neq 0 \Longrightarrow \frac{A_{jk}}{A_{kj}} = \frac{|\alpha_j|^2}{|\alpha_k|^2}. \tag{6.6.32}$$

Suppose we select some value $a > 0$ of $|\alpha_1|$. By the first part of Proposition 6.6.8, there is a path from α_1 to each α_ℓ through simple roots whose nearest neighbors satisfy (6.6.32), so $|\alpha_\ell|$ is then determined uniquely for each $\alpha_\ell \in \Sigma$. Then the identity

$$\langle \alpha_j, \alpha_k \rangle = \frac{1}{2} |\alpha_k|^2 A_{jk} \tag{6.6.33}$$

uniquely determines the inner product $\langle \alpha_j, \alpha_k \rangle$ for each $\alpha_j, \alpha_k \in \Sigma$. Since $\{\alpha_j \in \Sigma\}$ is a basis for $\mathfrak{h}'$, this uniquely determines the inner product on $\mathfrak{h}'$ and finishes the proof of Proposition 6.6.8. □

Let us note that the proof of Proposition 6.6.8 produces a factorization of the Cartan matrix,

$$A = 2GD, \tag{6.6.34}$$

where

$$D = \operatorname{diag}(|\alpha_1|^{-2}, \dots, |\alpha_m|^{-2}), \quad G = (G_{jk}), \quad G_{jk} = \langle \alpha_j, \alpha_k \rangle. \tag{6.6.35}$$

unique up to a positive scalar factor on D (and its inverse on G). All the diagonal entries of D are positive, and G is a positive definite matrix.

We now record some necessary conditions on a matrix A for it to be a Cartan matrix.

Proposition 6.6.9. *If $A = (A_{jk})$ is the Cartan matrix associated to the Lie algebra $\mathfrak{g}$ of a compact Lie group, whose complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$ is simple, then*

$$A_{jj} = 2 \quad \text{and} \quad j \neq k \Rightarrow A_{jk} \in \{0, -1, -2, -3\}, \tag{6.6.36}$$

$$A_{jk} \neq 0 \Longleftrightarrow A_{kj} \neq 0, \tag{6.6.37}$$

$$A_{jk}, A_{j\ell}, A_{k\ell} \neq 0 \Rightarrow \frac{A_{j\ell}}{A_{\ell j}} = \frac{A_{k\ell}}{A_{\ell k}} \cdot \frac{A_{jk}}{A_{kj}}, \tag{6.6.38}$$

$$\det A \neq 0. \tag{6.6.39}$$

Proof. The result (6.6.36) follows from (6.2.27). Each condition in (6.6.37) is equivalent to $\langle \alpha_j, \alpha_k \rangle \neq 0$. The conclusion in (6.6.38) follows from

$$\frac{|\alpha_j|^2}{|\alpha_\ell|^2} = \frac{|\alpha_k|^2}{|\alpha_\ell|^2} \cdot \frac{|\alpha_j|^2}{|\alpha_k|^2}, \tag{6.6.40}$$

in light of (6.6.32). Actually, we can rewrite (6.6.38) as

$$A_{j\ell} A_{\ell k} A_{kj} = A_{\ell j} A_{k\ell} A_{jk}, \tag{6.6.41}$$

and, given (6.6.37), we can say that this holds even in the absence of the nonvanishing hypothesis in (6.6.38).

The result (6.6.39) follows from (6.6.34)–(6.6.35). ☐

Note. In the example (6.6.9), $A_{jk} = A_{kj}$, but this is not always the case. Specializing (6.2.28), we have, when $|\alpha_j| \geq |\alpha_k|$,

$$\begin{aligned} A_{jk} = -1 &\Longleftrightarrow A_{kj} = -1, \\ A_{jk} = -2 &\Longleftrightarrow A_{kj} = -1, \\ A_{jk} = -3 &\Longleftrightarrow A_{kj} = -1. \end{aligned} \tag{6.6.42}$$

Let us construct the Cartan matrix associated to $\mathfrak{so}(5)$. As will be seen in §7.1, $\dim \mathfrak{so}(5) = 10$ and $\dim \mathfrak{h} = 2$, so there are 8 roots, and 4 of them are postive. As shown in Figure 7.1.1, with respect to a natural basis of $\mathfrak{h}$, they are

$$\beta_2^- = (0,1), \quad \alpha_4 = (1,-1), \quad \beta_2^+ = (1,0), \quad \alpha_2 = (1,1). \tag{6.6.43}$$

Then

$$\Sigma = \{\alpha_4, \beta_2^-\}, \tag{6.6.44}$$

and $\beta_2^+ = \alpha_4 + \beta_2^-$, $\alpha_2 = \beta_2^- + \beta_2^+$. We compute as above, with $\tilde{\alpha}_1 = \alpha_4$, $\tilde{\alpha}_2 = \beta_2^-$, so

$$\begin{aligned} A_{12} &= 2\,\frac{\langle \alpha_4, \beta_2^- \rangle}{\langle \beta_2^-, \beta_2^- \rangle} = -2, \\ A_{21} &= 2\,\frac{\langle \alpha_4, \beta_2^- \rangle}{\langle \alpha_4, \alpha_4 \rangle} = -1. \end{aligned} \tag{6.6.45}$$

Hence

$$A = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}. \tag{6.6.46}$$

The associated Dynkin diagram is given in Figure 6.6.1.

Let us see how the material laid out after Proposition 6.6.7 enables us to construct the root system (6.6.43) from the Cartan matrix (6.6.46). We will change notation to

$$\Sigma = \{\alpha_1, \alpha_2\}, \tag{6.6.47}$$

in place of $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$, given above by $\tilde{\alpha}_1 = \alpha_4, \tilde{\alpha}_2 = \beta_2^-$. We have

$$2\,\frac{\langle \alpha_1, \alpha_2 \rangle}{|\alpha_2|^2} = A_{12} = -2, \quad 2\,\frac{\langle \alpha_1, \alpha_2 \rangle}{|\alpha_1|^2} = A_{21} = -1, \tag{6.6.48}$$

hence

$$\frac{|\alpha_1|^2}{|\alpha_2|^2} = 2. \tag{6.6.49}$$

The angle θ_{12} between α_1 and α_2 satisfies

$$\cos^2 \theta_{12} = \frac{1}{4} A_{12} A_{21} = \frac{1}{2}; \tag{6.6.50}$$

cf. (6.2.15). Hence, taking the negative square root, by Proposition 6.6.2,

$$\cos \theta_{12} = \frac{\langle \alpha_1, \alpha_2 \rangle}{|\alpha_1| \cdot |\alpha_2|} = -\frac{1}{\sqrt{2}}, \quad \text{so} \quad \theta_{12} = \frac{3\pi}{4}. \tag{6.6.51}$$

We can take α_2 to be $(0,1) \in \mathbb{R}^2$, and then $\alpha_1 = (1,-1)$.

As for the members of Δ^+, we have

$$\begin{aligned} \text{level } 1: &\quad \alpha_1,\ \alpha_2, \\ \text{level } 2: &\quad \alpha_1 + \alpha_2, \\ \text{level } 3: &\quad \alpha_1 + 2\alpha_2. \end{aligned} \tag{6.6.52}$$

We explain these results. First, the set Σ of simple roots always makes up level 1. If Σ has just 2 elements, the only possible element of level 2 is their sum. One can check the α_1-string through α_2 to verify that this sum is a root, in the current setting. To investigate level 3, we examine two strings.

(I) The α_1-string through $\alpha_1 + \alpha_2$. Here $m = 1$ and

$$m - p = 2\,\frac{\langle \alpha_1 + \alpha_2, \alpha_1 \rangle}{|\alpha_1|^2} = 2 + A_{21}, \tag{6.6.53}$$

hence $m - p = 1$, so $p = 0$. This implies $\alpha_1 + \alpha_2 + \alpha_1$ is not a root.

(II) The α_2-string through $\alpha_1 + \alpha_2$. Again $m = 1$, and

$$m - p = 2\,\frac{\langle \alpha_1 + \alpha_2, \alpha_2 \rangle}{|\alpha_2|^2} = 2 + A_{12}, \tag{6.6.54}$$

hence $m - p = 0$, so $p = 1$. Thus $\alpha_1 + \alpha_2 + \alpha_2$ is a root. This exhausts level 3.

When A is given by (6.6.46), there are no level 4 roots. In fact, by Proposition 6.6.7, the only candidates for level 4 roots are $\alpha_1 + 2\alpha_2 + \alpha_1 = 2(\alpha_1 + \alpha_2)$ and $\alpha_1 + 2\alpha_2 + \alpha_2 = \alpha_1 + 3\alpha_2$. Proposition 6.2.4 implies the former is not a root, and the calculation $p = 1$ for the α_2-string through $\alpha_1 + \alpha_2$ implies $\alpha_1 + 3\alpha_2$ is not a root.

Thus the root system associated to the Cartan matrix (6.6.46) is as depicted in Figure 6.6.2. This figure is identical to Figure 7.1.1, except for the re-labeling of the roots.

As we have seen in (6.6.9), the Cartan matrix associated to SU(3) is

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. \tag{6.6.55}$$

Its root system is depicted in Figure 6.6.3. Here, the level 1 and level 2 roots are as in (6.6.52). The calculation (6.6.53) for the α_1-string through $\alpha_1+\alpha_2$ again gives $p = 0$, and this time the calculation for the α_2-string through $\alpha_1+\alpha_2$ also gives $p = 0$, so there are no level 3 roots. This is consistent with the observation that SU(3) has dimension 8 (and a 2-dimensional maximal torus), so it has 6 roots, 3 of them positive. (As we have also seen, these positive roots are ω_{jk}, $1 \le j < k \le 3$.)

Another candidate for a 2×2 Cartan matrix is

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}. \tag{6.6.56}$$

This turns out to be the Cartan matrix associated to the exceptional Lie group known as G_2. We construct its root system.

Use the notation (6.6.47) for Σ. This time, we have

$$2\,\frac{\langle \alpha_1, \alpha_2 \rangle}{|\alpha_2|^2} = A_{12} = -3, \quad 2\,\frac{\langle \alpha_1, \alpha_2 \rangle}{|\alpha_1|^2} = A_{21} = -1, \tag{6.6.57}$$

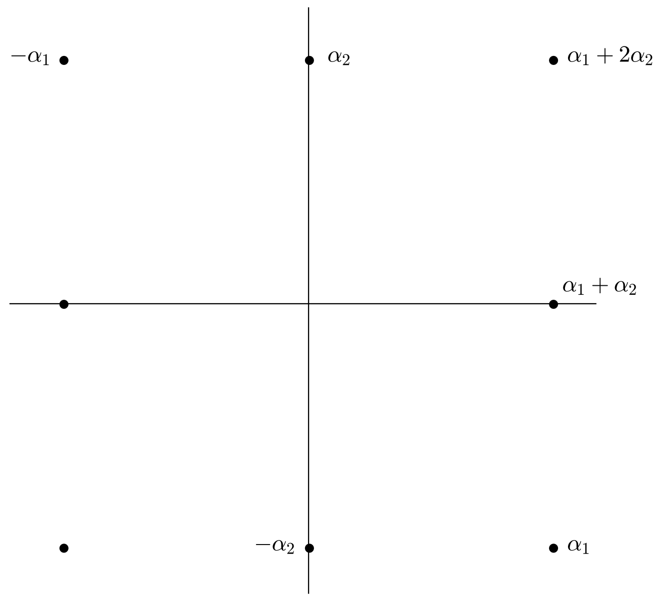

Figure 6.6.2. Root system of SO(5)

hence

$$\frac{|\alpha_1|^2}{|\alpha_2|^2} = 3. \tag{6.6.58}$$

The angle θ_{12} between α_1 and α_2 satisfies

$$\cos^2 \theta_{12} = \frac{1}{4} A_{12} A_{21} = \frac{3}{4}, \tag{6.6.59}$$

hence (again taking Proposition 6.6.2 into account)

$$\cos \theta_{12} = -\frac{\sqrt{3}}{2}, \quad \text{so} \quad \theta_{12} = \frac{5\pi}{6}. \tag{6.6.60}$$

We can take

$$\alpha_1 = (0, 1), \quad \text{hence} \quad \alpha_2 = \Big(\frac{\sqrt{3}}{6}, -\frac{1}{2}\Big), \tag{6.6.61}$$

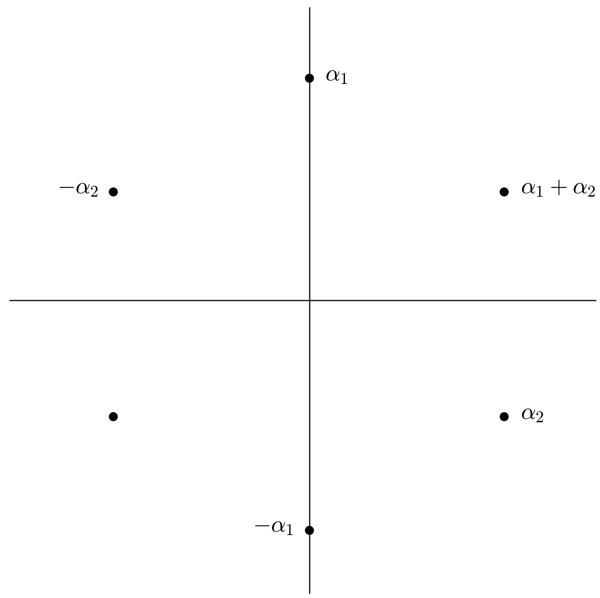

Figure 6.6.3. Root system of SU(3)

since $\langle\alpha_1,\alpha_2\rangle = (1/2)|\alpha_1|^2A_{21} = -1/2$ while $|\alpha_2|^2 = 1/3$, and $1/3 = 1/4 + 1/12$. As for the members of Δ^+, we have the following.

$$\begin{aligned} \text{level } 1: &\quad \alpha_1,\ \alpha_2, \\ \text{level } 2: &\quad \alpha_1+\alpha_2, \\ \text{level } 3: &\quad \alpha_1+2\alpha_2, \\ \text{level } 4: &\quad \alpha_1+3\alpha_2, \\ \text{level } 5: &\quad 2\alpha_1+3\alpha_2. \end{aligned} \tag{6.6.62}$$

The identities of levels 1 and 2 are by the same reasoning as applied to (6.6.52). To investigate level 3, we again examine two strings.

(I) The α_1-string through $\alpha_1+\alpha_2$. Then $m = 1$ and (6.6.53) applies, again yielding $p = 0$, so $2\alpha_1+\alpha_2$ is not a root.

(II) The α_2-string through $\alpha_1+\alpha_2$. Again $m = 1$, and (6.6.54) applies, this time yielding $p = 2$. Thus $\alpha_1+2\alpha_2$ is a root, the only level 3 root, and $\alpha_1+3\alpha_2$ is also a root (a level 4 root).

We have exhausted level 3. The only candidates for level 4 are $\alpha_1 + 2\alpha_2 + \alpha_2$ (seen above to be a root) and $\alpha_1 + 2\alpha_2 + \alpha_1 = 2(\alpha_1 + \alpha_2)$, which is not a root, by Propsition 36.4. This takes care of level 4.

Hence there are two candidates for level 5 roots, $2\alpha_1+3\alpha_2$ and $\alpha_1+4\alpha_2$. Again we have two strings to examine.

(I) The α_1-string through $\alpha_1 + 3\alpha_2$. Here $m = 0$, since $3\alpha_2$ is not a root. Then

$$m - p = 2\,\frac{\langle \alpha_1 + 3\alpha_2, \alpha_1\rangle}{|\alpha_1|^2} = 2 + 3A_{21}, \tag{6.6.63}$$

hence $m - p = -1$, so $p = 1$. Thus $2\alpha_1 + 3\alpha_2$ is a root.

(II) The α_2-string through $\alpha_1 + 3\alpha_2$. We have already seen that $\alpha_1 + 3\alpha_2$ is the end of the α_2-string through $\alpha_1 + \alpha_2$, so $\alpha_1 + 4\alpha_2$ is not a root.

There are no level 6 roots. Candidates would be $3\alpha_1 + 3\alpha_2 = 3(\alpha_1 + \alpha_2)$ and $2\alpha_1 + 4\alpha_2 = 2(\alpha_1 + 2\alpha_2)$, not roots by Proposition 6.2.4. The root system associated to the Cartan matrix (6.6.56) is depicted in Figure 6.6.4. Its Dynkin diagram is shown in Figure 6.6.5. Note that we have 12 roots, so $\dim G_2 = 14$.

We will identify G_2 as $\mathrm{Aut}(\mathbb{O})$ in Chapter 11.

We return to the classical groups and take a look at $\mathrm{SO}(n)$. For this, we peek ahead to §7.2. For $n = 2k$, the roots of $\mathfrak{so}(n)$ are given by (7.2.45). The positive roots are

$$\begin{array}{c} (0, \dots, 1, \dots, \varepsilon, \dots, 0) \quad (k\text{-tuple}), \quad \varepsilon = \pm 1, \\ \uparrow \qquad \uparrow \phantom{\dots, 0) \quad (k\text{-tuple}), \quad \varepsilon = \pm 1,} \\ i_1 \qquad i_2 \phantom{\dots, 0) \quad (k\text{-tuple}), \quad \varepsilon = \pm 1,} \end{array} \tag{6.6.64}$$

From this, one sees that the simple roots are

$$\alpha_1 = (1, -1, 0, \dots, 0),\ \ \alpha_2 = (0, 1, -1, \dots, 0), \quad \alpha_{k-1} = (0, \dots, 0, 1, -1), \tag{6.6.65}$$

and

$$\alpha_k = (0, \dots, 0, 1, 1). \tag{6.6.66}$$

Thus each $|\alpha_j|^2 = 2$ and $A_{j\ell} = A_{\ell j} = \langle \alpha_j, \alpha_\ell\rangle$. Differences from (6.6.8) are that

$$\langle \alpha_k, \alpha_{k-1}\rangle = 0, \quad \langle \alpha_k, \alpha_{k-2}\rangle = -1. \tag{6.6.67}$$

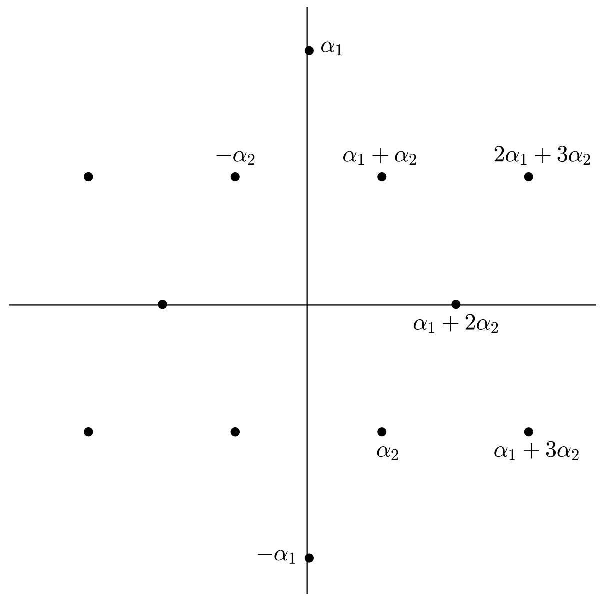

Figure 6.6.4. Root system for G_2

The Cartan matrix for $\mathrm{SO}(2k)$ is given by

$$\begin{pmatrix} 2 & -1 & & & & & \\ -1 & 2 & \ddots & & & & \\ & -1 & \ddots & -1 & & & \\ & & \ddots & 2 & -1 & & \\ & & & -1 & 2 & -1 & -1 \\ & & & & -1 & 2 & 0 \\ & & & & -1 & 0 & 2 \end{pmatrix}. \tag{6.6.68}$$

The associated Dynkin diagram is depicted in Figure 6.6.5.

As we show in §7.2, for $n = 2k+1$, the roots of $\mathfrak{so}(n)$ are given by (7.2.45) plus (7.2.50). The positive roots consist of (6.6.64) plus

$$(0,\dots,\underset{\substack{\uparrow\\ i}}{\varepsilon},\dots,0), \quad (k\text{-tuple}), \quad \varepsilon = \pm 1. \tag{6.6.69}$$

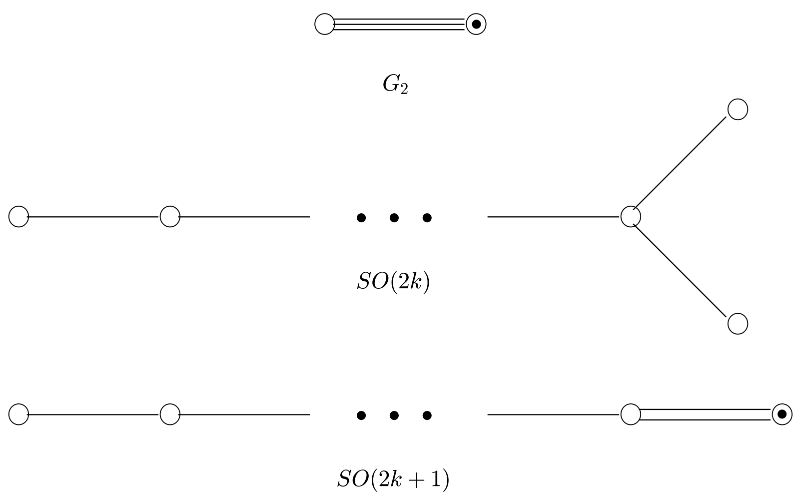

Figure 6.6.5. Dynkin diagrams of G_2, $SO(2k)$, and $SO(2k+1)$

Then the simple roots are given by (6.6.65) plus

$$\alpha_k = (0, \ldots, 0, 1). \tag{6.6.70}$$

In this case, $|\alpha_j|^2 = 2$ for $1 \le j \le k-1$ and $|\alpha_k|^2 = 1$. As opposed to (6.6.67),

$$A_{k-1,k} = 2\langle \alpha_k, \alpha_{k-1} \rangle = -2, \quad \langle \alpha_k, \alpha_{k-2} \rangle = 0. \tag{6.6.71}$$

The Cartan matrix for $\mathfrak{so}(2k+1)$ is given by

$$\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & \ddots & & \\ & -1 & \ddots & -1 & \\ & & \ddots & 2 & -2 \\ & & & -1 & 2 \end{pmatrix}. \tag{6.6.72}$$

The associated Dynkin diagram is depicted in Figure 6.6.5.

The Cartan-Killing classification

Compact Lie groups studied in this section fall into the following categories, known as the Cartan-Killing classification:

$A_n \quad SU(n+1)$

$B_n \quad SO(2n+1)$

$C_n \quad Sp(n)$

$D_n \quad SO(2n)$

In each case, the subscript n refers to the dimension of the maximal torus in the group. These groups (together with $U(n)$) are known as the classical compact Lie groups. The family $Sp(n)$ was introduced in §1.2. Their root systems will be analyzed in Chapter 10. In addition to these classical groups, there are five exceptional groups:

$G_2, \quad F_4, \quad E_6,\ E_7,\ E_8.$

As mentioned above, G_2 will appear in Chapter 11, as the automorphism group of the algebra of octonions $\mathbb{O}$.

To be more precise, the Cartan-Killing classification is a classification of the complexifications of the Lie algebras of these groups. Various noncompact Lie groups have Lie algebras with isomorphic complexifications, for example:

$A_n \quad Sl(n+1,\mathbb{R})$

$B_n \quad SO(2n,1)$

$C_n \quad Sp(n,\mathbb{R})$

$D_n \quad SO(2n-1,1).$

More material on these noncompact Lie groups can be found in [**38**].

6.7. Representations of disconnected compact Lie groups

Previous sections in this chapter have restricted attention to compact connected Lie groups. If G is not connected, the Lie algebra approach needs to be supplemented by other techniques. We look into some cases here.

One paradigmatic case is $O(n)$, which has two connected components,

$$O(n) = O_+(n) \cup O_-(n), \quad \det A = \pm 1 \text{ for } A \in O_\pm(n). \tag{6.7.1}$$

We also write $O_+(n) = SO(n)$. The case $O(2)$ was treated in Chapter 2, §2.5. More generally, we consider

$$G = G_+ \cup G_-, \quad e \in G_+, \tag{6.7.2}$$

where G_+ and G_- are connected, and G_+ contains the identity element e of the compact group G. It is useful to have the following.

Proposition 6.7.1. *In the setting of* (6.7.2)*, with $G_\pm$ compact and connected, the following hold.*
(1) $x, y \in G_+ \Longrightarrow x^{-1}, xy \in G_+$.
(2) $x \in G_+,\ y \in G \Longrightarrow y^{-1}xy \in G_+$.

Proof. Property (1) is clear for $x = e$, then follows by connectedness for all $x \in G_+$. The same strategy works to establish property (2). ☐

Proposition 6.7.1 implies G_+ is a normal subgroup of G, and we have

$$G/G_+ \approx \Gamma_2, \quad \Gamma_2 = \{\pm 1\}. \tag{6.7.3}$$

Thus we have the natural group homomorphism

$$\vartheta : G \longrightarrow \Gamma_2, \quad \vartheta(g) = \pm 1 \text{ for } g \in G_\pm. \tag{6.7.4}$$

When $G = O(n)$, we have $\vartheta(g) = \det g$.

Sometimes one has $G \approx G_+ \times \Gamma_2$, where the direct product of groups has the multiplication law

$$(g_1, a_1)(g_2, a_2) = (g_1 g_2, a_1 a_2), \quad g_j \in G_+,\ a_j \in \Gamma_2.$$

For example:

Proposition 6.7.2. *We have*

$$O(n) \approx O_+(n) \times \Gamma_2 \quad \text{if } n \text{ is odd.} \tag{6.7.5}$$

Proof. If $I \in \mathcal{L}(\mathbb{R}^n)$ is the identity, then $\det(-I) = (-1)^n$, so $-I \in O_-(n)$ when n is odd. This readily leads to the product group law. ☐

More generally, we have the following.

Proposition 6.7.3. *In the setting of* (6.7.2), *with* $G_\pm$ *compact and connected, we have*

$$G \approx G_+ \times \Gamma_2 \tag{6.7.6}$$

if and only if G_- *contains an element* $\tilde{e}$ *that commutes with each* $g \in G$ *and satisfies* $\tilde{e}^2 = e$.

When (6.7.6) holds, the description of the irreducible unitary representations of G follows from that of G_+, as a consequence of Proposition 2.8.11. In particular, this holds for $O(n)$ with n odd.

As we saw in §2.5, matters are very different for $O(2)$. We claim $O(n)$ does not have the direct product structure (6.7.6) if n is even. We prove this with the aid of the following result.

Proposition 6.7.4. *The representation of* $O(n)$ *on* $\mathbb{C}^n$ *arising from*

$$O(n) \subset M(n, \mathbb{R}) \subset M(n, \mathbb{C}) \tag{6.7.7}$$

is irreducible for each $n \geq 2$.

Proof. Assume $V \subset \mathbb{C}^n$ is a nonzero linear space invariant under $O(n)$. Take

$$v = \sum a_j e_j \in V, \quad v \neq 0 \ \ (a_j \in \mathbb{C}), \tag{6.7.8}$$

where $\{e_j\}$ is the standard basis of $\mathbb{R}^n$ Define $T_k \in O(n)$ by $T_k e_j = e_j$ if $j \neq k$, $T_k e_k = -e_k$. Then $T_k v \in V$, so $S_k v = (1/2)(v - T_k v) \in V$. Note that

$$S_k v = a_k v_k. \tag{6.7.9}$$

Hence $a_k e_k \in V$ for each k. Since $v \neq 0$, we have $e_j \in V$ for some j. Then applying elements of $O(n)$ yields $e_\ell \in V$ for each ℓ, so $V = \mathbb{C}^n$, and we have the irreducibility. □

Remark. A variant of this argument shows that the standard representation of $SO(n)$ on $\mathbb{C}^n$ is irreducible for $n \geq 3$. See §7.2 for a generalization, regarding representations of $SO(n)$ on $\Lambda^\ell \mathbb{C}^n$.

Corollary 6.7.5. *If* n *is even,* $O_-(n)$ *has no element that commutes with each* $g \in O(n)$.

Proof. By Proposition 6.7.4 and Schur's lemma, if $A \in \mathcal{L}(\mathbb{C}^n)$ commutes with each element of $O(n)$, then $A = \alpha I$ for some $\alpha \in \mathbb{C}$. If $A \in O(n)$, this forces $\alpha \in \{\pm 1\}$. But if n is even, $-I$ belongs to $O_+(n)$, not $O_-(n)$. □

We recall that a key ingredient in the analysis of $O(2)$ in §2.5 was an element $u \in O_-(2)$ that switched the two standard basis elements e_1 and e_2 of $\mathbb{R}^2$. Taking this as a cue, we assume we have

$$u \in G_-, \quad u^2 = e, \tag{6.7.10}$$

where $e \in G_+$ is the identity element. This yields an automorphism of G_+,

$$\tau : G_+ \longrightarrow G_+, \quad \tau(g) = ugu. \tag{6.7.11}$$

Note that $\tau^2(g) = g$. Note also that each element of G_- can be uniquely written as gu with $g \in G_+$. The group law on G is specified by

$$h(gu) = (hg)u, \quad (gu)h = g\tau(h)u, \quad g, h \in G_+. \tag{6.7.12}$$

To proceed, let ρ be an irreducible unitary representation of G_+ on V. Define a representation π_ρ of G on $V \oplus V$ as follows. First, set

$$\pi_\rho(g) = \begin{pmatrix} \rho(g) & \\ & \tilde{\rho}(g) \end{pmatrix}, \quad g \in G_+, \ \tilde{\rho}(g) = \rho \circ \tau(g). \tag{6.7.13}$$

Set

$$T = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \in \mathcal{L}(V \oplus V). \tag{6.7.14}$$

Note that, for $g \in G_+$,

$$\pi_\rho(g)T = \begin{pmatrix} & \rho(g) \\ \tilde{\rho}(g) & \end{pmatrix}, \quad T\pi_\rho(g) = \begin{pmatrix} & \tilde{\rho}(g) \\ \rho(g) & \end{pmatrix}. \tag{6.7.15}$$

Hence, for $g \in G_+$,

$$\pi_\rho(g)T = T\pi_{\tilde{\rho}}(g), \quad T\pi_\rho(g)T = \pi_{\tilde{\rho}}(g). \tag{6.7.16}$$

Hence setting

$$\pi_\rho(u) = T \tag{6.7.17}$$

defines π_ρ as a representation of G on $V \oplus V$. We have

$$\pi_\rho(gu) = \begin{pmatrix} & \rho(g) \\ \tilde{\rho}(g) & \end{pmatrix}, \quad g \in G_+. \tag{6.7.18}$$

We have the character formulas

$$\chi_{\pi_\rho}(g) = \chi_\rho(g) + \chi_{\tilde{\rho}}(g), \quad \chi_{\pi_\rho}(gu) = 0, \quad g \in G_+. \tag{6.7.19}$$

Hence

$$\begin{aligned} \|\chi_{\pi_\rho}\|^2_{L^2(G)} &= \frac{1}{2}\|\chi_\rho + \chi_{\tilde{\rho}}\|^2_{L^2(G_+)} \\ &= 1 + \mathrm{Re}(\chi_\rho, \chi_{\tilde{\rho}})_{L^2(G_+)}. \end{aligned} \tag{6.7.20}$$

Now

$$\begin{aligned} (\chi_\rho, \chi_{\tilde{\rho}})_{L^2(G_+)} = &\ 0, \ \text{ if } \rho \text{ not } \approx \tilde{\rho}, \\ &\ 1, \ \text{ if } \rho \approx \tilde{\rho}. \end{aligned} \tag{6.7.21}$$

This yields the following result.

Proposition 6.7.6. *If ρ is an irreducible unitary representation of G_+ and the representation π_ρ of G is defined by (6.7.13)–(6.7.18), then π_ρ is irreducible if and only if ρ and $\tilde{\rho}$ are not equivalent representations of G_+. If $\rho \approx \tilde{\rho}$, then π_ρ decomposes into two irreducible pieces.*

Recall how this works for $O(2)$ in §2.5. A complete set of irreducible representations of $SO(2)$ is given by

$$\rho_n(R(\theta)) = e^{in\theta}, \quad n \in \mathbb{Z}, \tag{6.7.22}$$

we have $u \in O_-(2)$, yielding $\tau(R(\theta)) = R(-\theta)$, hence $\tilde{\rho}_n = \rho_{-n}$. All the representations $\pi_n = \pi_{\rho_n}$ of $O(2)$ are irreducible except for the case $n = 0$, and $\pi_{-n} \approx \pi_n$. The reducible case π_0 splits into two irreducible factors.

Contrast this with the case $G = O(3)$, $u = -I$, $\tau(g) = g$. Then $\tilde{\rho} = \rho$ for each irreducible representation ρ of $SO(3)$, and none of the representations π_ρ are irreducible; they all split into two pieces. Of course, we know the irreducible representations of $O(3)$, thanks to Proposition 6.7.2. The same goes for $O(n)$ whenever n is odd.

In light of these considerations, we turn to the construction of irreducible representations of G in case $\rho \approx \tilde{\rho}$, say

$$\rho \circ \tau(g) = B^{-1}\rho(g)B, \quad \forall\, g \in G_+, \tag{6.7.23}$$

with $B \in \mathcal{L}(V)$ unitary. Replacing g by $\tau(g)$ and recalling that $\tau^2(g) = g$, we have

$$\rho(g) = B^{-1}\rho(\tau(g))B, \tag{6.7.24}$$

hence

$$\rho(g) = B^{-2}\rho(g)B^2, \quad \forall\, g \in G_+. \tag{6.7.25}$$

Since ρ is irreducible, we have $B^2 = \beta^2 I$ for some $\beta \in S^1 \subset \mathbb{C}$, and replacing B by $\beta^{-1}B$ (and relabeling), we arrange that (6.7.23) holds with

$$B^2 = I. \tag{6.7.26}$$

When (6.7.23) and (6.7.26) hold, we define the representations $\pi_\rho^\pm$ of G on V by

$$\begin{aligned} \pi_\rho^+(g) &= \rho(g), \quad \pi_\rho^+(gu) = \rho(g)B, \\ \pi_\rho^-(g) &= \rho(g), \quad \pi_\rho^-(gu) = -\rho(g)B, \end{aligned} \tag{6.7.27}$$

for $g \in G_+$, and note that the maps $\pi_\rho^\pm : G \to \mathcal{L}(V)$ preserve products, in light of the identities

$$\begin{aligned} \pi_\rho^+(g)\pi_\rho^+(hu) &= \rho(gh)B = \pi_\rho^+(ghu), \\ \pi_\rho^+(hu)\pi_\rho^+(g) &= \rho(h)B\rho(g) = \rho(h)\tilde{\rho}(g) = \rho(h\tau(g))B, \\ \pi_\rho^+(hug) &= \pi_\rho^+(h\tau(g)u) = \rho(h\tau(g))B, \end{aligned} \tag{6.7.28}$$

and similar identities involving π_ρ^-. These representations have characters

$$\chi_{\pi_\rho^\pm} = \begin{aligned} &\operatorname{Tr}(\rho), &&\text{on } G_+, \\ \pm &\operatorname{Tr}(\rho B), &&\text{on } G_-. \end{aligned} \tag{6.7.29}$$

We have

$$\|\chi_{\pi_\rho^\pm}\|^2_{L^2(G)} = \frac{1}{2} \int\limits_{G_+} \Big[|\operatorname{Tr} \rho(g)|^2 + |\operatorname{Tr} \rho(g) B|^2 \Big] \, dg. \tag{6.7.30}$$

As seen in Exercise 7 of §2.4,

$$\int\limits_{G_+} |\operatorname{Tr} \rho(g) B|^2 \, dg = \frac{1}{d_\rho} \operatorname{Tr} BB^* = 1. \tag{6.7.31}$$

Hence

$$\|\chi_{\pi_\rho^\pm}\|^2_{L^2(G)} = 1, \tag{6.7.32}$$

and we have that the representations $\pi_\rho^\pm$ of G on V are irreducible. Furthermore,

$$(\chi_{\pi_\rho^+}, \chi_{\pi_\rho^-})_{L^2(G)} = \frac{1}{2} \int\limits_{G_+} \Big[|\operatorname{Tr} \rho(g)|^2 - |\operatorname{Tr} \rho(g) B|^2 \Big] \, dg = 0, \tag{6.7.33}$$

so the representations π_ρ^+ and π_ρ^- are inequivalent. We also have, for $g \in G_+$,

$$\begin{aligned} \chi_{\pi_\rho}(g) &= 2\chi_\rho(g) = \chi_{\pi_\rho^+}(g) + \chi_{\pi_\rho^-}(g), \\ \chi_{\pi_\rho}(gu) &= \quad 0 \quad = \chi_{\pi_\rho^+}(gu) + \chi_{\pi_\rho^-}(gu), \end{aligned} \tag{6.7.34}$$

so

$$\pi_\rho \approx \pi_\rho^+ \oplus \pi_\rho^-, \tag{6.7.35}$$

in this case.

The following result summarizes our analysis of the irreducible unitary representations of $G = G_+ \cup G_-$.

Proposition 6.7.7. *Take compact $G = G_+ \cup G_-$, as in Proposition 6.7.1. Assume $u \in G_-$ satisfies* (6.7.10)*, and take τ as in* (6.7.11)*. Let $\{\rho_\alpha : \alpha \in \mathcal{I}\}$ denote a complete set of irreducible unitary representations of G_+. If $\rho_\alpha \approx \rho_\beta \circ \tau$, eliminate one of these, to form $\{\rho_\alpha : \alpha \in \mathcal{I}^b\}$.*

Then form $\pi_\alpha = \pi_{\rho_\alpha}$ as in (6.7.13)–(6.7.18)*. If $\rho_\alpha \approx \rho_\alpha \circ \tau$, decompose $\pi_\alpha \approx \pi_\alpha^+ \oplus \pi_\alpha^-$. Otherwise, keep π_α, $\alpha \in \mathcal{I}^b$.*

The representations so obtained form a complete set of irreducible unitary representations of G.

Proof. It just remains to check completeness. This follows from the fact that, as ρ ranges over ρ_α, $\alpha \in \mathcal{I}^b$, the matrix entries of π_ρ, given on G by (6.7.13) and (6.7.18), have dense linear span in $L^2(G)$. □

REMARK. The element $u \in G_-$ satisfying (6.7.10) need not be unique. Say also

$$v \in G_-, \quad v^2 = e, \quad \sigma(g) = vgv \ \text{ for } \ g \in G_+. \tag{6.7.36}$$

Then $v = hu$ for some $h \in G_+$, and $\sigma(g) = (hu)g(hu)^{-1} = h(ugu)h^{-1}$, so

$$\sigma(g) = h\tau(g)h^{-1}. \tag{6.7.37}$$

Hence, if ρ is a representation of G_+, we have

$$\rho \circ \sigma(g) = \rho(h)\, \rho \circ \tau(g)\, \rho(h)^{-1}, \tag{6.7.38}$$

so $\rho \circ \sigma$ and $\rho \circ \tau$ are equivalent representations. In particular,

$$\rho \approx \rho \circ \tau \Longleftrightarrow \rho \approx \rho \circ \sigma. \tag{6.7.39}$$

Representations of O(4). As seen in §4.1 (Chapter 4), as a consequence of the 2-fold covering homomorphism $p : SU(2) \times SU(2) \to SO(4)$, a complete set of irreducible unitary representations of $SO(4)$ is given by

$$\gamma_{k\ell} \circ p(g_1, g_2) = D_{k/2}(g_1) \otimes D_{\ell/2}(g_2), \quad k + \ell \text{ even}, \tag{6.7.40}$$

with $g_j \in SU(2)$. The condition $k + \ell$ even guarantees $\gamma_{k\ell} \circ p(\pm(I, I)) = I$. A maximal torus in $SO(4)$ is given by

$$\begin{pmatrix} R(\theta) & \\ & R(\varphi) \end{pmatrix}, \quad R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in SO(2). \tag{6.7.41}$$

As seen in the exercises for §4.1, we have

$$\operatorname{Tr} \gamma_{k\ell}\Bigl(\begin{pmatrix} R(\theta) & \\ & R(\varphi) \end{pmatrix}\Bigr) = \frac{\sin(k+1)(\theta+\varphi)/2}{\sin(\theta+\varphi)/2} \cdot \frac{\sin(\ell+1)(\theta-\varphi)/2}{\sin(\theta-\varphi)/2}. \tag{6.7.42}$$

Now we have the involution $\tau : SO(4) \to SO(4)$, given by conjugation by

$$\operatorname{diag}(-1, 1, 1, 1) \in O_-(4), \tag{6.7.43}$$

and a computation yields

$$\tau\Bigl(\begin{pmatrix} R(\theta) & \\ & R(\varphi) \end{pmatrix}\Bigr) = \begin{pmatrix} R(-\theta) & \\ & R(\varphi) \end{pmatrix}. \tag{6.7.44}$$

We have

$$\begin{aligned}
&\operatorname{Tr} \gamma_{k\ell} \circ \tau\Bigl(\begin{pmatrix} R(\theta) & \\ & R(\varphi) \end{pmatrix}\Bigr) \\
&= \frac{\sin(k+1)(\theta-\varphi)/2}{\sin(\theta-\varphi)/2} \cdot \frac{\sin(\ell+1)(\theta+\varphi)/2}{\sin(\theta+\varphi)/2} \\
&= \operatorname{Tr} \gamma_{\ell k}\Bigl(\begin{pmatrix} R(\theta) & \\ & R(\varphi) \end{pmatrix}\Bigr).
\end{aligned} \tag{6.7.45}$$

Hence

$$\gamma_{k\ell} \circ \tau \approx \gamma_{\ell k}, \tag{6.7.46}$$

and

$$\gamma_{k\ell} \circ \tau \approx \gamma_{k\ell} \Longleftrightarrow k = \ell. \tag{6.7.47}$$

Thus Proposition 6.7.7 gives the following conclusion.

Proposition 6.7.8. *A complete set of irreducible unitary representations of $O(4)$ is given as follows. Let*

$$\gamma_{k\ell}, \quad k, \ell \in \mathbb{N},\ k+\ell \text{ even.} \tag{6.7.48}$$

denote the complete set of irreducible unitary representations of $SO(4)$ produced in §4.1 and described above. With (6.7.46) *in mind, pare this down to*

$$(k,\ell) \in \mathcal{I}^b, \quad \text{i.e.,}\ \ k, \ell \in \mathbb{N},\ k+\ell \text{ even},\ k \geq \ell. \tag{6.7.49}$$

Then form $\pi_{k\ell} = \pi_{\gamma_{k\ell}}$ as in (6.7.13)–(6.7.18)*. If $k = \ell$, decompose*

$$\pi_{kk} \approx \pi_{kk}^{+} \oplus \pi_{kk}^{-}, \tag{6.7.50}$$

as indicated above. Otherwise, we keep $\pi_{k\ell}$. The representations so obtained form a complete set of irreducible unitary representations of $O(4)$.

We will investigate representations of $O(n)$ for even $n > 4$ in Chapter 7.

Many interesting compact Lie groups have more than two connected components. Let's consider a compact Lie group with k connected components. Denote by G_e the connected component of G containing the identity element e. Parallel to Proposition 6.7.1 we have:

Proposition 6.7.9. *If G_e is the connected component of the identity in G, then G_e is a normal subgroup of G.*

If G has k connected components, it follows that

$$G/G_e \approx \Gamma, \tag{6.7.51}$$

where Γ is a group with k elements, and we have a natural group homomorphism

$$\vartheta : G \longrightarrow \Gamma, \quad \operatorname{Ker}\vartheta = G_e. \tag{6.7.52}$$

Here is an example of a subgroup of $U(n)$ with k connected components.

The groups $S_kU(n)$. Let us set

$$S_kU(n) = \{g \in U(n) : \det g \in \Gamma_k\}, \quad \Gamma_k = \{e^{2\pi i\ell/k} : \ell \in \mathbb{Z}/(k)\}. \tag{6.7.53}$$

Elements of $G = S_kU(n)$ that are scalar multiples of I have the form

$$u_k^\nu, \quad \nu \in \mathbb{Z}/(nk), \quad u_k = e^{2\pi i/nk} I. \tag{6.7.54}$$

Note that

$$\det u_k = e^{2\pi i/k}, \quad u_k^{nk} = I. \tag{6.7.55}$$

Each connected component of G contains n of these elements. In particular, $G_e = SU(n)$ contains all powers of $u_k^k = e^{2\pi i/n}$. Parallel to (4.5.14), we have the exact sequence of groups

$$1 \longrightarrow \Gamma_n \longrightarrow \Gamma_{nk} \times SU(n) \longrightarrow S_kU(n) \longrightarrow 1, \tag{6.7.56}$$

where

$$\Gamma_{nk} = \{u_k^\nu : \nu \in Z/(nk)\}, \tag{6.7.57}$$

and

$$\Gamma_n = \{(\omega, g) \in \Gamma_{nk} \times SU(n) : g = \omega^{-1} I,\ \omega^n = 1\}, \tag{6.7.58}$$

a cyclic group of order n, generated by

$$(\zeta^{-1}, \zeta I), \quad \zeta = e^{2\pi i/n} = u_k^k. \tag{6.7.59}$$

We now look at the representations of $S_kU(n)$. Let $\{\rho_\alpha : \alpha \in \mathcal{I}\}$ denote a complete set of irreducible unitary representations of SU(n). (These were classified in Chapter 4.) By Proposition 2.8.11, a complete set of irreducible unitary representations of $\Gamma_{nk} \times SU(n)$ is given by $\{\pi_{m\alpha} : m \in \mathbb{Z}/(nk),\ \alpha \in \mathcal{I}\}$, defined by

$$\pi_{m\alpha}(\omega, g) = \omega^m \rho_\alpha(g), \quad \omega = u_k^\nu. \tag{6.7.60}$$

Such a representation of $\Gamma_{nk} \times SU(n)$ produces a representation of $S_kU(n)$ if and only if $\pi_{m\alpha}(\Gamma_n) = I$, i.e., if and only if

$$\rho_\alpha(\zeta I) = \zeta^m I, \tag{6.7.61}$$

with ζ as in (6.7.59). Now, since ζI is in the center of $SU(n)$, it follows that for each $\alpha \in \mathcal{I}$, $\rho_\alpha(\zeta I)$ is a scalar that is an nth root of unity, i.e.,

$$\rho_\alpha(\zeta I) = \zeta^\mu I, \quad \mu = \mu(\alpha) \in \mathbb{Z}. \tag{6.7.62}$$

Thus $\pi_{m\alpha}$ in (6.7.60) gives a representation of $S_kU(n)$ if and only if

$$m = \mu(\alpha), \quad \text{mod} \ \ n. \tag{6.7.63}$$

We have the following:

Proposition 6.7.10. *Let* $\{\rho_\alpha : \alpha \in \mathcal{I}\}$ *denote a complete set of irreducible unitary representations of* $SU(n)$*. Then each irreducible unitary representation of* $S_kU(n)$ *has the form* $\pi_{m\alpha}$*, given by*

$$\pi_{m\alpha}(\omega g) = \omega^m \rho_\alpha(g), \quad \omega = u_k^\nu, \ g \in SU(n), \tag{6.7.64}$$

for $\alpha \in \mathcal{I}$*,* $m \in \mathbb{Z}/(nk)$*, subject to the constraint* (6.7.62)–(6.7.63)*, with* $\zeta = e^{2\pi i/n}$*.*

6.A. Maximal tori

A torus $\mathbb{T}$ in a Lie group G is a compact abelian subgroup, and it is the image of its Lie algebra under the exponential map. Hence $\mathbb{T}$ is isomorphic to the quotient

$$\mathcal{T} = \mathbb{R}^n/\mathbb{Z}^n. \tag{6.A.1}$$

The group $\mathbb{T}$ is a multiplicative subgroup of G, but it is convenient to treat $\mathcal{T}$ as an additive group.

We begin with the following useful general property of tori.

Proposition 6.A.1. *If $\mathbb{T}$ is a torus, there exists $t_0 \in \mathbb{T}$ such that*

$$\{t_0^k : k \in \mathbb{Z}\} \ \ \textit{is dense in } \mathbb{T}. \tag{6.A.2}$$

For a proof, it suffices to show that there exists $\alpha \in \mathcal{T}$ such that

$$\{k\alpha : k \in \mathbb{Z}\} \ \ \text{is dense in } \mathcal{T}. \tag{6.A.3}$$

To produce such α, we note that $\mathbb{R}$ is a vector space over $\mathbb{Q}$ satisfying

$$\dim_{\mathbb{Q}} \mathbb{R} = \infty, \tag{6.A.4}$$

since $\mathbb{R}$ is uncountable. Now take $\alpha_0 = 1, \alpha_1, \dots, \alpha_n \in \mathbb{R}$, linearly independent over $\mathbb{Q}$, and set $\alpha = (\alpha_1, \dots, \alpha_n) \pmod{\mathbb{Z}^n}$. We claim (6.A.3) holds. A convenient route to this result involves considering the linear operator

$$T_\alpha : L^2(\mathcal{T}) \longrightarrow L^2(\mathcal{T}), \quad T_\alpha f(x) = f(x - \alpha). \tag{6.A.5}$$

We show that T_α has the following ergodic property.

Lemma 6.A.2. *Take α as above. If $f \in L^2(\mathcal{T})$ and $T_\alpha f = f$, then f is constant.*

Proof. We expand f as a Fourier series, with Fourier coefficients

$$\hat{f}(\ell) = \int_{\mathcal{T}} f(x) e^{-2\pi i \ell \cdot x}\, dx, \quad \ell \in \mathbb{Z}^n. \tag{6.A.6}$$

Then

$$\widehat{T_\alpha f}(\ell) = e^{-2\pi i \ell \cdot \alpha} \hat{f}(\ell), \tag{6.A.7}$$

so

$$T_\alpha f = f, \ \hat{f}(\ell) \neq 0 \Longrightarrow \ell \cdot \alpha \in \mathbb{Z}. \tag{6.A.8}$$

Given the linear independence of $\{\alpha_j : 0 \le j \le n\}$ over $\mathbb{Q}$, we see that if $T_\alpha f = f$, then $\hat{f}(\ell) = 0$ for $\ell \neq 0$, so f is constant. $\square$

We can now prove (6.A.3). Indeed, if this denseness fails, we have $\{k\alpha : k \in \mathbb{Z}\}$ disjoint from $B_\varepsilon(x_0)$, for some $x_0 \in \mathcal{T},\ \varepsilon > 0$. Hence

$$\mathcal{O} = \bigcup_{k\in\mathbb{Z}} B_{\varepsilon/2}(k\alpha) \text{ is disjoint from } B_{\varepsilon/2}(x_0). \tag{6.A.9}$$

Then

$$f = \chi_{\mathcal{O}} \Longrightarrow T\alpha f = f \text{ and } f = 0 \text{ on } B_{\varepsilon/2}(x_0), \tag{6.A.10}$$

contradicting Lemma 6.A.2.

To proceed, we say a torus $\mathbb{T} \subset G$ is a conjugating torus for G if each $x \in G$ is conjugate to some element of $\mathbb{T}$, or equivalently if

$$\bigcup_{g\in G} g^{-1}\mathbb{T}g = G. \tag{6.A.11}$$

It is elementary that

$$\mathbb{T} = \left\{ u(\theta) = \begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix} : \theta_j \in \mathbb{R} \right\} \tag{6.A.12}$$

is a conjugating torus for $U(n)$. Furthermore,

$$\mathbb{T} = \{u(\theta) : \theta_1 + \cdots + \theta_n = 0\} \tag{6.A.13}$$

is a conjugating torus for $SU(n)$. These results follow from the fact that for each $A \in U(n)$, $\mathbb{C}^n$ has an orthonormal basis of eigenvectors of A. Similarly, a use of Proposition 2.4.4, applied to $A \in SO(n)$, shows that
(6.A.14)

$$\mathbb{T} = \left\{ R_k(\theta) = \begin{pmatrix} R(\theta_1) & & \\ & \ddots & \\ & & R(\theta_k) \end{pmatrix} : \theta_j \in \mathbb{R} \right\}, \quad R(\theta_j) = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix},$$

is a conjugating torus for $SO(2k)$, and

$$\mathbb{T} = \left\{ \begin{pmatrix} R_k(\theta) & \\ & 1 \end{pmatrix} : \theta \in \mathbb{R}^k \right\} \tag{6.A.15}$$

is a conjugating torus for $SO(2k+1)$. In Chapter 10 it is proved that

(6.A.16) the torus (6.A.12) is also a conjugating torus for $Sp(n)$.

Knowing that one has a conjugating torus has the following implication.

Proposition 6.A.3. *Let G be a compact Lie group and assume $\mathbb{T} \subset G$ is a conjugating torus. Let $\mathbb{T}'$ be another torus in G. Then there exists $g \in G$ such that*

$$\mathbb{T}' \subset g^{-1}\mathbb{T}g. \tag{6.A.17}$$

Consequently $\mathbb{T}$ *is a maximal torus. In such a case, whenever* $\mathbb{T}' \subset G$ *is a maximal torus, one has* $g\mathbb{T}'g^{-1} = \mathbb{T}$*, for some* $g \in G$*.*

Proof. Pick $t_1 \in \mathbb{T}'$ such that $\{t_1^k : k \in \mathbb{Z}\}$ is dense in $\mathbb{T}'$. Since $\mathbb{T}$ is conjugating, there exists $g \in G$ such that

$$\begin{aligned} gt_1g^{-1} \in \mathbb{T}, \text{ hence } gt_1^k g^{-1} \in \mathbb{T},\ \forall\, k \in \mathbb{Z}, \\ \text{hence } g\mathbb{T}'g^{-1} \subset \mathbb{T}, \end{aligned} \tag{6.A.18}$$

so (6.A.17) holds. Finally, if the inclusion in (6.A.17) is strict, $\mathbb{T}'$ cannot be maximal. □

The following is a useful topological consequence of the existence of a conjugating torus.

Proposition 6.A.4. *Let* G *be a compact Lie group. If* G *has a conjugating torus* $\mathbb{T}$*, then* G *is connected.*

Proof. Take $x \in G$, and then take $g \in G$ such that $x \in g^{-1}\mathbb{T}g = \mathbb{T}'$. Then $\mathbb{T}'$ is a torus in G containing the identity element e, so there is a continuous path from x to e. □

Corollary 6.A.5. *The groups* $U(n)$*,* $SU(n)$ *and* $SO(n)$ *are all connected.*

In Chapter 10 we will show that (6.A.12) is also a conjugating torus for $Sp(n)$, obtaining incidentally that $Sp(n)$ is connected.

Chapter 7

The orthogonal groups $\mathrm{SO}(n)$ and their representations

We begin our treatment of the structure and representation theory of the groups $\mathrm{SO}(n)$ in §7.1 with a discussion of the cases $2 \le n \le 5$. Of these the case $\mathrm{SO}(2) \approx \mathbb{T}^1$ is elementary, and the cases $n = 3$ and 4 have been treated in §4.1, where they are related to $\mathrm{SU}(2)$. After a brief recollection of this, §7.1 concentrates on $\mathrm{SO}(5)$. This group has rank 2 and dimension 10, hence 8 roots, denoted α_j $(1 \le j \le 4)$ and $\beta_j^{\pm}$ $(1 \le j \le 2)$. We see that the adjoint representation of $\mathrm{SO}(5)$ is irreducible, with highest weight α_2 and that the standard representation on $\mathbb{C}^5$ is irreducible, with highest weight β_2^+. Hence, there are unique irreducible unitary representations of $\mathrm{SO}(5)$ with highest weights $j\alpha_2 + k\beta_2^+$, for arbitrary $j, k \in \mathbb{Z}^+$. Now, these elements of $\mathfrak{h}'$ are only "half" of the dominant integral weights, as specified by the Theorem of the Highest Weight in §6.2. The rest of the dominant integral weights will be associated to representations of the double cover $\mathrm{Spin}(5)$ later in this chapter. In this connection, recall from §4.1 that $\mathrm{SU}(2)$ is a double cover of $\mathrm{SO}(3)$ and $\mathrm{SU}(2) \times \mathrm{SU}(2)$ is a double cover of $\mathrm{SO}(4)$.

We move to the study of representations of $\mathrm{SO}(n)$ for general n in §7.2. As in Chapter 4, we use as building blocks the representations Λ^ℓ on $\mathbb{C}^n$, given by

$$\Lambda^\ell(g)v_1 \wedge \cdots \wedge v_\ell = gv_1 \wedge \cdots \wedge gv_\ell, \tag{7.0.1}$$

this time for $g \in \mathrm{SO}(n)$. Of course, the details are different for $\mathrm{SO}(n)$ than they were for $\mathrm{SU}(n)$. For one, now Λ^ℓ is irreducible for each $\ell \in \{0, \ldots, n\}$

except when n is even and $\ell = n/2$. Furthermore, for ℓ in this range, the representations Λ^ℓ and $\Lambda^{n-\ell}$ of SO(n) are equivalent. These phenomena arise thanks to the Hodge star operator. Also, for $n = 2k$, we have

$$\Lambda^k\mathbb{C}^{2k} = \Lambda^k_+\mathbb{C}^{2k} \oplus \Lambda^k_-\mathbb{C}^{2k}, \tag{7.0.2}$$

where $\Lambda^k_\pm\mathbb{C}^{2k}$ are eigenspaces for this star operator, and SO($2k$) acts on each summand, irreducibly. We analyze the weight vectors and weights of these representations, and identify the highest weights. We note that Λ^2 is equivalent to the adjoint representation of SO(n), which reveals the structure of the roots of $\mathfrak{so}(n)$. A description of the roots splits onto two cases: $n = 2k$ and $n = 2k+1$. We identify the dominant integral weights for SO(n) in these two cases. Extending material from §7.1, we see that "half" of these weights are positive integral combinations of the highest weights of representations described above. Again, the rest of these weights will be associated to representations of Spin(n).

Sections 7.3–7.6 are devoted to the construction of Spin(n), as a double cover of SO(n), and a study of its representations. We begin in §7.3 with the construction of Clifford algebras. Then §7.4 produces Spin(n) as a subset of $\mathcal{C}\ell(n,0)$, shows that it is a connected Lie group, and produces a surjective, 2-to-1 homomorphism

$$\tau : \mathrm{Spin}(n) \longrightarrow \mathrm{SO}(n). \tag{7.0.3}$$

Section 7.5 constructs spinor representations $D^\pm_{1/2}$ of Spin($2k$) onto $S_\pm(2k)$, where

$$S_+(2k) = \bigoplus_{j \text{ even}} \Lambda^j_{\mathbb{C}}\mathbb{C}^k, \quad S_-(2k) = \bigoplus_{j \text{ odd}} \Lambda^j_{\mathbb{C}}\mathbb{C}^k, \tag{7.0.4}$$

which are irreducible. There is a natural inclusion $\mathrm{Spin}(2k-1) \hookrightarrow \mathrm{Spin}(2k)$, yielding an irreducible representation of Spin($2k-1$) on $S_+(2k)$ (and also one on $S_-(2k)$, but for Spin($2k-1$) these two are equivalent). In §7.6 we study the weight spaces and highest weights of the spinor representations, and show that these together with the ones produced in §7.2 generate all the dominant integral weights, which implies that we have all the irreducible unitary representations of Spin(n). As a corollary, we have a "non-topological" proof that Spin(n) is simply connected.

7.1. Representations of SO(n), $n \le 5$

Before proceeding to general results, in the next section, here we describe maximal tori of $G = \mathrm{SO}(n)$, the root space decompositions, and the Weyl groups, when $n \le 5$. We start with $n = 2$. We have

$$\mathrm{SO}(2) \approx S^1, \quad \mathfrak{so}(2) = \mathfrak{h} \approx \mathbb{R}, \quad \text{no roots.} \tag{7.1.1}$$

Moving on to $n = 3$, as shown in §4.1, there is a 2-fold covering map

$$\mathrm{SU}(2) \longrightarrow \mathrm{SO}(3), \quad \text{hence } \ \mathfrak{so}(3) \approx \mathfrak{su}(2), \tag{7.1.2}$$

we have $\mathfrak{h}$ spanned by X_1, and root vectors $X_\pm = X_2 \mp iX_3$, with X_j as in (4.1.3). Hence

$$\dim \mathfrak{so}(3) = 3, \quad \operatorname{rank} \mathfrak{so}(3) = 1, \quad \text{there are 2 roots.} \tag{7.1.3}$$

Next, SO(4) was also studied in §4.1. As shown there, there is a 2-fold covering map

$$\mathrm{SU}(2) \times \mathrm{SU}(2) \longrightarrow \mathrm{SO}(4), \quad \text{hence } \ \mathfrak{so}(4) \approx \mathfrak{su}(2) \oplus \mathfrak{su}(2). \tag{7.1.4}$$

On each factor we have a piece of $\mathfrak{h}$ and a couple of root vectors (in the complexification), so

$$\dim \mathfrak{so}(4) = 6, \quad \operatorname{rank} \mathfrak{so}(4) = 2, \quad \text{there are 4 roots.} \tag{7.1.5}$$

As a warm-up for studying SO(5), we take a second look at the maximal torus of SO(4) and the roots of $\mathfrak{so}(4)$. We have

$$\mathbb{T} = \left\{ \begin{pmatrix} R_{\theta_1} & 0 \\ 0 & R_{\theta_2} \end{pmatrix} : \theta_j \in \mathbb{R}/2\pi\mathbb{Z} \right\}, \quad R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{7.1.6}$$

Equivalently,

$$\mathfrak{h} = \left\{ D_{a,b} = \begin{pmatrix} aJ & 0 \\ 0 & bJ \end{pmatrix} : a, b \in \mathbb{R} \right\}, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{7.1.7}$$

To get the root spaces $\mathfrak{g}_\alpha$, we decompose the following linear complement to $\mathfrak{h}_\mathbb{C}$ in $\mathfrak{so}(4)_\mathbb{C}$,

$$\left\{ A_C = \begin{pmatrix} 0 & C \\ -C^t & 0 \end{pmatrix} : C \in \mathrm{M}(2, \mathbb{C}) \right\}, \tag{7.1.8}$$

into 4 pieces, each of complex dimension 1, joint eigenvectors for the $\operatorname{ad} \mathfrak{h}$ action. A computation gives

$$[D_{a,b}, A_C] = \begin{pmatrix} 0 & aJC - bCJ \\ -bJC^t + aC^tJ & 0 \end{pmatrix}. \tag{7.1.9}$$

Thus we look for $C_k \in \mathrm{M}(2, \mathbb{C})$ such that

$$aJC_k - bC_kJ = i\alpha_k(a, b)C_k, \tag{7.1.10}$$

or equivalently,

$$JC_k = i\alpha_k(1,0)C_k, \quad C_kJ = i\alpha_k(0,1)C_k. \tag{7.1.11}$$

Such matrices can be found by inspection from the formulas

$$J\begin{pmatrix}1\\ \pm i\end{pmatrix} = \mp i\begin{pmatrix}1\\ \pm i\end{pmatrix}, \quad (1,\pm i)J = \pm i(1,\pm i). \tag{7.1.12}$$

One obtains

(7.1.13)

$$C_1 = \begin{pmatrix}1 & i\\ i & -1\end{pmatrix}, \quad C_2 = \begin{pmatrix}1 & -i\\ -i & -1\end{pmatrix}, \quad C_3 = \begin{pmatrix}1 & -i\\ i & 1\end{pmatrix}, \quad C_4 = \begin{pmatrix}1 & i\\ -i & 1\end{pmatrix},$$

for which (7.1.10)–(7.1.11) hold with

(7.1.14)

$$\alpha_1(a,b) = -(a+b), \quad \alpha_2(a,b) = a+b, \quad \alpha_3(a,b) = -(a-b), \quad \alpha_4(a,b) = a-b.$$

We now tackle the case $n = 5$. A maximal torus of SO(5) is given by

$$\mathbb{T} = \left\{ \begin{pmatrix} R_{\theta_1} & & \\ & R_{\theta_2} & \\ & & 1 \end{pmatrix} : \theta_j \in \mathbb{R}/2\pi\mathbb{Z} \right\}, \tag{7.1.15}$$

with R_θ as in (7.1.6). In this case,

$$\mathfrak{h} = \left\{ D_{a,b} = \begin{pmatrix} aJ & & \\ & bJ & \\ & & 0 \end{pmatrix} : a, b \in \mathbb{R} \right\}, \tag{7.1.16}$$

with J as in (7.1.7). Parallel to (7.1.5), we have

$$\dim \mathfrak{so}(5) = 10, \quad \operatorname{rank} \mathfrak{so}(5) = 2, \quad \text{there are 8 roots.} \tag{7.1.17}$$

Four of the root spaces are spanned by

$$\begin{pmatrix} 0 & C_j & \\ -C_j^t & 0 & \\ & & 0 \end{pmatrix}, \tag{7.1.18}$$

with C_j as in (7.1.13). The corresponding roots $\alpha \in \mathfrak{h}' \approx \mathbb{R}^2$, with $\mathfrak{h} \approx \{(a,b) : a, b \in \mathbb{R}\}$ via (7.1.16), are again given by (7.1.14).

The other 4 root spaces are 1-dimensional complex subspaces of

$$\left\{ E_{v,w} = \begin{pmatrix} & & v \\ & & w \\ -v^t & -w^t & 0 \end{pmatrix} : v, w \in \mathbb{C}^2 \right\}. \tag{7.1.19}$$

A computation gives

$$[D_{a,b}, E_{v,w}] = \begin{pmatrix} & & aJv \\ & & bJw \\ av^tJ & bw^tJ & 0 \end{pmatrix}. \tag{7.1.20}$$

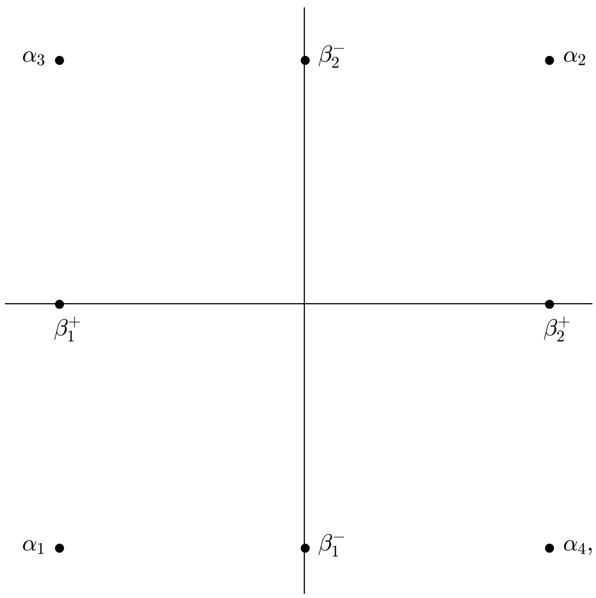

Figure 7.1.1. Roots of SO(5)

Referring to (7.1.12), we take

$$v_j = w_j = \begin{pmatrix} 1 \\ (-1)^{j-1}i \end{pmatrix}, \tag{7.1.21}$$

to get

$$[D_{a,b}, E_{v_j,0}] = i\beta_j^+(a,b)E_{v_j,0} \tag{7.1.22}$$

with

$$\beta_j^+(a,b) = (-1)^j a, \tag{7.1.23}$$

and

$$[D_{a,b}, E_{0,w_j}] = i\beta_j^-(a,b)E_{0,w_j}, \tag{7.1.24}$$

with

$$\beta_j^-(a,b) = (-1)^j b. \tag{7.1.25}$$

In summary, the 8 roots of $\mathfrak{so}(5)$ are α_j $(1 \leq j \leq 4)$, given by (7.1.14), and $\beta_j^\pm$ $(1 \leq j \leq 2)$, given by (7.1.23) and (7.1.24). These roots, expanded with respect to the basis dual to $\{D_{1,0}, D_{0,1}\}$ of $\mathfrak{h}$, from (7.1.16), are depicted in Figure 7.1.1.

Take a look at the (image under $\overline{\mathcal{W}}$ of) the elements of the Weyl group, acting on $\mathfrak{h}'$. In particular, the set of reflections $\{S_\alpha : \alpha \in \Delta\}$, given by (6.3.24)–(6.3.25), generate a group that coincides with the symmetry group of the square, D_4. Since each element $\mathcal{W}(g),\ g \in N(\mathbb{T})$, must permute the roots and act as an orthogonal transformation on $\mathfrak{h}'$, we see that $W(G)$ is generated by $\{S_\alpha : \alpha \in \Delta\}$ in this case, illustrating Proposition 6.3.4. We also have

$$W(\mathrm{SO}(5)) \approx D_4. \tag{7.1.26}$$

Another readily verifiable result is that the unique non-raisable weight vector (up to scaling) is the element C_2 of $\mathfrak{g}_{\alpha_2}$. Hence the adjoint representation of SO(5) on $\mathfrak{so}(5)_{\mathbb{C}}$ is irreducible. Equivalently,

$$\mathfrak{so}(5)_{\mathbb{C}} \ \text{ is simple,} \tag{7.1.27}$$

as a complex Lie algebra.

Having looked at the roots of $\mathfrak{so}(5)$, i.e., the weights for the adjoint representation, we next turn to the standard representation of SO(5) on $\mathbb{C}^5$, and decompose $\mathbb{C}^5 = \oplus V_\lambda$, where

$$V_\lambda = \{v \in \mathbb{C}^5 : Hv = i\lambda(H)v,\ \forall\, H \in \mathfrak{h}\}. \tag{7.1.28}$$

Here H takes the form $D_{a,b}$ of (7.1.16). We have

$$D_{a,b}\begin{pmatrix}1\\ \pm i\\ 0\\ 0\\ 0\end{pmatrix} = \mp ia\begin{pmatrix}1\\ \pm i\\ 0\\ 0\\ 0\end{pmatrix}, \quad \text{weights } \ \lambda_1^{\mp}(a,b) = \mp a, \tag{7.1.29}$$

$$D_{a,b}\begin{pmatrix}0\\ 0\\ 1\\ \pm i\\ 0\end{pmatrix} = \mp ib\begin{pmatrix}0\\ 0\\ 1\\ \pm i\\ 0\end{pmatrix}, \quad \text{weights } \ \lambda_2^{\mp}(a,b) = \mp b, \tag{7.1.30}$$

and

$$D_{a,b}\begin{pmatrix}0\\ 0\\ 0\\ 0\\ 1\end{pmatrix} = 0, \quad \text{weight } \ \lambda_3(a,b) = 0. \tag{7.1.31}$$

In summary:

$$\begin{aligned} &\lambda_1^+ = a = \beta_2^+, \quad \lambda_1^- = -a = \beta_1^+ \\ &\lambda_2^+ = b = \beta_2^-, \quad \lambda_2^- = -b = \beta_1^-, \quad \lambda_3 = 0. \end{aligned} \tag{7.1.32}$$

The highest weight is $\lambda_1^+ = \beta_2^+$, and this can be seen to be the only non-raisable weight. Hence the standard representation of SO(5) on $\mathbb{C}^5$ is irreducible. This is a special case of the fact that the standard representation of SO(n) on $\mathbb{C}^n$ is irreducible whenever $n \geq 3$, which will be proven in §7.2.

At this point we have identified α_2 and β_2^+ as highest weights of irreducible unitary representations of SO(5). Of course, $0 \in \mathfrak{h}'$ is the highest weight of the trivial representation. We can produce other elements of $\mathfrak{h}'$ known to be highest weights of irreducible representations of SO(5), using the observation that

$$\begin{aligned} &\mu_1,\ \mu_2 \text{ highest weights for irreducible representations of } G \\ &\Longrightarrow \mu_1 + \mu_2 \text{ highest weight for an irreducible representation.} \end{aligned} \tag{7.1.33}$$

Cf. Proposition 6.2.2. See Figure 7.1.2 for a depiction of the elements so produced, depicted by two black dots (representing α_2 and β_2^+) and a collection of circles (representing the other non-negative integral combinations of α_2 and β_2^+). These elements of $\mathfrak{h}'$ are not all the dominant integral weights, as specified in the Theorem of the Highest Weight in §6.2. It is readily checked that $\alpha_2/2$ is also dominant integral, and the collection of all dominant integral weights is the set of non-negative integral combinations of $\alpha_2/2$ and β_2^+. The additional dominant integral weights are depicted as diamonds in Figure 7.1.2.

In this context, the group to which the Theorem of the Highest Weight applies is not SO(5), but its simply connected double cover. In general, for $n \geq 3$, SO(n) has a simply connected double cover

$$\mathrm{Spin}(n) \longrightarrow \mathrm{SO}(n). \tag{7.1.34}$$

This group will be constructed in §7.4, following prerequisite material on Clifford algebras presented in §7.3. Section 7.5 will present spinor representations of Spin(n). For $n = 5$, there will be such a representation with highest weight $\alpha_2/2$.

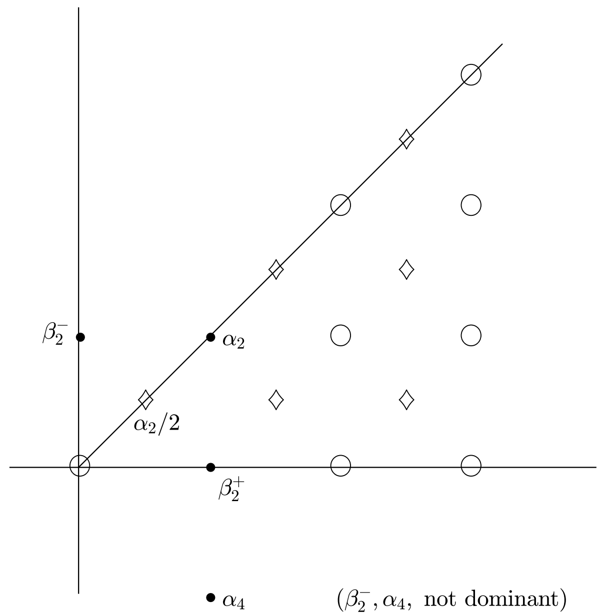

Figure 7.1.2. Dominant integral weights for SO(5)

Exercises

1. Set

$$\mathfrak{g} = \mathfrak{so}(5) = \{X \in M(5,\mathbb{R}) : X^* = -X\},$$

$$\mathfrak{k} = \mathfrak{so}(4) = \left\{ \begin{pmatrix} Y & \\ & 0 \end{pmatrix} : Y \in M(4,\mathbb{R}), Y^* = -Y \right\},$$

$$\mathfrak{p} = \left\{ \begin{pmatrix} & v \\ -v^t & 0 \end{pmatrix} : v \in \mathbb{R}^4 \right\}.$$

Show that

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}, \quad [\mathfrak{k},\mathfrak{k}] \subset \mathfrak{k}. \quad [\mathfrak{k},\mathfrak{p}] \subset \mathfrak{p}, \quad [\mathfrak{p},\mathfrak{p}] \subset \mathfrak{k}.$$

For the later, compute

$$\left[\begin{pmatrix} & v \\ -v^t & 0 \end{pmatrix}, \begin{pmatrix} & w \\ -w^t & 0 \end{pmatrix} \right] = \begin{pmatrix} wv^t - vw^t & \\ & 0 \end{pmatrix}.$$

2. Denote the groups generated by $\mathfrak{g}$ and $\mathfrak{k}$ by $G = SO(5)$ and $K \approx SO(4)$. Given $\begin{pmatrix} B & \\ & 1 \end{pmatrix} = A \in K$ and $Z = \begin{pmatrix} & v \\ -v^t & \end{pmatrix} \in \mathfrak{p}$, compute

$$AZA^{-1}, \quad \text{representing } K \text{ on } \mathfrak{p}.$$

3. Take

$$Z = \begin{pmatrix} & v \\ -v^t & 0 \end{pmatrix}, \quad v \in \mathbb{R}^4,\ |v| = 1.$$

Show that

$$Z^2 = -\begin{pmatrix} P_v & \\ & 1 \end{pmatrix} = -Q_v, \quad Z = Q_v Z,$$

where $Q_v = vv^t$ is the orthogonal projection of $\mathbb{R}^4$ onto Span(v). Deduce that

$$e^{tZ} = Q_v^\perp + \big[(\cos t)I + (\sin t)Z\big]Q_v.$$

7.2. Representations of SO(n), general n

We present some results on irreducible representations of SO(n) valid for general n. To start, note that SO(n) (and more generally Gl(n)) acts on $\Lambda^\ell \mathbb{R}^n$ for each $\ell \in \{0, 1, \dots, n\}$, via

$$\Lambda^\ell(g)\, v_1 \wedge \cdots \wedge v_\ell = gv_1 \wedge \cdots \wedge gv_\ell. \tag{7.2.1}$$

This extends by complexification to $\Lambda^\ell(g) : \Lambda^\ell \mathbb{C}^n \to \Lambda^\ell \mathbb{C}^n$.

Proposition 7.2.1. *The representation Λ^ℓ of* SO(n) *on $\Lambda^\ell \mathbb{C}^n$ is irreducible for each $\ell \in \{0, \dots, n\}$, except when n is even and $\ell = n/2$.*

Proof. To start, we assume $\ell < n/2$. Let $\{e_j : 1 \le j \le n\}$ be the standard basis of $\mathbb{R}^n$, hence of $\mathbb{C}^n$. Assume $V \subset \Lambda^\ell \mathbb{C}^n$ is a complex linear subspace, invariant under the SO(n) action, and assume $V \neq 0$. Pick a nonzero $\varphi \in V$, and write

$$\varphi = \sum a_{i_1 \cdots i_\ell}\, e_{i_1} \wedge \cdots \wedge e_{i_\ell}, \tag{7.2.2}$$

the sum taken over ℓ-tuples satisfying $1 \le i_1 < \cdots < i_\ell \le n$.

Suppose there is only one nonzero term, so we can assume

$$\varphi = e_{i_1} \wedge \cdots \wedge e_{i_\ell}. \tag{7.2.3}$$

Then as g runs over elements $E^\theta_\sigma \in \mathrm{SO}(n)$ given by

$$E^\theta_\sigma e_j = \theta_j e_{\sigma(j)}, \quad \sigma \in S_n,\ \theta_j = \pm 1, \tag{7.2.4}$$

such that $\det E^\theta_\sigma = \theta_1 \cdots \theta_n (\operatorname{sgn} \sigma) = 1$, we have $\Lambda^\ell(g)(e_{i_1} \wedge \cdots \wedge e_{i_\ell})$ running over $\pm e_{j_1} \wedge \cdots \wedge e_{j_n}$ for all multiindices satisfying $1 \le j_1 < \cdots < j_\ell \le n$, hence $V = \Lambda^\ell \mathbb{C}^n$.

Next suppose φ as in (7.2.2) belongs to V and at least 2 of the coefficients are nonzero, say $a_{i_1 \cdots i_\ell} \neq 0$ and $a_{j_1 \cdots j_\ell} \neq 0$, with $(i_1, \dots, i_\ell) \neq (j_1, \dots, j_\ell)$. As long as $\ell < n/2$, there exist $a, b \in \{1, \dots, n\}$ such that

$$a \in \{i_1, \dots, i_\ell\}, \quad a \notin \{j_1, \dots, j_\ell\}, \quad b \notin \{i_1, \dots, i_\ell\} \cup \{j_1, \dots, j_\ell\}. \tag{7.2.5}$$

Choose $g \in \mathrm{SO}(n)$ so that

$$ge_a = -e_a, \quad ge_b = -e_b, \quad ge_j = e_j \ \text{ otherwise}. \tag{7.2.6}$$

Then

$$\psi = \varphi + \Lambda^\ell(g)\varphi \tag{7.2.7}$$

has fewer nonzero coefficients than φ, but it has at least one. An induction finishes the irreducibility proof for $\ell < n/2$. □

To take care of the case $n/2 < \ell \le n$, we have the following.

Proposition 7.2.2. *For $0 \le \ell \le n$, the representations Λ^ℓ of* SO(n) *on $\Lambda^\ell \mathbb{C}^n$ and $\Lambda^{n-\ell}$ of* SO(n) *on $\Lambda^{n-\ell} \mathbb{C}^n$ are equivalent.*

Proof. We bring in the Hodge star operator

$$* : \Lambda^\ell \mathbb{R}^n \longrightarrow \Lambda^{n-\ell}\mathbb{R}^n, \tag{7.2.8}$$

defined for $\psi \in \Lambda^\ell \mathbb{R}^n$ by

$$\varphi \wedge *\psi = \langle \varphi, \psi \rangle\, \omega, \quad \forall\, \varphi \in \Lambda^\ell \mathbb{R}^n, \tag{7.2.9}$$

where $\omega \in \Lambda^n \mathbb{R}^n$ is the "volume element" $e_1 \wedge \cdots \wedge e_n$ and $\langle\ ,\ \rangle$ is the natural inner product on $\Lambda^\ell \mathbb{R}^n$ specified as follows. An inner product on a real vector space V induces an isomorphism $V \to V'$, which gives an isomorphism $\Lambda^\ell V \to \Lambda^\ell V' \approx (\Lambda^\ell V)'$, hence an inner product on $\Lambda^\ell V$. In the case $V = \mathbb{R}^n$ with standard orthonormal basis $\{e_j : 1 \le j \le n\}$, the set $\{e_{i_1} \wedge \cdots \wedge e_{i_\ell} : 1 \le i_1 < \cdots < i_\ell \le n\}$ is an orthonormal basis for $\Lambda^\ell \mathbb{R}^n$. With such a specification, we have

$$* \circ \Lambda^\ell(g) = \Lambda^{n-\ell}(g) \circ * \tag{7.2.10}$$

whenever $g \in \mathrm{Gl}(n, \mathbb{R})$ preserves the inner product and ω, i.e., whenever $g \in \mathrm{SO}(n)$. Having this, we extend

$$* : \Lambda^\ell \mathbb{C}^n \longrightarrow \Lambda^{n-\ell}\mathbb{C}^n \tag{7.2.11}$$

by $\mathbb{C}$-linearity, and (7.2.10) continues to hold. To complete the proof of Proposition 7.2.2, we note that $*$ in (7.2.11) is an isomorphism. In fact, a calculation gives

$$*e_{i_1} \wedge \cdots \wedge e_{i_\ell} = (\operatorname{sgn} \pi) e_{j_1} \wedge \cdots \wedge e_{j_{n-\ell}}, \tag{7.2.12}$$

where $\{i_1, \dots, i_\ell, j_1, \dots, j_{n-\ell}\} = \{1, \dots, n\}$, and the permutation π puts this set of indices in standard order. It follows that

$$** = (-1)^{\ell(n-\ell)} \ \text{ on } \ \Lambda^\ell \mathbb{C}^n. \tag{7.2.13}$$

$\square$

Proposition 7.2.2 finishes the proof of all the statements in Proposition 7.2.1 about the action of SO(n) on $\Lambda^\ell \mathbb{C}^n$ for $\ell \ne n/2$. In case $n = 2k$ and $\ell = k$, we have the SO($2k$) action commuting with $* : \Lambda^k \mathbb{C}^{2k} \to \Lambda^k \mathbb{C}^{2k}$. Note that if $1 \le i_1 < \cdots < i_k \le 2k$,

$$*e_{i_1} \wedge \cdots e_{i_k} = \pm e_{j_1} \wedge \cdots \wedge e_{j_k}, \quad \{i_1, \dots, i_k\} \cup \{j_1, \dots, j_k\} = \{1, \dots, 2k\}, \tag{7.2.14}$$

so $*$ is not a multiple of the identity. According to (7.2.13),

$$*^2 = (-1)^{k^2} = (-1)^k \ \text{ on } \ \Lambda^k \mathbb{C}^{2k}. \tag{7.2.15}$$

Hence

$$\begin{aligned} k \text{ even} &\Longrightarrow \operatorname{Spec} * = \{\pm 1\} \ \text{ on } \ \Lambda^k \mathbb{C}^{2k}, \\ k \text{ odd} &\Longrightarrow \operatorname{Spec} * = \{\pm i\} \ \text{ on } \ \Lambda^k \mathbb{C}^{2k}. \end{aligned} \tag{7.2.16}$$

(As an aside, the definitions imply that $*$ is orthogonal on $\Lambda^\ell\mathbb{R}^n$ and hence unitary on $\Lambda^\ell\mathbb{C}^n$. Consequently, by (7.2.15), $*$ is self-adjoint on $\Lambda^k\mathbb{C}^{2k}$ for k even and skew-adjoint on $\Lambda^k\mathbb{C}^{2k}$ for k odd.) We see that $\Lambda^k\mathbb{C}^{2k}$ breaks up into two pieces under the SO($2k$) action:

$$\begin{aligned}\Lambda^k\mathbb{C}^{2k} &= \Lambda^k_+\mathbb{C}^{2k} \oplus \Lambda^k_-\mathbb{C}^{2k},\\ \Lambda^k_\pm\mathbb{C}^{2k} &= \pm 1 \ \text{ eigenspace of } * \text{ for } k \text{ even}\\ &\quad\ \pm i \ \text{ eigenspace of } * \text{ for } k \text{ odd.}\end{aligned} \tag{7.2.17}$$

We also have

(7.2.18)

$$\begin{aligned} g\in \mathrm{O}(n),\ \det g = -1 &\Longrightarrow * \circ \Lambda^\ell(g) = -\Lambda^{n-\ell}(g)\circ *, \quad \forall\, \ell\in\{1,\dots,n\}\\ &\Longrightarrow \Lambda^k(g) : \Lambda^k_+\mathbb{C}^{2k} \xrightarrow{\approx} \Lambda^k_-\mathbb{C}^{2k},\end{aligned}$$

the latter when $n = 2k$.

The following is one complement to Proposition 7.2.1.

Proposition 7.2.3. *The action Λ^k of* O($2k$) *on $\Lambda^k\mathbb{C}^{2k}$ is irreducible.*

Proof. This is a variation of the proof of Proposition 7.2.1. Say $V \subset \Lambda^k\mathbb{C}^{2k}$ is invariant under the O($2k$) action. Take nonzero $\varphi \in V$, represented as in (7.2.2). If φ has the form (7.2.3), the argument given before implies $V = \Lambda^k\mathbb{C}^{2k}$. If there are at least two nonzero coefficients in (7.2.2), say $a_{i_1\cdots i_k}$ and $a_{j_1\cdots j_k}$, in this situation we take $a \in \{1,\dots,2k\}$ such that $a \in \{i_1,\dots,i_k\}$ but $a \notin \{j_1,\dots,j_k\}$, and in place of (40.6) define $g \in \mathrm{O}(n)$ by

$$ge_a = -e_a, \quad ge_j = e_j \ \text{ otherwise.} \tag{7.2.19}$$

Then, as in (7.2.7),

$$\psi = \varphi + \Lambda^k(g)\varphi \tag{7.2.20}$$

has fewer non-vanishing coefficients than φ, but is does have at least one. As in Proposition 7.2.1, an induction finishes the proof of irreducibility. $\square$

Here is another complement to Proposition 7.2.1.

Proposition 7.2.4. *The representations $\Lambda^k_\pm$ of* SO($2k$) *on $\Lambda^k_\pm\mathbb{C}^{2k}$ are irreducible.*

Proof. Take the case $\Lambda^k_+\mathbb{C}^{2k}$. Suppose $V_+ \subset \Lambda^k_+\mathbb{C}^{2k}$ is nonzero and invariant under the SO($2k$) action. Take $g_0 \in \mathrm{O}(2k)$ with $\det g_0 = -1$, and set

$$V_- = \Lambda^k(g_0)V_+, \tag{7.2.21}$$

a subspace of $\Lambda^k_-\mathbb{C}^{2k}$, by (7.2.18). Consider

$$V = V_+ \oplus V_- \subset \Lambda^k\mathbb{C}^{2k}. \tag{7.2.22}$$

We have

$$g \in \mathrm{O}(2k) \Longrightarrow \Lambda^k(g) : V \to V, \tag{7.2.23}$$

and hence, by Proposition 7.2.2, $V = \Lambda^k \mathbb{C}^{2k}$. This forces $V_+ = \Lambda_+^k \mathbb{C}^{2k}$, and proves irreducibility of Λ_+^k. The treatment of Λ_-^k is similar. □

We next consider the weights and weight spaces for the representations Λ^ℓ. We take the following maximal torus in SO(n). Assume $n = 2k$ or $n = 2k + 1$. For $1 \le j \le k$, define $R_j(\theta)$ by

$$\begin{aligned} R_j(\theta)e_{2j-1} &= (\cos\theta)e_{2j-1} + (\sin\theta)e_{2j} \\ R_j(\theta)e_{2j} &= -(\sin\theta)e_{2j-1} + (\cos\theta)e_{2j} \\ R_j(\theta)e_i &= e_i \quad \text{for} \quad i \notin \{2j-1, 2j\}. \end{aligned} \tag{7.2.24}$$

Then we take

$$\mathbb{T} = \{R_1(\theta_1)\cdots R_k(\theta_k) : \theta_j \in \mathbb{R}/2\pi\mathbb{Z}\}. \tag{7.2.25}$$

The Lie algebra $\mathfrak{h}$ of $\mathbb{T}$ is spanned by $\{J_{2j-1,2j} : 1 \le j \le k\}$, where, for $1 \le i < j \le k$,

$$J_{ij}e_i = e_j, \quad J_{ij}e_j = -e_i, \quad J_{ij}e_m = 0 \quad \text{for} \quad m \notin \{i, j\}. \tag{7.2.26}$$

Let us also set

$$E_j = J_{2j-1,2j}. \tag{7.2.27}$$

We prepare to calculate $d\Lambda^\ell(E_j)$ on $\Lambda^\ell\mathbb{C}^n$. We assume $0 \le \ell < n/2$, $n = 2k$ or $2k + 1$. It is convenient to pass from the standard basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$ (hence of $\mathbb{C}^n$) to the orthonormal basis $\{u_1, \dots, u_n\}$ of $\mathbb{C}^n$, given by

$$\begin{aligned} u_{2j-1} &= \frac{1}{\sqrt{2}}(e_{2j-1} - ie_{2j}), \\ u_{2j} &= \frac{1}{\sqrt{2}}(e_{2j-1} + ie_{2j}), \quad 1 \le j \le k, \\ u_n &= e_n, \quad \text{if} \quad n = 2k + 1. \end{aligned} \tag{7.2.28}$$

We have

$$\begin{aligned} E_j u_{2j-1} &= iu_{2j-1}, \\ E_j u_{2j} &= -iu_{2j}, \\ E_j u_i &= 0, \quad \text{if} \quad i \notin \{2j-1, 2j\}. \end{aligned} \tag{7.2.29}$$

Since

$$\begin{aligned} d\Lambda^\ell(E_j)u_{i_1} \wedge \cdots \wedge u_{i_\ell} = E_j u_{i_1} \wedge u_{i_2} \wedge \cdots \wedge u_{i_\ell} + \cdots \\ + u_{i_1} \wedge \cdots \wedge u_{i_{\ell-1}} \wedge E_j u_{i_\ell}, \end{aligned} \tag{7.2.30}$$

we get

(7.2.31)

$$
\begin{aligned}
&d\Lambda^\ell(E_j)u_{i_1} \wedge \cdots \wedge u_{i_\ell} \\
&= \begin{array}{rl} i\, u_{i_1} \wedge \cdots \wedge u_{i_\ell}, & \text{if } \ 2j-1 \in \{i_1, \dots, i_\ell\}, \text{ and } 2j \notin \{i_1, \dots, i_\ell\}, \\ -i\, u_{i_1} \wedge \cdots \wedge u_{i_\ell}, & \text{if } \ 2j \in \{i_1, \dots, i_\ell\}, \text{ and } 2j-1 \notin \{i_1, \dots, i_\ell\}, \\ 0 & \text{otherwise.} \end{array}
\end{aligned}
$$

Hence we have the following.

Proposition 7.2.5. *If $\ell < n/2$, then the monomials*

$$\{u_{i_1} \wedge \cdots \wedge u_{i_\ell} : 1 \le i_1 < \cdots < i_\ell \le n\}$$

form a basis of weight vectors, with weights determined by (7.2.31), *for the representation Λ^ℓ of* SO(n) *on $\Lambda^\ell \mathbb{C}^n$. In particular, this representation has highest weight given by the following k-tuple, if $n = 2k$ or $2k+1$:*

$$(1, \dots, 1, 0, \dots, 0) \quad (\ell \textit{ ones}), \tag{7.2.32}$$

with highest weight vector

$$u_1 \wedge u_3 \wedge \cdots \wedge u_{2\ell-1}. \tag{7.2.33}$$

Recall Proposition 7.2.2, which then takes care of the cases $n/2 < \ell \le n$. Finally, consider the case $n = 2k$, $\ell = k$. The calculations (7.2.28)–(7.2.31) still apply. We have weight vectors

$$
\begin{aligned}
&\varphi_0 = u_1 \wedge u_3 \wedge \cdots \wedge u_{2k-1}, \quad \text{weight} \ \ (1, \dots, 1, 1) \ \ (k\text{-tuple}), \\
&\varphi_1 = u_1 \wedge u_3 \wedge \cdots \wedge u_{2k-3} \wedge u_{2k}, \quad \text{weight} \ \ (1, \dots, 1, -1).
\end{aligned} \tag{7.2.34}
$$

These are the two highest weights for the representation Λ^k of SO($2k$) on $\Lambda^k \mathbb{C}^{2k}$. It follows that φ_0 is the highest weight vector for the representation Λ^k_σ on $\Lambda^k_\sigma \mathbb{C}^{2k}$, for some choice of sign $\sigma = \pm$ (the reader can have fun figuring out which choice). By (7.2.18), if $\varphi_0 \in \Lambda^k_\sigma \mathbb{C}^{2k}$, then $\varphi_1 \in \Lambda^k_{-\sigma} \mathbb{C}^{2k}$, as one can see by taking $g \in \mathrm{O}(2k)$ to switch e_{2k-1} and e_{2k} and fix the other e_j. Thus φ_1 must be the highest weight vector for the representation $\Lambda^k_{-\sigma}$ of SO($2k$) on $\Lambda^k_{-\sigma} \mathbb{C}^{2k}$. We summarize:

Proposition 7.2.6. *The representations $\Lambda^k_\pm$ of* SO($2k$) *on $\Lambda^k_\pm \mathbb{C}^{2k}$ have highest weights given by k-tuples*

$$(1, \dots, 1, 1) \quad \textit{and} \quad (1, \dots, 1, -1), \tag{7.2.35}$$

in some order.

We next take a second look at $\Lambda^2 \mathbb{C}^n$. This has the following significance. There is an isomorphism

$$
\begin{aligned}
A &: \Lambda^2 \mathbb{R}^n \longrightarrow \mathrm{Skew}(n) = \mathfrak{so}(n) \\
A &: \Lambda^2 \mathbb{C}^n \longrightarrow \mathfrak{so}_{\mathbb{C}}(n),
\end{aligned} \tag{7.2.36}
$$

defined by

$$A(u \wedge v)x = \langle u, x \rangle v - \langle v, x \rangle u, \tag{7.2.37}$$

for $u, v, x \in \mathbb{R}^n$, and extended by $\mathbb{C}$-linearity. Note that if $u, v, x, y \in \mathbb{R}^n$, then $\langle A(u \wedge v)x, y \rangle = \langle u, x \rangle \langle v, y \rangle - \langle v, x \rangle \langle u, y \rangle$, giving the asserted skew-symmetry. Now, given $g \in \mathrm{SO}(n)$ (or, more generally, $g \in \mathrm{O}(n)$),

$$\begin{aligned} gA(u \wedge v)g^{-1}x &= \langle u, g^{-1}x \rangle gv - \langle v, g^{-1}x \rangle gu \\ &= \langle gu, x, \rangle gv - \langle gv, x \rangle gu \\ &= A(gu \wedge gv)x. \end{aligned} \tag{7.2.38}$$

In other words, A intertwines the representations Λ^2 and Ad. We record the consequence:

Proposition 7.2.7. *The representation Λ^2 of* SO(n) *on $\Lambda^2\mathbb{C}^n$ is unitarily equivalent to the adjoint representation of* SO(n) *on $\mathfrak{so}_{\mathbb{C}}(n)$.*

Thus the study of Λ^2 has the potential to reveal information about the structure of the Lie algebra $\mathfrak{so}(n)$. In particular, the nonzero weights of Λ^2 are the *roots* of $\mathfrak{so}(n)$.

In our second look at Λ^2, we relabel the basis (7.2.28) of $\mathbb{C}^n$ as follows. For convenience, assume $n \geq 4$. Set

$$v_{j\varepsilon} = \frac{1}{\sqrt{2}}(e_{2j-1} - i\varepsilon e_{2j}), \quad 1 \leq j \leq k, \quad \varepsilon = \pm 1. \tag{7.2.39}$$

Then $\{v_{j\varepsilon} : 1 \leq j \leq k,\ \varepsilon = \pm 1\}$ forms a basis of $\mathbb{C}^n$ if $n = 2k$. If $n = 2k+1$, complete the basis by taking

$$v_n = e_n \quad \text{if} \quad n = 2k+1. \tag{7.2.40}$$

Parallel to (7.2.29), we have

$$\begin{aligned} E_j v_{i\varepsilon} &= i\varepsilon\delta_{ij} v_{i\varepsilon}, \\ E_j v_n &= 0 \quad \text{if} \quad n = 2k+1. \end{aligned} \tag{7.2.41}$$

In the current situation, a basis for $\Lambda^2\mathbb{C}^n$ is given by

$$\{v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2} : 1 \leq i_1 < i_2 \leq k,\ \varepsilon_1, \varepsilon_2 = \pm 1\} \cup \{v_{i,1} \wedge v_{i,-1} : 1 \leq i \leq k\}, \tag{7.2.42}$$

if $n = 2k$, and if $n = 2k+1$ we complete the basis by taking

$$\{v_{i\varepsilon} \wedge v_n : 1 \leq i \leq k,\ \varepsilon = \pm 1\}. \tag{7.2.43}$$

Now we calculate $d\Lambda^2(E_j)$ on this basis.

First, we have

$$d\Lambda^2(E_j)(v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2}) = i(\delta_{ji_1}\varepsilon_1 + \delta_{ji_2}\varepsilon_2) v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2}, \tag{7.2.44}$$

so each $v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2}$ is a weight vector, with weight

$$(7.2.45)\qquad \begin{array}{c} (0,\dots,\varepsilon_1,\dots,\varepsilon_2,\dots,0) \\ \quad\uparrow \qquad\ \uparrow \\ \quad i_1 \qquad\ i_2 \end{array} \qquad (k\text{-tuple}).$$

This weight is positive if and only if $\varepsilon_1 = 1$. Next,

$$(7.2.46)\qquad v_{i,1} \wedge v_{i,-1} = ie_{2i-1} \wedge e_{2i},$$

and

$$(7.2.47)\qquad d\Lambda^2(E_j)(v_{i,1} \wedge v_{i,-1}) = 0,$$

so each $v_{i,1} \wedge v_{i,-1}$ is a weight vector with weight

$$(7.2.48)\qquad (0,\dots,0) \qquad (k\text{-tuple}).$$

If $n = 2k$, the weights given by (7.2.45) and (7.2.48) are all the weights. If $n = 2k+1$, we also have

$$(7.2.49)\qquad d\Lambda^2(E_j)(v_{i\varepsilon} \wedge v_n) = i\delta_{ij}\varepsilon\, v_{i\varepsilon} \wedge v_n,$$

so $v_{i\varepsilon} \wedge v_n$ is a weight vector, with weight

$$(7.2.50)\qquad \begin{array}{c} (0,\dots,\varepsilon,\dots,0) \\ \uparrow \\ i \end{array} \qquad (k\text{-tuple}).$$

Taking Proposition 7.2.7 into account, we have:

Proposition 7.2.8. *The roots of $\mathfrak{so}(n)$ are given by* (7.2.45) *if $n = 2k$. If $n = 2k+1$, the roots are given by* (7.2.45) *and* (7.2.50). *The positive roots are given by*

$$(7.2.51)\qquad \begin{array}{c} (0,\dots,1,\dots,\varepsilon_2,\dots,0) \\ \quad\uparrow \qquad\ \uparrow \\ \quad i_1 \qquad\ i_2 \end{array} \qquad (k\text{-tuple}),$$

for $1 \le i_1 < i_2 \le k$, if $n = 2k$, and if $n = 2k+1$, also

$$(7.2.52)\qquad \begin{array}{c} (0,\dots,1,\dots,0) \\ \uparrow \\ i \end{array} \qquad (k\text{-tuple}),$$

for $1 \le i \le k$.

It is useful to record the image under $A : \Lambda^2\mathbb{C}^n \to \mathfrak{so}_{\mathbb{C}}(n)$ of the weight vectors given in (7.2.44), (7.2.47), and (7.2.49). The definition (7.2.37) readily yields

$$(7.2.53)\qquad A(e_i \wedge e_j) = J_{ij},$$

which is defined in (7.2.26). Hence $v_{i,1} \wedge v_{i,-1} = ie_{2i-1} \wedge e_{2i} \Rightarrow$

$$A(v_{i,1} \wedge v_{i,-1}) = iJ_{2i-1,2i}, \quad 1 \le i \le k, \tag{7.2.54}$$

a basis of $\mathfrak{h}_{\mathbb{C}}$, which is expected since the weights (7.2.48) are zero.

Next $v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2} = (1/2)(e_{2i_1-1} \wedge e_{2i_2-1} - \varepsilon_1\varepsilon_2 e_{2i_1} \wedge e_{2i_2} - i\varepsilon_1 e_{2i_1} \wedge e_{2i_2-1} - i\varepsilon_2 e_{2i_1-1} \wedge e_{2i_2}) \Rightarrow$

$$\begin{aligned} &A(v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2}) \\ &= \frac{1}{2}\Big(J_{2i_1-1,2i_2-1} - \varepsilon_1\varepsilon_2 J_{2i_1,2i_2} - i\varepsilon_1 J_{2i_1,2i_2-1} - i\varepsilon_2 J_{2i_1-1,2i_2}\Big), \\ &1 \le i_1 < i_2 \le k, \quad \varepsilon_1, \varepsilon_2 = \pm 1. \end{aligned} \tag{7.2.55}$$

These elements span root spaces with roots given by (7.2.45). In case $n = 2k = 4$, we have

$$\begin{aligned} A(v_{1\varepsilon_1} \wedge v_{2\varepsilon_2}) &= \frac{1}{2}\Big(J_{13} - \varepsilon_1\varepsilon_2 J_{24} - i\varepsilon_1 J_{23} - i\varepsilon_2 J_{14}\Big) \\ &= \frac{1}{2}\begin{pmatrix} & & -1 & i\varepsilon_2 \\ & & i\varepsilon_1 & \varepsilon_1\varepsilon_2 \\ 1 & -i\varepsilon_1 & & \\ -i\varepsilon_2 & -\varepsilon_1\varepsilon_2 & & \end{pmatrix}. \end{aligned} \tag{7.2.56}$$

Compare the root space calculation (7.1.8)–(7.1.13).

If $n = 2k$, the spaces spanned by elements of the form (7.2.55) give all the root spaces. If $n = 2k+1$, we also have the images under A of (7.2.49). Then $v_{1\varepsilon} \wedge v_n = (1/\sqrt{2})(e_{2i-1} \wedge e_n - i\varepsilon e_{2i} \wedge e_n) \Rightarrow$

$$A(v_{i\varepsilon} \wedge v_n) = \frac{1}{\sqrt{2}}\Big(J_{2i-1,n} - i\varepsilon J_{2i,n}\Big). \tag{7.2.57}$$

These elements span root spaces with roots given by (7.2.50). In case $n = 2k+1 = 5$, we have

$$\begin{aligned} A(v_{1\varepsilon} \wedge v_5) &= \frac{1}{\sqrt{2}}\Big(J_{15} - i\varepsilon J_{25}\Big) \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} & & & & -1 \\ & & & & i\varepsilon \\ & & & & 0 \\ & & & & 0 \\ 1 & -i\varepsilon & 0 & 0 & 0 \end{pmatrix}, \end{aligned} \tag{7.2.58}$$

with a similar result for $A(v_{2\varepsilon} \wedge v_5)$. Compare the root space calculations (7.1.19)–(7.1.21).

Recall from §6.2 the definition of a dominant integral weight, namely an element $\lambda \in \mathfrak{h}'$ such that

$$2\,\frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} \tag{7.2.59}$$

is a non-negative integer, for each positive root α. Note that $\langle\alpha, \alpha\rangle = 2$ for all roots of the form (7.2.51) and $\langle\alpha, \alpha\rangle = 1$ for all roots of the form (7.2.52). As a consequence of Proposition 7.2.8, we have the following.

Proposition 7.2.9. *The dominant integral weights for $\mathfrak{so}(2k)$ are given by k-tuples of the form $(d_1, \dots, d_k)$, satisfying*

$$d_1 \geq \cdots \geq d_{k-1} \geq |d_k|, \tag{7.2.60}$$

where either all the components d_j are integers or they are all (non-integral) half-integers. The dominant integral weights for $\mathfrak{so}(2k+1)$ are given by such k-tuples, satisfying

$$d_1 \geq \cdots \geq d_k \geq 0, \tag{7.2.61}$$

instead of (7.2.60).

Proof. For each positive root of the form (7.2.51), the condition that (7.2.59) belong to $\mathbb{Z}^+$ is that $\langle\lambda, \alpha\rangle \in \mathbb{Z}^+$, hence, if $\lambda = (d_1, \dots, d_k)$, that $d_j + \varepsilon d_\ell \in \mathbb{Z}^+$, for $1 \leq j < \ell \leq k$, $\varepsilon = \pm 1$, i.e.,

$$d_j + d_\ell,\ d_j - d_\ell \in \mathbb{Z}^+, \quad \text{for } 1 \leq j < \ell \leq k. \tag{7.2.62}$$

This requires $d_j \geq d_\ell$ for $1 \leq j < \ell \leq k$, and it also requires $2d_j \in \mathbb{Z}^+$ for $1 \leq j < k$. Hence d_1 is either an integer or a (non-integral) half integer, and then (7.2.62) requires the same property of each d_j, $1 \leq j \leq k$. With $j = k-1$, (7.2.62) requires $d_{k-1} \geq |d_k|$, and we have (7.2.60). This takes care of SO($2k$).

For SO($2k+1$), we need also consider the positive roots of the form (40.51). If α is such a root, membership of (7.2.59) in $\mathbb{Z}^+$ requires

$$2d_j \in \mathbb{Z}^+, \quad \text{for } 1 \leq j \leq k. \tag{7.2.63}$$

Hence we have (7.2.61). ☐

The dominant integral weights described above are non-negative integral combinations of the highest weight representations of SO(n) described above, provided d_j are all integers, as is seen upon recalling that the previously obtained highest weights are the k-tuples

$$(1, 0, \dots, 0),\ (1, 1, 0, \dots, 0), \dots, (1, \dots, 1, 1) \ \text{ and } \ (1, \dots, 1, -1) \tag{7.2.64}$$

when $n = 2k$, for $\Lambda^\ell\mathbb{C}^{2k}$, $1 \leq \ell \leq k-1$, and $\Lambda^k_\pm\mathbb{C}^{2k}$, and they are

$$(1, 0, \dots, 0),\ (1, 1, 0, \dots, 0), \dots, (1, \dots, 1, 1) \tag{7.2.65}$$

when $n = 2k+1$, for $\Lambda^\ell\mathbb{C}^{2k+1}$, $1 \leq \ell \leq k$.

The dominant integral weights involving half integral d_j are non-negative integral combinations of these plus the k-tuples

$$\Big(\frac{1}{2}, \dots, \frac{1}{2}, \frac{1}{2}\Big) \ \text{ and } \ \Big(\frac{1}{2}, \dots, \frac{1}{2}, -\frac{1}{2}\Big), \tag{7.2.66}$$

for $n = 2k$, and

$$\left(\frac{1}{2}, \dots, \frac{1}{2}, \frac{1}{2}\right), \tag{7.2.67}$$

for $n = 2k+1$. Constructions of representations of two-fold covers of SO(n) with these highest weights will be given in §§7.5–7.6.

We next specify the Weyl group $W(\mathrm{SO}(n))$ for each $n \geq 3$, or more precisely its image under $\overline{\mathcal{W}}$ in $\mathrm{Gl}(\mathfrak{h}')$ (defined by (6.3.6)–(6.3.7)), which we will denote $\overline{\mathcal{W}}(\mathrm{SO}(n))$. Recall that for $n = 2k$ the roots of $\mathfrak{so}(n)$ are given by (7.2.45); denote such roots as

$$\varepsilon_1 E'_{i_1} + \varepsilon_2 E'_{i_2} = \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}, \tag{7.2.68}$$

where $1 \leq i_1 < i_2 \leq k$, $\varepsilon_1, \varepsilon_2 \in \{\pm 1\}$ and $\{E'_1, \dots, E'_k\}$ is the basis of $\mathfrak{h}'$ dual to the basis $\{E_1, \dots, E_k\}$, specified by (7.2.27) (which is orthonormal with respect to an Ad-invariant inner product on $\mathfrak{so}(n)$). By Proposition 6.3.2, the following reflections belong to $\overline{\mathcal{W}}(\mathrm{SO}(2k))$:

$$\rho_{i_1\varepsilon_1 i_2\varepsilon_2}(\lambda) = \lambda - \langle \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}, \lambda\rangle \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}, \tag{7.2.69}$$

since $\langle \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}, \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}\rangle = 2$. Note that

$$\begin{aligned} \rho_{i_1\varepsilon_1 i_2\varepsilon_2} E'_{i_1} &= -\varepsilon_1\varepsilon_2 E'_{i_2}, \\ \rho_{i_1\varepsilon_1 i_2\varepsilon_2} E'_{i_2} &= -\varepsilon_1\varepsilon_2 E'_{i_1}, \\ \rho_{i_1\varepsilon_1 i_2\varepsilon_2} E'_{\ell} &= E'_{\ell}, \quad \ell \notin \{i_1, i_2\}. \end{aligned} \tag{7.2.70}$$

Noting that (7.2.70) is a function of $-\varepsilon_1\varepsilon_2$, we can relabel these reflections as

$$\{R_{i_1 i_2 \varepsilon} : 1 \leq i_1 < i_2 \leq k,\ \varepsilon = \pm 1\}, \tag{7.2.71}$$

given by

$$\begin{aligned} R_{i_1 i_2 \varepsilon} E'_{i_1} &= \varepsilon E'_{i_2}, \\ R_{i_1 i_2 \varepsilon} E'_{i_2} &= \varepsilon E'_{i_1}, \\ R_{i_1 i_2 \varepsilon} E'_{\ell} &= E'_{\ell}, \quad \ell \notin \{i_1, i_2\}. \end{aligned} \tag{7.2.72}$$

By Proposition 6.3.4, $\overline{\mathcal{W}}(\mathrm{SO}(2k))$ is the group generated by the set of reflections (7.2.71).

The roots of $\mathrm{SO}(2k+1)$ are given by (7.2.68) plus

$$\varepsilon E'_i, \quad 1 \leq i \leq k, \quad \varepsilon = \pm 1. \tag{7.2.73}$$

Thus, in addition to the reflections (7.2.71), $\overline{\mathcal{W}}(\mathrm{SO}(2k+1))$ contains the set of reflections

$$\{R_i : 1 \leq i \leq k\}, \tag{7.2.74}$$

given by

$$R_i E_i' = -E_i', \quad R_i E_\ell' = E_\ell', \quad \ell \neq i. \tag{7.2.75}$$

By Proposition 6.3.4, $\overline{\mathcal{W}}(\mathrm{SO}(2k+1))$ is the group generated by the set of reflections given in (7.2.71) and (7.2.74). It follows that

$$\overline{\mathcal{W}}(\mathrm{SO}(2k+1)) = \{E_\sigma^\theta : \sigma \in S_k,\ \theta = (\pm 1, \dots, \pm 1)\}, \tag{7.2.76}$$

where, as in (7.2.4), $E_\sigma^\theta \in \mathrm{End}(\mathfrak{h}')$ is defined by

$$E_\sigma^\theta E_j' = \theta_j E_{\sigma(j)}'. \tag{7.2.77}$$

Given this, we have

Proposition 7.2.10. *For $k \geq 1$, $\overline{\mathcal{W}}(\mathrm{SO}(2k+1))$ is the group of transformations of $\mathfrak{h}' \approx \mathbb{R}^k$ that are symmetries of the k-dimensional cube*

$$Q^k = \{x \in \mathbb{R}^k : -1 \leq x_i \leq 1,\ \textit{for } 1 \leq i \leq k\}. \tag{7.2.78}$$

Proof. Each transformation given by (7.2.77) clearly produces a symmetry of Q^k. Conversely, each symmetry S of Q^k is an orthogonal transformation of $\mathbb{R}^k$ that is uniquely specified by the image under S of the ordered basis $(E_1', \dots, E_k')$. This image is necessarily of the form

$$\theta_1 E_{\sigma(1)}', \dots, \theta_k E_{\sigma(k)}' \tag{7.2.79}$$

for some permutation σ of $\{1, \dots, k\}$ and some $\theta_j \in \{\pm 1\}$, so S is of the form (7.2.77). $\square$

Remark. Inspection of (7.2.72) shows that

$$\overline{\mathcal{W}}(\mathrm{SO}(2k)) = \{E_\sigma^\theta : \sigma \in S_k,\ \theta_1 \cdots \theta_k = 1\}. \tag{7.2.80}$$

Returning to the issue of highest weights, we recall the central role that the Lie group homomorphisms

$$\gamma^\alpha : \mathrm{SU}(2) \longrightarrow G, \tag{7.2.81}$$

defined for each root α of $\mathfrak{g}$, played in §6.2, in the statement of the Theorem of the Highest Weight. We record some results on γ^α when $G = \mathrm{SO}(n)$, starting with the case $n = 2k$.

Proposition 7.2.11. *For each root α of $\mathfrak{so}(2k)$,*

$$\gamma^\alpha : \mathrm{SU}(2) \longrightarrow \mathrm{SO}(2k) \quad \textit{is injective.} \tag{7.2.82}$$

Proof. As noted in §6.2, the kernel of γ^α is either $\{I\}$ or $\{\pm I\} \subset \mathrm{SU}(2)$. Now, by (6.2.4),

$$\gamma^\alpha(\operatorname{Exp} tX_1) = \operatorname{Exp}(tX_1^\alpha), \tag{7.2.83}$$

and, by (4.1.2), $\operatorname{Exp}(2\pi X_1) = -I \in \mathrm{SU}(2)$, so our task is to examine $\operatorname{Exp}(2\pi X_1^\alpha) \in \mathrm{SO}(n)$. Now, by (6.2.3) and (6.1.6), X_1^α is specified as the element of $\mathfrak{h}$ satisfying

$$\langle h, X_1^\alpha \rangle = \frac{\alpha(h)}{\langle \alpha, \alpha \rangle}, \quad \forall\, h \in \mathfrak{h}. \tag{7.2.84}$$

Given that α has the form (7.2.45) (so $\langle \alpha, \alpha \rangle = 2$), we deduce that

$$2X_1^\alpha = \varepsilon_1 E_{i_1} + \varepsilon_2 E_{i_2}, \tag{7.2.85}$$

with $i_1 \neq i_2,\ \varepsilon_j = \pm 1$, and $E_j = J_{2j-1,2j}$, as in (7.2.27). Hence

$$\operatorname{Exp}(2tX_1^\alpha) = e^{t\varepsilon_1 J_{2i_1-1,2i_1}}\, e^{t\varepsilon_2 J_{2i_2-1,2i_2}} \in \mathrm{SO}(n), \tag{7.2.86}$$

and in particular, for α of the form (7.2.45),

$$\begin{aligned} \gamma^\alpha(-I) = \operatorname{Exp}(2\pi X_1^\alpha) &= e^{\pi\varepsilon_1 J_{2i_1-1,2i_1}}\, e^{\pi\varepsilon_2 J_{2i_2-1,2i_2}} \\ &\neq I \quad \text{in} \quad \mathrm{SO}(n), \end{aligned} \tag{7.2.87}$$

since

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Longrightarrow e^{\pm \pi J} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7.2.88}$$

□

Here is the contrasting result for SO($2k+1$).

Proposition 7.2.12. *For each root* α *of* $\mathfrak{so}(2k+1)$ *of the form* (7.2.45),

$$\gamma^\alpha : \mathrm{SU}(2) \longrightarrow \mathrm{SO}(2k+1) \quad \textit{is injetive.} \tag{7.2.89}$$

However, for each root of $\mathfrak{so}(2k+1)$ *of the form* (7.2.50),

$$\gamma^\alpha(-I) = I \quad \textit{in} \quad \mathrm{SO}(2k+1), \tag{7.2.90}$$

hence $\gamma^\alpha(\mathrm{SU}(2))$ *is isomorphic to* SO(3).

Proof. If the root α has the form (7.2.45), the arguments proving Proposition 40.11 apply, to give (7.2.89). If α has the form (7.2.50), then $\langle \alpha, \alpha \rangle = 1$, and in place of (7.2.85) we have

$$X_1^\alpha = \varepsilon E_i, \tag{7.2.91}$$

again with $\varepsilon = \pm 1$ and $E_i = J_{2i-1,2i}$. Hence

$$\operatorname{Exp}(tX_1^\alpha) = e^{t\varepsilon J_{2i-1,2i}} \in \mathrm{SO}(2k+1), \tag{7.2.92}$$

so

$$\begin{aligned} \gamma^\alpha(-I) = \operatorname{Exp}(2\pi X_1^\alpha) &= e^{2\pi\varepsilon J_{2i-1,2i}} \\ &= I, \quad \text{in} \quad \mathrm{SO}(2k+1), \end{aligned} \tag{7.2.93}$$

as asserted in (7.2.90). □

Note that (7.2.90) implies that, for each root of $\mathfrak{so}(2k+1)$ of the form (7.2.50), and each representation π of SO($2k+1$), $\pi^\alpha = \pi \circ \gamma^\alpha$ is an SO(3) representation, so in place of (7.2.63), we have $d_j \in \mathbb{Z}^+$, for $1 \le j \le k$. This is consistent with our observation about the irreducible representations of SO(n) whose existence was proved above. Such an argument does not apply to representations of SO($2k$).

However, we can see directly that each d_j is an integer, for each irreducible representation of SO(n), as follows. If π is a representation of SO(n) on V and $v \in V_\lambda$ is a weight vector, associated to the weight $\lambda = (d_1, \dots, d_k)$, with $n = 2k$ or $2k+1$, then, for each $j \in \{1, \dots, k)$, $E_j = J_{2j-1,2j}$, as in (7.2.27),

$$d\pi(E_j)v = id_j v, \tag{7.2.94}$$

hence

$$\pi(e^{tJ_{2j-1,2j}})v = e^{itd_j}v. \tag{7.2.95}$$

Since, parallel to (7.2.93),

$$e^{2\pi J_{2j-1,2j}} = I \in \mathrm{SO}(n), \tag{7.2.96}$$

this requires $d_j \in \mathbb{Z}$. Hence we have the following.

Proposition 7.2.13. *If $n = 2k$ or $2k+1$, the highest weights of irreducible unitary representations of* SO(n) *are precisely the k-tuples $(d_1, \dots, d_k)$ satisfying* (7.2.60) *if $n = 2k$,* (7.2.61) *if $n = 2k+1$, and*

$$d_j \in \mathbb{Z}, \quad \forall\, j \in \{1, \dots, k\}. \tag{7.2.97}$$

Proof. The existence of such representations follows, via Proposition 6.2.2, from the results (7.2.64)–(7.2.65). The necessity of (7.2.97) has just been established. □

7.3. Clifford algebras

Let V be a finite dimensional, real vector space and $Q : V \times V \to \mathbb{R}$ a symmetric bilinear form. The Clifford algebra $\mathcal{C}\ell(V, Q)$ is an associative algebra, with unit 1, generated by V, and satisfying the anticommutation relations

$$uv + vu = -2Q(u, v) \cdot 1, \quad \forall\, u, v \in V. \tag{7.3.1}$$

Formally, we construct $\mathcal{C}\ell(V, Q)$ as

$$\mathcal{C}\ell(V, Q) = \otimes^* V/\mathcal{I}, \tag{7.3.2}$$

where $\otimes^* V$ is the tensor algebra:

$$\otimes^* V = \mathbb{R} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots, \tag{7.3.3}$$

and

$$\begin{aligned} \mathcal{I} &= \text{ two-sided ideal generated by } \{u \otimes v + v \otimes u + 2Q(u, v)1 : u, v \in V\} \\ &= \text{ two-sided ideal generated by } \{e_j \otimes e_k + e_k \otimes e_j + 2Q(e_j, e_k)1\}, \end{aligned} \tag{7.3.4}$$

where $\{e_j\}$ is a basis of V. Note that

$$Q = 0 \Longrightarrow \mathcal{C}\ell(V, Q) \approx \Lambda^* V \quad \text{(the exterior algebra)}. \tag{7.3.5}$$

Here is a fundamental property of $\mathcal{C}\ell(V, Q)$.

Proposition 7.3.1. *Let $\mathcal{A}$ be an associative algebra with unit, and let*

$$M : V \longrightarrow \mathcal{A} \tag{7.3.6}$$

be a linear map satisfying

$$M(u)M(v) + M(v)M(u) = -2Q(u, v)1, \tag{7.3.7}$$

for each $u, v \in V$ (or equivalently for all $u = e_j$, $v = e_k$, where $\{e_j\}$ is a basis of V). Then M extends to a homomorphism

$$M : \mathcal{C}\ell(V, Q) \longrightarrow \mathcal{A}, \quad M(1) = 1. \tag{7.3.8}$$

Proof. Given (7.3.6), there is a homomorphism $\widetilde{M} : \otimes^* V \to \mathcal{A}$ extending M, with $\widetilde{M}(1) = 1$. The relation (7.3.7) implies $\widetilde{M} = 0$ on $\mathcal{I}$, so it descends to $\otimes^* V/\mathcal{I} \to \mathcal{A}$, giving (7.3.8). □

From here on we require Q to be nondegenerate. Thus each Clifford algebra $\mathcal{C}\ell(V, Q)$ we consider will be isomorphic to one of the following. Take $V = \mathbb{R}^n$, with standard basis $\{e_1, \dots, e_n\}$, take $p, q \geq 0$ such that $p + q = n$, and take $Q(u, v) = \sum_{j \leq p} u_j v_j - \sum_{j > p} u_j v_j$, where $u = \sum u_j e_j$ and $v = \sum v_j e_j$. In such a case, $\mathcal{C}\ell(V, Q)$ is denoted $\mathcal{C}\ell(p, q)$.

We also define the complexification of $\mathcal{C}\ell(V,Q)$:

$$\mathbb{C}\ell(V,Q)=\mathbb{C}\otimes\mathcal{C}\ell(V,Q). \tag{7.3.9}$$

(We tensor over $\mathbb{R}$.) Note that taking $e_j\mapsto ie_j$ for $p+1\le j\le n$ gives, whenever $p+q=n$,

$$\mathbb{C}\ell(p,q)\approx\mathbb{C}\ell(n,0),\quad \text{which we denote } \ \mathbb{C}\ell(n). \tag{7.3.10}$$

Use of the anticommutator relations (7.3.1) show that if $\{e_1,\dots,e_n\}$ is a basis of V, then each element $u\in\mathcal{C}\ell(V,Q)$ can be written in the form

$$u=\sum_{i_\nu=0\text{ or }1}a_{i_1\cdots i_n}e_1^{i_1}\cdots e_n^{i_n}, \tag{7.3.11}$$

or, equivalently, in the form

$$u=\sum_{k=0}^{n}\sum_{j_1<\cdots<j_k}\tilde{a}_{j_1\cdots j_k}e_{j_1}\cdots e_{j_k}. \tag{7.3.12}$$

(By convention the $k=0$ summand in (41.12) is $\tilde{a}_\emptyset\cdot 1$.) In other words, we see that

$$\{e_{j_1}\cdots e_{j_k}:0\le k\le n,\ j_1<\cdots<j_k\} \tag{7.3.13}$$

spans $\mathcal{C}\ell(V,Q)$. Again, by convention, the subset of (7.3.13) for which $k=0$ is $\{1\}$. It is very useful to know that the following is true.

Proposition 7.3.2. *The set* (7.3.13) *is a basis of* $\mathcal{C}\ell(V,Q)$.

This is true for all Q, but we will restrict attention to nondegenerate Q. Since we know that (7.3.13) spans, the assertion is that the dimension of $\mathcal{C}\ell(p,q)$ is 2^n when $p+q=n$. By (7.3.10), it suffices to show this for $\mathcal{C}\ell(n,0)$, and we can assume $\{e_1,\dots,e_n\}$ is the standard orthonormal basis of $\mathbb{R}^n$ Note that the assertion for $Q=0$ corresponding to Proposition 7.3.2 is that

$$\{e_{j_1}\wedge\cdots\wedge e_{j_k}:0\le k\le n,j_1<\cdots<j_k\}\ \text{ is a basis of }\ \Lambda^*\mathbb{R}^n, \tag{7.3.14}$$

where $\{e_1,\dots,e_n\}$ is the standard basis of $\mathbb{R}^n$. We will use this in our proof of Proposition 7.3.2. See §B.5 for a proof of (7.3.14).

Given that (7.3.14) is true, we can define a linear map

$$\alpha:\Lambda^*\mathbb{R}^n\longrightarrow\mathcal{C}\ell(n,0) \tag{7.3.15}$$

by $\alpha(1)=1$ and

$$\alpha(e_{j_1}\wedge\cdots\wedge e_{j_k})=e_{j_1}\cdots e_{j_k}, \tag{7.3.16}$$

when $1\le j_1<\cdots<j_k\le n$. The content of Proposition 7.3.2 is that α is a linear isomorphism. On the way to proving this, we construct a representation of $\mathcal{C}\ell(n,0)$ on $\Lambda^*\mathbb{R}^n$, of interest in its own right.

To construct this representation, i.e., homomorphism of algebras

$$M : \mathcal{C}\ell(n,0) \longrightarrow \operatorname{End}(\Lambda^*\mathbb{R}^n), \tag{7.3.17}$$

we begin with a linear map

$$M : \mathbb{R}^n \longrightarrow \operatorname{End}(\Lambda^*\mathbb{R}^n), \tag{7.3.18}$$

defined on the basis $\{e_1, \dots, e_n\}$ as follows. Define

$$\wedge_j : \Lambda^k\mathbb{R}^n \longrightarrow \Lambda^{k+1}\mathbb{R}^n, \quad \iota_j : \Lambda^k\mathbb{R}^n \longrightarrow \Lambda^{k-1}\mathbb{R}^n \tag{7.3.19}$$

by

$$\wedge_j(e_{j_1} \wedge \cdots \wedge e_{j_k}) = e_j \wedge e_{j_1} \wedge \cdots \wedge e_{j_k}, \tag{7.3.20}$$

and

$$\begin{aligned} \iota_j(e_{j_1} \wedge \cdots \wedge e_{j_k}) &= \\ (-1)^{\ell-1} e_{j_1} \wedge \cdots \wedge \widehat{e_{j_\ell}} \wedge \cdots \wedge e_{j_k} \quad &\text{if } \ j = j_\ell, \\ 0 \qquad\qquad &\text{if } \ j \notin \{j_1, \dots, j_k\}. \end{aligned} \tag{7.3.21}$$

Here the symbol $\widehat{e_{j_\ell}}$ signifies that e_{j_ℓ} is removed from the product.

Remark. If $\Lambda^*\mathbb{R}^n$ has the inner product such that (7.3.14) is an orthonormal basis, then ι_j is the adjoint of $\wedge_j$.

A calculation (left to the reader) gives the following anticommutator relations for these operators:

$$\begin{aligned} \wedge_j \wedge_k + \wedge_k \wedge_j &= 0, \\ \iota_j \iota_k + \iota_k \iota_j &= 0, \\ \wedge_j \iota_k + \iota_k \wedge_j &= \delta_{jk}. \end{aligned} \tag{7.3.22}$$

Now we define M in (7.3.18) by

$$M(e_j) = M_j = \wedge_j - \iota_j. \tag{7.3.23}$$

From (7.3.22) we get

$$M_j M_k + M_k M_j = -2\delta_{jk}. \tag{7.3.24}$$

Hence Proposition 7.3.1 applies to give the homomorphism of algebras (7.3.17), with $M(1) = I$, the identity operator.

We can now prove Proposition 7.3.2. We define a linear map

$$\beta : \mathcal{C}\ell(n,0) \longrightarrow \Lambda^*\mathbb{R}^n, \quad \beta(u) = M(u)1. \tag{7.3.25}$$

Recalling the map α from (7.3.15)–(7.3.16), we have

$$\begin{aligned} \beta \circ \alpha(e_{j_1} \wedge \cdots \wedge e_{j_k}) &= M(e_{j_1} \cdots e_{j_k})1 \\ &= M(e_{j_1}) \cdots M(e_{j_k})1. \end{aligned} \tag{7.3.26}$$

Now $M(e_{j_k})1 = e_{j_k},\ M(e_{j_{k-1}})e_{j_k} = e_{j_{k-1}} \wedge e_{j_k}$ if $j_{k-1} < j_k$, and inductively we see that

$$j_1 < \cdots < j_k \Longrightarrow M(e_{j_1}) \cdots M(e_{j_k})1 = e_{j_1} \wedge \cdots \wedge e_{j_k}. \tag{7.3.27}$$

It follows that α and β are inverses, and that each is a linear isomorphism. This proves Proposition 7.3.2 (granted (7.3.14)).

We next characterize $\mathcal{C}\ell(p,q)$ for small p and q. For starters, $\mathcal{C}\ell(1,0)$ and $\mathcal{C}\ell(0,1)$ are linear spaces of the form

$$\{a + be_1 : a, b \in \mathbb{R}\}. \tag{7.3.28}$$

In $\mathcal{C}\ell(1,0),\ e_1^2 = -1$, so

$$\mathcal{C}\ell(1,0) \approx \mathbb{C}, \quad e_1 \leftrightarrow i. \tag{7.3.29}$$

Meanwhile, in $\mathcal{C}\ell(0,1),\ e_1^2 = 1$, so $\mathcal{C}\ell(0,1)$ is of the form

$$\begin{gathered}\{\alpha f_+ + \beta f_- : \alpha, \beta \in \mathbb{R}\} \\ f_\pm = \frac{1 \pm e_1}{2} \Rightarrow f_\pm^2 = f_\pm,\ f_+ f_- = f_- f_+ = 0,\end{gathered} \tag{7.3.30}$$

and we have

$$\mathcal{C}\ell(0,1) \approx \mathbb{R} \oplus \mathbb{R} \approx C_{\mathbb{R}}(\{+,-\}), \tag{7.3.31}$$

the space of real valued functions on the two-point set $\{+,-\}$.

Next, $\mathcal{C}\ell(2,0), \mathcal{C}\ell(1,1)$, and $\mathcal{C}\ell(0,2)$ are linear spaces of the form

$$\{a + be_1 + ce_2 + de_1e_2 : a, b, c, d \in \mathbb{R}\}. \tag{7.3.32}$$

In $\mathcal{C}\ell(2,0),\ e_1^2 = e_2^2 = (e_1e_2)^2 = -1$, and also $e_2(e_1e_2) = e_1$, while $(e_1e_2)e_1 = e_2$, which are the algebraic relations satisfied by i, j, k in the algebra $\mathbb{H}$ of quaternions, defined by (1.2.1)–(1.2.3). Hence

$$\mathcal{C}\ell(2,0) \approx \mathbb{H} = \{a + bi + cj + dk\}. \tag{7.3.33}$$

In $\mathcal{C}\ell(0,2),\ e_1^2 = e_2^2 = 1$, while $(e_1e_2)^2 = -1$. Meanwhile $e_2(e_1e_2) = -e_1$ and $(e_1e_2)e_1 = -e_2$, and we have

$$\begin{aligned}&\mathcal{C}\ell(0,2) \approx \mathrm{M}(2,\mathbb{R}) \\ &\quad = \Bigl\{aI + b\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + c\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} + d\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} : a, b, c, d \in \mathbb{R}\Bigr\}.\end{aligned} \tag{7.3.34}$$

It turns out that also

$$\mathcal{C}\ell(1,1) \approx M(2,\mathbb{R}).$$

We leave this to the reader.

Using (7.3.31) and (7.3.34), we find the complexified algebras

$$\mathbb{C}\ell(1) \approx \mathbb{C} \oplus \mathbb{C}, \quad \mathbb{C}\ell(2) \approx \mathrm{M}(2,\mathbb{C}). \tag{7.3.35}$$

These results are special cases of the following:

Proposition 7.3.3. *The complex Clifford algebras* $\mathbb{C}\ell(n)$ *have the properties*

$$\begin{aligned}\mathbb{C}\ell(2k) &\approx \mathrm{M}(2^k,\mathbb{C}),\\ \mathbb{C}\ell(2k+1) &\approx \mathrm{M}(2^k,\mathbb{C})\oplus \mathrm{M}(2^k,\mathbb{C}).\end{aligned} \tag{7.3.36}$$

Proposition 7.3.3 follows inductively from (7.3.35) and the following result.

Proposition 7.3.4. *For* $n \in \mathbb{N}$, *we have isomorphisms of algebras*

$$\mathbb{C}\ell(n+2) \approx \mathbb{C}\ell(n)\otimes \mathbb{C}\ell(2). \tag{7.3.37}$$

In turn, Proposition 7.3.4 follows from:

Proposition 7.3.5. *For* $n \in \mathbb{N}$, *we have isomorphisms of algebras*

$$\mathcal{C}\ell(n,0)\otimes \mathcal{C}\ell(0,2) \approx \mathcal{C}\ell(0,n+2). \tag{7.3.38}$$

It remains to prove (7.3.38). To do this, we construct a homomorphism of algebras

$$M : \mathcal{C}\ell(0,n+2) \longrightarrow \mathcal{C}\ell(n,0)\otimes \mathcal{C}\ell(0,2). \tag{7.3.39}$$

Once it is checked that M is onto, a dimension count guarantees it is an isomorphism.

To produce (7.3.39), we start with a linear map

$$M : \mathbb{R}^{n+2} \longrightarrow \mathcal{C}\ell(n,0)\otimes \mathcal{C}\ell(0,2), \tag{7.3.40}$$

defined by

$$\begin{aligned} Me_j &= M_j = e_j\otimes e_{n+1}e_{n+2}, \quad 1\le j\le n,\\ Me_j &= M_j = 1\otimes e_j, \qquad j = n+1, n+2.\end{aligned} \tag{7.3.41}$$

Here we take $\{e_1,\dots,e_n\}$ to generate $\mathcal{C}\ell(n,0)$ and $\{e_{n+1},e_{n+2}\}$ to generate $\mathcal{C}\ell(0,2)$. To extend M in (7.3.40) to (7.3.39), we need to establish the anticommutation relations

$$M_jM_k + M_kM_j = 2\delta_{jk}, \quad 1\le j,k\le n+2. \tag{7.3.42}$$

To get this for $1\le j,k\le n$, we use the computations

$$\begin{aligned} &(e_{n+1}e_{n+2})^2 = -e_{n+1}^2e_{n+2}^2 = -1,\\ &(e_j\otimes e_{n+1}e_{n+2})(e_k\otimes e_{n+1}e_{n+2}) = e_je_k\otimes (e_{n+1}e_{n+2})^2 = -e_je_k\otimes 1,\end{aligned} \tag{7.3.43}$$

which yield

$$\begin{aligned} 1\le j,k\le n \Rightarrow M_jM_k + M_kM_j &= -(e_je_k\otimes 1 + e_ke_j\otimes 1)\\ &= 2\delta_{jk},\end{aligned} \tag{7.3.44}$$

as desired. Next we have

$$\begin{aligned}
&1 \le j \le n \Longrightarrow \\
&M_j M_{n+1} + M_{n+1} M_j \\
&= (e_j \otimes e_{n+1} e_{n+2})(1 \otimes e_{n+1}) + (1 \otimes e_{n+1})(e_j \otimes e_{n+1} e_{n+2}) \\
&= e_j \otimes e_{n+1} e_{n+2} e_{n+1} + e_j \otimes e_{n+1} e_{n+1} e_{n+2} \\
&= 0,
\end{aligned} \tag{7.3.45}$$

since $e_{n+1}e_{n+2} = -e_{n+2}e_{n+1}$. Similarly one gets $M_j M_{n+2} + M_{n+2} M_j = 0$ for $1 \le j \le n$. Next,

$$M_{n+1} M_{n+1} = (1 \otimes e_{n+1})(1 \otimes e_{n+1}) = 1 \otimes e_{n+1}^2 = 1, \tag{7.3.46}$$

and similarly $M_{n+2} M_{n+2} = 1$. Finally,

$$\begin{aligned}
&M_{n+1} M_{n+2} + M_{n+2} M_{n+1} \\
&= (1 \otimes e_{n+1})(1 \otimes e_{n+2}) + (1 \otimes e_{n+2})(1 \otimes e_{n+1}) \\
&= 1 \otimes (e_{n+1} e_{n+2} + e_{n+2} e_{n+1}) \\
&= 0.
\end{aligned} \tag{7.3.47}$$

This establishes (7.3.42). Hence, by Proposition 7.3.1, M extends to the algebra homomorphism (7.3.39) (with $M1 = I$). It is routine to verify that the elements on the right side of (7.3.41) generate $\mathcal{C}\ell(n,0) \otimes \mathcal{C}\ell(0,2)$, so M in (7.3.39) is onto, hence an isomorphism. This completes the proof of Proposition 7.3.5, hence Propositions 7.3.3–7.3.4.

Remark. The following companions to (7.3.38),

$$\begin{aligned}
\mathcal{C}\ell(0,n) \otimes \mathcal{C}\ell(2,0) &\approx \mathcal{C}\ell(n+2,0), \\
\mathcal{C}\ell(p,q) \otimes \mathcal{C}\ell(1,1) &\approx \mathcal{C}\ell(p+1,q+1),
\end{aligned} \tag{7.3.48}$$

have essentially the same proof. From (7.3.38) and (7.3.48) it follows that

$$\mathcal{C}\ell(n+8,0) \approx \mathcal{C}\ell(n,0) \otimes \mathcal{C}\ell(0,2) \otimes \mathcal{C}\ell(2,0) \otimes \mathcal{C}\ell(0,2) \otimes \mathcal{C}\ell(2,0). \tag{7.3.49}$$

Meanwhile, by (7.3.33)–(7.3.34),

$$\mathcal{C}\ell(0,2) \otimes \mathcal{C}\ell(2,0) \approx \mathrm{M}(2,\mathbb{R}) \otimes \mathbb{H}. \tag{7.3.50}$$

This, together with the isomorphism

$$\mathbb{H} \otimes \mathbb{H} \approx \mathrm{M}(4,\mathbb{R}), \tag{7.3.51}$$

leads to

$$\mathcal{C}\ell(n+8,0) \approx \mathcal{C}\ell(n,0) \otimes \mathrm{M}(16,\mathbb{R}). \tag{7.3.52}$$

For use in §7.5, we note that $\mathcal{C}\ell(V,Q)$ has a $\mathbb{Z}/(2)$ grading,

$$\mathcal{C}\ell(V,Q) = \mathcal{C}\ell^0(V,Q) \oplus \mathcal{C}\ell^1(V,Q), \tag{7.3.53}$$

where elements of $\mathcal{C}\ell^0(V,Q)$ have the form (7.3.12) with k restricted to be even and elements of $\mathcal{C}\ell^1(V,Q)$ have such a form with k restricted to be odd. In view of (7.3.1), we have

$$\begin{aligned} \mathcal{C}\ell^0 \cdot \mathcal{C}\ell^0 \subset \mathcal{C}\ell^0, \quad \mathcal{C}\ell^0 \cdot \mathcal{C}\ell^1 \subset \mathcal{C}\ell^1, \\ \mathcal{C}\ell^1 \cdot \mathcal{C}\ell^0 \subset \mathcal{C}\ell^1, \quad \mathcal{C}\ell^1 \cdot \mathcal{C}\ell^1 \subset \mathcal{C}\ell^0. \end{aligned} \tag{7.3.54}$$

If $V = \mathbb{R}^n$ with its standard positive definite inner product, we write (41.53) as

$$\mathcal{C}\ell(n,0) = \mathcal{C}\ell^0(n,0) \oplus \mathcal{C}\ell^1(n,0). \tag{7.3.55}$$

The complexification of $\mathcal{C}\ell^j(V,Q)$ is denoted $\mathbb{C}\ell^j(V,Q)$, and that of $\mathcal{C}\ell^j(n,0)$ is denoted $\mathbb{C}\ell^j(n)$.

7.4. The groups Spin(n)

We will construct Spin(n) as a subset of $\mathcal{C}\ell(n,0)$. A more general construction produces groups $\mathrm{Spin}(p,q) \subset \mathcal{C}\ell(p,q)$, but we will not deal with this here; cf. [**38**], [**26**] for material on this. Let us take $V = \mathbb{R}^n$, with the standard basis $\{e_1, \dots, e_n\}$ and inner product defined by $Q(e_j, e_k) = \langle e_j, e_k \rangle = \delta_{jk}$. We start with the observation that if $v \in V$, $\langle v, v\rangle = 1$, then, for $x \in V$,

$$\begin{aligned} \tau(v)x &= vxv \\ &= -xvv - 2\langle v, x\rangle v \\ &= x - 2\langle v, x\rangle v. \end{aligned} \tag{7.4.1}$$

Hence $\tau(v) : V \to V$ is reflection across the hyperplane $(v)^\perp$. With this in mind, we set

$$\mathrm{Pin}(n) = \{v_1 \cdots v_k \in \mathcal{C}\ell(n,0) : k \in \mathbb{N}, v_j \in \mathbb{R}^n,\ \langle v_j, v_j\rangle = 1\}, \tag{7.4.2}$$

and define

$$\tau : \mathrm{Pin}(n) \longrightarrow \mathrm{O}(n) \tag{7.4.3}$$

by

$$\begin{aligned} \tau(v_1 \cdots v_k)x &= v_1 \cdots v_k x v_k \cdots v_1 \\ &= \tau(v_1) \cdots \tau(v_k)x, \end{aligned} \tag{7.4.4}$$

so $\tau(v_1 \cdots v_k)$ is a product of k reflections of the form (7.4.1).

We need to show that (7.4.4) is well defined, independently of the representation of an element of Pin(n) as a particular product. The following takes care of this.

Lemma 7.4.1. *If $v_j \in \mathbb{R}^n$ are unit vectors, so $u = v_1 \cdots v_k \in \mathrm{Pin}(n)$, then*

$$v_1 \cdots v_k = 1 \Longrightarrow \tau(v_1 \cdots v_k) = I.$$

Proof. First we note that if $v_1 \cdots v_k = 1$, then k is even. In fact, if k is odd, we have from (7.3.54) that $v_1 \cdots v_k \in \mathcal{C}\ell^1(n,0)$. Now that we know k must be even, we have $(v_1 \cdots v_k)(v_k \cdots v_1) = 1$, so in such a case

$$\tau(u)x = uxu^{-1},$$

which is well defined in $\mathcal{C}\ell(n,0)$ for any invertible $u \in \mathcal{C}\ell(n,0)$, independently of the representation of u. $\square$

Note that $\det \tau(v_1 \cdots v_k) = (-1)^k$. Hence, if we set

$$\mathrm{Spin}(n) = \{v_1 \cdots v_k : k \in 2\mathbb{N},\ v_j \in \mathbb{R}^n,\ \langle v_j, v_j\rangle = 1\}, \tag{7.4.5}$$

we have

$$\tau : \mathrm{Spin}(n) \longrightarrow \mathrm{SO}(n). \tag{7.4.6}$$

Note that Pin(n) and Spin(n) are groups, since $(v_1 \cdots v_k)(v_k \cdots v_1) = \pm 1$. Also $\pm v_1 v_1 = \mp 1$, so $\{\pm 1\} \subset \text{Spin}(n)$. The following result is fundamental.

Proposition 7.4.2. *The maps* (7.4.3) *and* (7.4.6) *are surjective, and*

$$\text{Ker}\, \tau = \{\pm 1\}. \tag{7.4.7}$$

First we discuss the surjectivity. For $v \in \mathbb{R}^n$, $\tau(v)$ in (7.4.1) is a reflection, and each reflection on $\mathbb{R}^n$ has this form. Hence the surjectivity follows from:

Proposition 7.4.3. *Each $A \in \mathrm{O}(n)$ is a product of reflections.*

Proof. We make use of the basic linear algebra result that given such A, there exist $Q \in \mathrm{O}(n)$ such that $Q^{-1}AQ$ has the block diagonal form

$$Q^{-1}AQ = \begin{pmatrix} R_{\theta_1} & & & & & & & & \\ & \ddots & & & & & & & \\ & & R_{\theta_k} & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix}. \tag{7.4.8}$$

Since $Q\tau Q^{-1}$ is a reflection whenever τ is a reflection, it suffices to show that the right side of (7.4.8) is a product of reflections. It suffices to consider the separate blocks. In particular, we show that each rotation

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{7.4.9}$$

is a product of two reflections in $\mathbb{R}^2$, of the form

$$\rho(v)x = x - 2\langle v, x\rangle v, \quad x, v \in \mathbb{R}^2,\ |v| = 1. \tag{7.4.10}$$

In fact, it is readily verified that

$$v \in \mathbb{R}^2,\ |v| = 1 \Longrightarrow \rho(R_{\theta/2}v)\rho(v) = R_\theta. \tag{7.4.11}$$

As far as the diagonal entries ± 1 represented as reflections, this is obvious, so the surjectivity assertion of Proposition 7.4.2 is proven. $\square$

Our next task is to establish (7.4.7). To tackle this, suppose $v_j \in \mathbb{R}^n$ are unit vectors such that

$$\tau(u) = I, \quad u = v_1 \cdots v_k. \tag{7.4.12}$$

Since each $\tau(v_j)$ has determinant -1, k is even in (7.4.12). Therefore we have $(v_k \cdots v_1)(v_1 \cdots v_k) = 1$. Referring to (7.4.4), we have

$$\begin{aligned} &\tau(u)x = x, \quad \forall\, x \in \mathbb{R}^n \\ &\Leftrightarrow ux = xu, \quad \forall\, x \in \mathbb{R}^n, \\ &\Leftrightarrow xux = -|x|^2 u, \quad \forall\, x \in \mathbb{R}^n \\ &\Leftrightarrow u = -e_j u e_j, \end{aligned} \tag{7.4.13}$$

for the standard orthonormal basis $\{e_j\}$ of $\mathbb{R}^n$. Now, using Proposition 7.3.2, set

$$u = \sum_{i_\nu = 0 \text{ or } 1} a_{i_1 \cdots i_n} e_1^{i_1} \cdots e_n^{i_n}. \tag{7.4.14}$$

We have, with $i_1 + \cdots + i_n = 2\ell$,

$$\begin{aligned} &- e_j (e_1^{i_1} \cdots e_j^{i_j} \cdots e_n^{i_n}) e_j \\ &= (-1)^{2\ell - i_j + 1} e_1^{i_1} \cdots e_j^{i_j + 2} \cdots e_n^{i_n} \\ &= (-1)^{2\ell - i_j + 2} e_1^{i_1} \cdots e_j^{i_j} \cdots e_n^{i_n}. \end{aligned} \tag{7.4.15}$$

Hence, for u as in (7.4.14), if (7.4.13) holds, then

$$u = \sum_{i_\nu = 0 \text{ or } 1} (-1)^{i_j} a_{i_1 \cdots i_n} e_1^{i_1} \cdots e_n^{i_n}, \quad \forall\, j. \tag{7.4.16}$$

Given Proposition 7.3.2, we deduce that, for u of the form (7.4.14),

$$\tau(u) = I \Longrightarrow a_{i_1 \cdots i_n} = 0, \quad \text{except for } a_{0 \cdots 0}, \tag{7.4.17}$$

which gives (7.4.7), and completes the proof of Proposition 7.4.2.

The next result is an important complement to Proposition 7.4.2, to the effect that (7.4.6) presents Spin(n) as a *connected* double cover of SO(n).

Proposition 7.4.4. *For each $n \ge 2$, Spin(n) is connected.*

Proof. Since we know SO(n) is connected, it suffices to show that there is a continuous path in Spin(n) from 1 to -1. Set

$$\gamma(t) = e_1 \cdot \big((\cos t) e_1 + (\sin t) e_2 \big), \quad 0 \le t \le \pi. \tag{7.4.18}$$

We have $\gamma : [0, \pi] \to \mathrm{Spin}(n)$, and

$$\gamma(0) = -1, \quad \gamma(\pi) = 1, \tag{7.4.19}$$

so Proposition 7.4.4 is proven. □

We examine the Lie algebra spin(n) of Spin(n), i.e., the tangent space to Spin(n) at 1. The Lie algebra $\mathfrak{so}(n)$ of SO(n) is spanned by elements

$$J_{jk} = -E_{jk} + E_{kj}, \quad j < k, \tag{7.4.20}$$

where $E_{jk} \in \mathrm{M}(n, \mathbb{R})$ is defined by

$$E_{jk} e_\ell = \delta_{k\ell} e_j. \tag{7.4.21}$$

The element J_{jk} generates the group $R_{jk}(t) = e^{tJ_{jk}}$ of rotations in the $e_j - e_k$ plane, given by

$$\begin{aligned} R_{jk}(t)e_j &= (\cos t)e_j + (\sin t)e_k \\ R_{jk}(t)e_k &= -(\sin t)e_j + (\cos t)e_k \\ R_{jk}(t)e_\ell &= e_\ell, \qquad \ell \notin \{j, k\}. \end{aligned} \tag{7.4.22}$$

Comparing (7.4.1) with (7.4.10)–(7.4.11), we see that

$$R_{jk}(t) = \tau\big(R_{jk}(t/2)e_j \cdot e_j\big) = \tau(-R_{jk}(t/2)e_j \cdot e_j). \tag{7.4.23}$$

The curves $\gamma_{jk}(t) = -R_{jk}(t/2)e_j \cdot e_j$ are curves in Spin(n) through the group identity 1. Since $R'_{jk}(0) = J_{jk}$, we have

$$\gamma'_{jk}(0) = \frac{1}{2} e_j e_k \in T_1 \operatorname{Spin}(n) = \operatorname{spin}(n). \tag{7.4.24}$$

From (42.23),

$$\begin{aligned} R'_{jk}(0) &= -\frac{1}{2} D\tau(1)(R'_{jk}(0)e_j \cdot e_j) \\ &= \frac{1}{2} D\tau(1)(e_j e_k), \end{aligned} \tag{7.4.25}$$

hence

$$d\tau(e_j e_k) = 2J_{jk}. \tag{7.4.26}$$

Hence

$$\operatorname{spin}(n) = \operatorname{Span}\{e_j e_k : j < k\} \subset \mathcal{C}\ell(n, 0). \tag{7.4.27}$$

Note that, for $j < k$,

$$\begin{aligned} (e_j e_k)^2 = -1 \Longrightarrow e^{t e_j e_k} &= \cos t + (\sin t) e_j e_k \\ &= -(\cos t) e_j^2 + (\sin t) e_j e_k \\ &= -\big((\cos t)e_j + (\sin t)e_k\big) \cdot e_j \\ &= -R_{jk}(t) e_j \cdot e_j, \end{aligned} \tag{7.4.28}$$

so we recover the result implicit in (7.4.23)–(7.4.26) that the one-parameter group in Spin(n) generated by $e_j e_k$ is

$$\operatorname{Exp}(t e_j e_k) = -R_{jk}(t) e_j \cdot e_j. \tag{7.4.29}$$

Either by a calculation or by applying analogues of reasoning done in §3.2, we see that the Lie bracket on spin(n) is given by

$$[e_j e_k, e_\ell e_m] = e_j e_k e_\ell e_m - e_\ell e_m e_j e_k, \tag{7.4.30}$$

and

$$d\tau([e_j e_k, e_\ell e_m]) = [2J_{jk}, 2J_{\ell m}]. \tag{7.4.31}$$

The reader is invited to verify that the right side of (7.4.30) belongs to the space described in (7.4.27).

The space

$$\mathfrak{h} = \operatorname{Span}\{E_j : 1 \le j \le k\} \subset \mathfrak{so}(n), \quad E_j = J_{2j-1,2j}, \tag{7.4.32}$$

is the Lie algebra of a maximal torus of SO(n), when $n = 2k$ or $n = 2k+1$. The preimage under $d\tau$ is

$$\tilde{\mathfrak{h}} = \operatorname{Span}\{e_{2j-1}e_{2j} : 1 \le j \le k\} \subset \operatorname{spin}(n). \tag{7.4.33}$$

By (7.4.28)–(7.4.29), we can say that there exists $\varepsilon > 0$ such that, with

$$\begin{gathered} \Psi : \prod_{1\le j<k\le n} \mathbb{R} \longrightarrow \operatorname{Spin}(n), \\ \Psi((t_{jk})) = \prod_{j<k} \operatorname{Exp}(t_{jk}e_j e_k) = \prod_{j<k} -R_{jk}(t_{jk})e_j \cdot e_j, \end{gathered} \tag{7.4.34}$$

there is a neighborhood $\mathcal{O}$ of $1 \in \operatorname{Spin}(n)$ such that

$$\Psi : \prod_{1\le j<k\le n} (-\varepsilon, \varepsilon) \longrightarrow \mathcal{O}, \quad \text{diffeomorphically.} \tag{7.4.35}$$

Alternative description of Spin(n)

Since $\mathcal{C}\ell(n,0)$ is a finite-dimensional associative algebra over $\mathbb{R}$ with unit, the set $\mathcal{C}\ell^{\rm inv}(n,0)$ of invertible elements is a nonempty open subset, forming a multiplicative group. We have a representation ρ of $\mathcal{C}\ell^{\rm inv}(n,0)$ on $\mathcal{C}\ell(n,0)$, given by

$$\rho(u)w = uwu^{-1}, \quad u \in \mathcal{C}\ell^{\rm inv}(n,0),\ w \in \mathcal{C}\ell(n,0). \tag{7.4.36}$$

The set

$$\mathcal{P}(n,0) = \{u \in \mathcal{C}\ell^{\rm inv}(n,0) : \rho(u) : \mathbb{R}^n \to \mathbb{R}^n\} \tag{7.4.37}$$

(regarding $\mathbb{R}^n \subset \mathcal{C}\ell(n,0)$) is a subgroup of $\mathcal{C}\ell^{\rm inv}(n,0)$, containing $\{v \in \mathbb{R}^n : |v| = 1\}$, by (7.4.21). We can describe

$$\begin{gathered} \operatorname{Spin}(n) = \\ \text{subgroup of } \mathcal{P}(n,0) \text{ generated by } \{v_1 v_2 : v_j \in \mathbb{R}^n,\ |v_j| = 1\}, \end{gathered} \tag{7.4.38}$$

and

$$\operatorname{Pin}(n) = \operatorname{Spin}(n) \cup e_1 \cdot \operatorname{Spin}(n). \tag{7.4.39}$$

7.5. Spinor representations

Let V be an n-dimensional real vector space, with a positive definite inner product $\langle\ ,\ \rangle$. We want to associate a representation of $\mathcal{C}\ell(V,\langle\ ,\ \rangle)$ and associated objects on a space of spinors, which we will define below. To construct this space we need some extra structure on V.

First consider the case where n is even, i.e., $n = 2k$. We assume there is given a complex structure on V, i.e., a linear map $J : V \to V$ satisfying $J^2 = -I$, and that J is an isometry with respect to $\langle\ ,\ \rangle$. We denote by $\mathcal{V}$ the k-dimensional complex vector space (V, J), and we endow $\mathcal{V}$ with a hermitian inner product

$$(u, v) = \langle u, v\rangle + i\langle u, Jv\rangle. \tag{7.5.1}$$

We set

$$S = S(V, \langle\ ,\ \rangle, J) = \Lambda^*_{\mathbb{C}}\mathcal{V} = \bigoplus_{j=0}^{k} \Lambda^j_{\mathbb{C}}\mathcal{V}. \tag{7.5.2}$$

The inner product (7.5.1) defines a conjugate linear isomorphism $\mathcal{V} \to \mathcal{V}'$, which gives a conjugate linear isomorphism $\Lambda^*_{\mathbb{C}}\mathcal{V} \to \Lambda^*_{\mathbb{C}}\mathcal{V}' \approx (\Lambda^*_{\mathbb{C}}\mathcal{V})'$, *hence a hermitian inner product on $\Lambda^*_{\mathbb{C}}\mathcal{V}$. Concretely, if $\{v_1, \dots, v_n\}$ is an orthonormal basis of $\mathcal{V}$, then

$$\{v_{j_1} \wedge \cdots \wedge v_{j_\ell} : j_1 < \cdots < j_\ell\} \text{ is an orthonormal basis of } \Lambda^\ell_{\mathbb{C}}\mathcal{V}. \tag{7.5.3}$$

We may as well take $V = \mathbb{R}^n$, $\{e_j : 1 \le j \le n\}$ the standard basis, with $\langle e_j, e_k\rangle = \delta_{jk}$, and define J by $Je_{2j-1} = e_{2j}$, $Je_{2j} = -e_{2j-1}$, $1 \le j \le k$. (Recall $n = 2k$). Then $(V, J) = \mathbb{C}^k$, with orthonormal basis $\{v_j = e_{2j} : 1 \le j \le k\}$, and

$$S = \Lambda^*\mathbb{C}^k. \tag{7.5.4}$$

In order to define a representation of $\mathcal{C}\ell(n, 0)$ on S, we produce an $\mathbb{R}$-linear map

$$M : \mathbb{R}^n \longrightarrow \mathrm{End}_{\mathbb{C}}(\Lambda^*_{\mathbb{C}}\mathcal{V}), \tag{7.5.5}$$

in the form

$$M(v) = \wedge_v - j_v, \tag{7.5.6}$$

where

$$\wedge_v : \Lambda^\ell_{\mathbb{C}}\mathcal{V} \longrightarrow \Lambda^{\ell+1}_{\mathbb{C}}\mathcal{V}, \quad \wedge_v\varphi = v \wedge \varphi, \tag{7.5.7}$$

with v interpreted as an element of $\mathcal{V}$, and

$$j_v : \Lambda^{\ell+1}_{\mathbb{C}} \longrightarrow \Lambda^\ell_{\mathbb{C}}\mathcal{V}, \quad j_v\psi = (\wedge_v)^*\psi, \tag{7.5.8}$$

that is,

$$(v \wedge \varphi, \psi) = (\varphi, j_v \psi), \quad \varphi \in \Lambda^\ell_{\mathbb{C}} \mathcal{V},\ \psi \in \Lambda^{\ell+1}_{\mathbb{C}} \mathcal{V}. \tag{7.5.9}$$

We claim that, for $u, v \in \mathbb{R}^n$,

$$M(u)M(v) + M(v)M(u) = -2\langle u, v\rangle I. \tag{7.5.10}$$

It suffices to show that

$$M(v)^2 = -\langle v, v\rangle I, \tag{7.5.11}$$

and insert $u \pm v$ into this identity. To prove (7.5.11), we can assume $\langle v, v\rangle = 1$. Pick an orthonormal basis $\{v_1, \dots, v_k\}$ for $\mathbb{C}^k$ with $v_1 = v$. Then use of (43.9) establishes that, for $j_1 < \cdots < j_{\ell+1}$,

$$\begin{aligned} j_v(v_{j_1} \wedge \cdots \wedge v_{j_{\ell+1}}) = v_{j_2} \wedge \cdots \wedge v_{j_{\ell+1}} \quad &\text{if } j_1 = 1, \\ 0 \quad &\text{if } j_1 > 1. \end{aligned} \tag{7.5.12}$$

Since $\wedge_v^2 = 0$ and (hence) $j_v^2 = 0$, we get

$$M(v)^2 = -(\wedge_v j_v + j_v \wedge_v) = -\langle v, v\rangle I, \tag{7.5.13}$$

the last identity via (7.5.12).

Now Proposition 7.3.1 implies M extends to a homomorphism of algebras

$$M : \mathcal{C}\ell(2k, 0) \longrightarrow \mathrm{End}_{\mathbb{C}}(\Lambda^* \mathbb{C}^k), \tag{7.5.14}$$

which in turn extends to a $\mathbb{C}$-linear algebra homomorphism

$$M : \mathbb{C}\ell(2k) \longrightarrow \mathrm{End}_{\mathbb{C}}(\Lambda^* \mathbb{C}^k). \tag{7.5.15}$$

The following is fundamental.

Proposition 7.5.1. *In* (7.5.15)*, M is an isomorphism of algebras.*

Proof. Note that $\dim_{\mathbb{C}} \Lambda^* \mathbb{C}^k = 2^k$; hence

$$\dim_{\mathbb{C}} \mathrm{End}_{\mathbb{C}}(\Lambda^* \mathbb{C}^k) = 2^{2k} = \dim_{\mathbb{C}} \mathbb{C}\ell(2k). \tag{7.5.16}$$

Thus it suffices to prove M is injective. Clearly $M(v) \neq 0$ for nonzero $v \in \mathbb{R}^n$, and $\operatorname{Ker} M$ must be a two-sided ideal in $\mathbb{C}\ell(2k)$. Recall from (7.3.36) that $\mathbb{C}\ell(2k) \approx \mathrm{M}(2^k, \mathbb{C})$. It is a fact that, for each $m \in \mathbb{N}$,

$$\mathrm{M}(m, \mathbb{C}) \ \text{ has no proper two-sided ideals;} \tag{7.5.17}$$

i.e., $\mathrm{M}(m, \mathbb{C})$ is *simple*. See §B.6 for a proof. This finishes the proof of Proposition 7.5.1. ☐

The algebra homomorphism M in (7.5.15) restricts to $\mathrm{Pin}(2k) \subset \mathcal{C}\ell(2k, 0)$, yielding a group homomorphism

$$D_{1/2} : \mathrm{Pin}(2k) \longrightarrow \mathrm{Gl}(\Lambda^* \mathbb{C}^k), \tag{7.5.18}$$

i.e., a representation of $\mathrm{Pin}(2k)$ on $\Lambda^* \mathbb{C}^k$. Since the linear span of $\mathrm{Pin}(2k)$ (over $\mathbb{R}$) is $\mathcal{C}\ell(2k, 0)$, we have from Proposition 7.5.1 that:

Corollary 7.5.2. *The representation* $D_{1/2}$ *of* $\mathrm{Pin}(2k)$ *on* $S = \Lambda^*\mathbb{C}^k$ *is irreducible.*

Remark. The operator $D_{1/2}(g)$ is unitary for each $g \in \mathrm{Pin}(2k)$. In fact, if $v \in \mathbb{R}^{2k}$ and $\langle v, v\rangle = 1$, then $M(v)$ is skew-adjoint and $M(v)^2 = -I$, so $\operatorname{Spec} M(v) \subset \{\pm i\}$, and hence $M(v)$ is unitary.

The restriction of $D_{1/2}$ to $\mathrm{Spin}(2k)$ is not irreducible. In fact, the spaces

$$\bigoplus_{j \text{ even}} \Lambda^j_{\mathbb{C}}\mathbb{C}^k = S_+(2k), \quad \bigoplus_{j \text{ odd}} \Lambda^j_{\mathbb{C}}\mathbb{C}^k = S_-(2k) \tag{7.5.19}$$

are invariant under the action of $D_{1/2}$ restricted to $\mathrm{Spin}(2k)$, and more generally under the action of M restricted to $\mathcal{C}\ell^0(2k, 0)$, defined as in (7.3.53)–(7.3.55). We have

$$M : \mathbb{C}\ell^0(2k) \longrightarrow \mathrm{End}_{\mathbb{C}}(S_+(2k)) \oplus \mathrm{End}_{\mathbb{C}}(S_-(2k)). \tag{7.5.20}$$

Note that

$$v \in \mathbb{R}^{2k},\ v \neq 0 \Longrightarrow M(v) : S_\pm(2k) \to S_\mp(2k), \tag{7.5.21}$$

while left multiplication by v takes $\mathbb{C}\ell^0(2k)$ to $\mathbb{C}\ell^1(2k)$, and these maps are all isomorphisms. We have

$$\dim_{\mathbb{C}} \mathrm{End}_{\mathbb{C}}\, S_+(2k) \oplus \mathrm{End}_{\mathbb{C}}\, S_-(2k) = 2^{2k-1} = \dim_{\mathbb{C}} \mathbb{C}\ell^0(2k). \tag{7.5.22}$$

We already know from Proposition 7.5.1 that M in (7.5.20) is injective, so it is an isomorphism. We deduce the following.

Corollary 7.5.3. *The representation* $D_{1/2}$ *restricted to* $\mathrm{Spin}(2k)$ *splits into two factors:*

$$D^\pm_{1/2} : \mathrm{Spin}(2k) \longrightarrow \mathrm{Gl}(S_\pm(2k)), \tag{7.5.23}$$

and both are irreducible.

We next discuss the spinor representation of $\mathrm{Spin}(2k-1)$. If $\{e_1, \dots, e_{2k}\}$ is the standard basis of $\mathbb{R}^{2k}$, and $\mathbb{R}^{2k-1} = \mathrm{Span}\,\{e_1, \dots, e_{2k-1}\}$, the map

$$\mathbb{R}^{2k-1} \longrightarrow \mathcal{C}\ell^0(2k, 0), \quad v \mapsto ve_{2k}, \quad v \in \mathbb{R}^{2k-1} \tag{7.5.24}$$

satisfies the analogue of (7.3.7) and hence, via Proposition 7.3.1, gives rise to a homomorphism of algebras

$$\kappa : \mathcal{C}\ell(2k-1, 0) \longrightarrow \mathcal{C}\ell^0(2k, 0). \tag{7.5.25}$$

Explicitly,

$$\begin{aligned}\kappa\Big(\sum_{j_1<\cdots<j_\ell} a_{j_1\cdots j_\ell} e_{j_1}\cdots e_{j_\ell}\Big)\\ = \sum_{j_1<\cdots<j_\ell} a_{j_1\cdots j_\ell} e_{j_1} e_{2k}\cdots e_{j_\ell} e_{2k}.\end{aligned} \tag{7.5.26}$$

Now, since $e_j e_{2k} = -e_{2k} e_j$ for $j < 2k$ and $e_{2k}^2 = -1$,

$$\begin{aligned} & e_{j_1} e_{2k} e_{j_2} e_{2k} \cdots e_{j_\ell} e_{2k} \\ &\quad = e_{j_1} e_{j_2} e_{j_3} e_{2k} \cdots e_{j_\ell} e_{2k} \\ &\quad = \cdots \\ &\quad = e_{j_1} e_{j_2}\cdots e_{j_\ell} \qquad \text{if } \ell \text{ is even} \\ &\quad \ e_{j_1} e_{j_2}\cdots e_{j_\ell} e_{2k} \quad \text{if } \ell \text{ is odd,}\end{aligned} \tag{7.5.27}$$

so

$$\begin{aligned} &\kappa\Big(\sum_{\ell=0}^{2k-1}\sum_{j_1<\cdots<j_\ell} a_{j_1\cdots j_\ell} e_{j_1}\cdots e_{j_\ell}\Big) \\ &= \sum_{\ell \text{ even}}\sum_{j_1<\cdots<j_\ell} a_{j_1\cdots j_\ell} e_{j_1}\cdots e_{j_\ell} \\ &\quad + \sum_{\ell \text{ odd}}\sum_{j_1<\cdots<j_\ell} a_{j_1\cdots j_\ell} e_{j_1}\cdots e_{j_\ell} e_{2k}.\end{aligned} \tag{7.5.28}$$

The map κ is thus clearly injective. Since the dimensions match, it is an isomorphism.

The inclusion $\mathrm{Pin}(2k-1) \subset \mathcal{C}\ell(2k-1, 0)$ gives an inclusion

$$\mathrm{Pin}(2k-1) \hookrightarrow \mathrm{Spin}(2k), \tag{7.5.29}$$

and restricting $D^+_{1/2}$ gives a representation

$$D^+_{1/2} : \mathrm{Pin}(2k-1) \longrightarrow \mathrm{Gl}(S_+(2k)). \tag{7.5.30}$$

There is also a representation $D^-_{1/2}$ of $\mathrm{Pin}(2k-1)$ on $S_-(2k))$, but these two are intertwined by the isomorphism $M(e_{2k}) : S_+(2k) \to S_-(2k)$.

Proposition 7.5.4. *The representation* (7.5.30) *is irreducible.*

Proof. In (7.5.25), κ is an isomorphism, in (7.5.20), M is an isomorphism, and $\mathrm{Pin}(2k-1)$ spans $\mathcal{C}\ell(2k-1, 0)$. □

Remark. Restriction of (7.5.30) to $\mathrm{Spin}(2k-1)$ gives a representation

$$D^+_{1/2} : \mathrm{Spin}(2k-1) \longrightarrow \mathrm{Gl}(S_+(2k)). \tag{7.5.31}$$

In §7.6 we show that this is also irreducible.

7.6. Weight spaces for the spinor representations

In §7.5 we constructed representations of $\mathrm{Spin}(n)$ on $S_\pm(n) = \Lambda^{\mathrm{even/odd}}\mathbb{C}^k$ in case $n = 2k$ and on $S_+(n) = \Lambda^{\mathrm{even}}\mathbb{C}^k$ in case $n = 2k-1$. Here we will show that the monomials in these subspaces of $\Lambda^*\mathbb{C}^k$ are weight vectors, compute the weights, and identify the highest weights. Our ordered basis of $\tilde{\mathfrak{h}}$, the Lie algebra of a maximal torus in $\mathrm{Spin}(n)$ described in (7.4.33), will be

$$\Bigl\{\frac{1}{2}e_1e_2, \frac{1}{2}e_3e_4, \dots\Bigr\} = \Bigl\{\frac{1}{2}e_{2j-1}e_{2j} : 1 \le j \le k\Bigr\}, \tag{7.6.1}$$

given $n = 2k$ or $2k+1$. Recall that this maps via $d\tau$ to the ordered basis

$$\{J_{2j-1,2j} : 1 \le j \le k\} \tag{7.6.2}$$

for the Lie algebra of a maximal torus of $\mathrm{SO}(n)$, described in (7.4.20); cf. (7.4.26).

We first treat the case $n = 2k$. To get started, note that

$$\begin{aligned} &\gamma_{ij}(t) = e^{t(e_ie_j)} = (\cos t)1 + (\sin t)e_ie_j \\ &\Longrightarrow D^\pm_{1/2}(\gamma_{ij}(t))\varphi = (\cos t)\varphi + (\sin t)M(e_i)M(e_j)\varphi \\ &\Longrightarrow dD^\pm_{1/2}(e_ie_j)\varphi = M(e_i)M(e_j)\varphi = M_iM_j\varphi, \end{aligned} \tag{7.6.3}$$

for $1 \le i,j \le 2k,\ \varphi \in \Lambda^*\mathbb{C}^k$, and, as in (7.5.6), $M(v) = \wedge_v - j_v,\ v \in \mathbb{R}^{2k}$, and we introduce simplified notation

$$M_i = M(e_i) = \wedge_{e_i} - j_{e_i} = \wedge_i - j_i. \tag{7.6.4}$$

Hence

$$dD^\pm_{1/2}(e_{2j-1}e_{2j})\varphi = M_{2j-1}M_{2j}\varphi = (\wedge_{2j-1} - j_{2j-1})(\wedge_{2j} - j_{2j})\varphi. \tag{7.6.5}$$

Since the wedge product is in $\Lambda^*\mathbb{C}^k$, and in $\mathbb{C}^k$ we have $e_{2j} = ie_{2j-1}$, it follows that $\wedge_{2j}\wedge_{2j-1} = 0$, and similarly $j_{2j}j_{2j-1} = 0$, while $\wedge_{2j-1} = -i\wedge_{2j}$ and $j_{2j-1} = ij_{2j}$. Hence

$$\begin{aligned} dD^\pm_{1/2}(e_{2j-1}e_{2j})\varphi &= -i(j_{2j}\wedge_{2j} - \wedge_{2j}\, j_{2j})\varphi \\ &= i(1 - 2j_{2j}\wedge_{2j})\varphi, \end{aligned} \tag{7.6.6}$$

the last identity using $\wedge_{2j}j_{2j} + j_{2j}\wedge_{2j} = 1$.

Let us take φ to be a monomial in $\Lambda^\ell\mathbb{C}^k$, with respect to the basis of $\mathbb{C}^k$ given by

$$\{v_j : 1 \le j \le k\}, \quad v_j = e_{2j} = ie_{2j-1}. \tag{7.6.7}$$

We have

$$dD^\pm_{1/2}(e_{2j-1}e_{2j}) = iQ_j \tag{7.6.8}$$

where, for $1 \le i_1 < \cdots < i_\ell \le k$,

$$Q_j\, v_{i_1} \wedge \cdots \wedge v_{i_\ell} = \begin{array}{ll} -\, v_{i_1} \wedge \cdots \wedge v_{i_\ell} & \text{if } j \notin \{i_1, \ldots, i_\ell\} \\ +\, v_{i_1} \wedge \cdots \wedge v_{i_\ell} & \text{if } j \in \{i_1, \ldots, i_\ell\}. \end{array} \tag{7.6.9}$$

In particular,

$$\frac{1}{2} Q_j\, v_1 \wedge \cdots \wedge v_k = \frac{1}{2} v_1 \wedge \cdots \wedge v_k, \quad \forall\, j \in \{1, \ldots, k\}, \tag{7.6.10}$$

and

$$\frac{1}{2} Q_j\, v_1 \wedge \cdots \wedge v_{k-1} = \begin{array}{ll} +\,\frac{1}{2} v_1 \wedge \cdots \wedge v_{k-1} & \text{if } j \in \{1, \ldots, k-1\} \\ -\,\frac{1}{2} v_1 \wedge \cdots \wedge v_{k-1} & \text{if } j = k. \end{array} \tag{7.6.11}$$

These calculations prove the following:

Proposition 7.6.1. *The vector $v_1 \wedge \cdots \wedge v_k$ is a highest weight vector with weight*

$$\Big(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2}\Big). \tag{7.6.12}$$

The vector $v_1 \wedge \cdots \wedge v_{k-1}$ is a weight vector with weight

$$\Big(\frac{1}{2}, \ldots, \frac{1}{2}, -\frac{1}{2}\Big). \tag{7.6.13}$$

The weight (7.6.12) is the highest weight for $D^+_{1/2}$ on $S_+(2k)$ if k is even and for $D^-_{1/2}$ on $S_-(2k)$ if k is odd, and vice-versa for the weight (7.6.13).

We turn to the case $n = 2k - 1$. Then we replace the basis (7.6.1) by

$$\Big\{\frac{1}{2} e_{2j-1} e_{2j} : 1 \le j \le k-1\Big\}. \tag{7.6.14}$$

Recalling the description (7.5.24)–(7.5.29) of the representation $D^+_{1/2}$ of $\mathrm{Spin}(2k-1)$ on $S_+(2k) = \Lambda^{\mathrm{even}}\mathbb{C}^k$, we bring in the following counterpoint to (7.6.3), for $1 \le i, j \le 2k - 1$:

$$\begin{aligned} \gamma_{ij}(t) &= e^{t(e_i e_{2k} e_j e_{2k})} \\ &= e^{t(e_i e_j)} \\ &= (\cos t) 1 + (\sin t) e_i e_j, \end{aligned} \tag{7.6.15}$$

since, for $i, j < 2k$, we have $e_i e_{2k} e_j e_{2k} = -e_1 e_j e_{2k}^2 = e_i e_j$. Hence $dD^+_{1/2}(e_i e_j)$ is given exactly by the formula (7.6.3), and the calculations (7.6.4)–(7.6.10) need essentially no further changes. We have

$$dD^+_{1/2}(e_{2j-1} e_{2j}) = iQ_j, \quad 1 \le j \le k-1, \tag{7.6.16}$$

where, for $1 \le i_1 < \cdots < i_\ell \le k$, $Q_j\, v_{i_1} \wedge \cdots \wedge v_{i_\ell}$ is given by (44.9). We thus have the following counterpart to Proposition 7.6.1.

Proposition 7.6.2. *The representation $D^+_{1/2}$ of* $\mathrm{Spin}(2k-1)$ *on* $S_+(2k)$ *has highest weight vector* $v_1 \wedge \cdots \wedge v_\kappa$*, with* $\kappa = k$ *if* k *is even and* $k-1$ *if* k *is odd, and its highest weight is given by the* $(k-1)$*-tuple*

$$\Bigl(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2}\Bigr). \tag{7.6.17}$$

We now deduce the following refinement of Proposition 7.5.4.

Corollary 7.6.3. *The representation $D^+_{1/2}$ of* $\mathrm{Spin}(2k-1)$ *on* $S_+(2k)$ *is irreducible.*

Proof. If not, there would be an irreducible component with highest weight different from (7.6.17). However, none of the other weights arising in (7.6.8)–(7.6.9) for the representation $D^+_{1/2}$ of $\mathrm{Spin}(2k-1)$ are dominant integral weights, so they cannot be highest weights of a representation of $\mathrm{Spin}(2k-1)$. ☐

In view of the tensor product result of Proposition 6.2.2, we can combine Propositions 7.6.1–7.6.2 with Proposition 7.2.13, on representations of $\mathrm{SO}(n)$, to obtain highest weights for representations of $\mathrm{Spin}(n)$. A comparison with Proposition 7.2.9 gives the following definitive result.

Proposition 7.6.4. *Assume* $n \geq 3$*. The highest weights of irreducible representations of* $\mathrm{Spin}(n)$ *are precisely the dominant integral weights of* $\mathfrak{so}(n)$*.*

The following is an incidental corollary.

Corollary 7.6.5. *If* $n \geq 3$*,* $\mathrm{Spin}(n)$ *is simply connected.*

Proof. Since the center of $\mathfrak{so}(n)$ is trivial for $n \geq 3$, a general result established in §E.3 implies the universal covering group $\widetilde{G}$ of $\mathrm{SO}(n)$ is compact. We have a covering homomorphism $\mu : \widetilde{G} \to \mathrm{Spin}(n)$, and claim μ is an isomorphism. Each irreducible representation π of $\mathrm{Spin}(n)$ yields an irreducible representation $\pi \circ \mu$ of $\widetilde{G}$, with the same derived action on $\mathfrak{g} = \mathrm{spin}(n) = \mathfrak{so}(n)$, and the same highest weight. If μ were not an isomorphism, there would have to be other irreducible representations of $\widetilde{G}$, by the Peter-Weyl theorem, but there are no further dominant integral weights available. Thus μ must be an isomorphism. ☐

One can also give a purely topological proof of this simple connectivity, though it requires background in homotopy theory, which can be found in [**36**]. The argument goes as follows. The quotient result

$$\mathrm{SO}(n+1)/\,\mathrm{SO}(n) = S^n \tag{7.6.18}$$

yields a homotopy exact sequence

$$\pi_{k+1}(S^n) \longrightarrow \pi_k(\mathrm{SO}(n)) \longrightarrow \pi_k(\mathrm{SO}(n+1)) \longrightarrow \pi_k(S^n), \tag{7.6.19}$$

where $\pi_k(M)$ is the group of homotopy classes of continuous maps $S^k \to M$. See [**36**], p. 91 for this.

Now $0 < k < n \Rightarrow \pi_k(S^n) = 0$, and $k + 1 < n \Rightarrow \pi_{k+1}(S^n) = 0$, so

$$n \geq 3 \Longrightarrow \pi_1(\mathrm{SO}(n)) \approx \pi_1(\mathrm{SO}(n+1)). \tag{7.6.20}$$

We see directly from the 2-to-1 homomorphism $\mathrm{SU}(2) \to \mathrm{SO}(3)$ that $\pi_1(\mathrm{SO}(3)) \approx \mathbb{Z}/(2)$, so

$$n \geq 3 \Longrightarrow \pi_1(\mathrm{SO}(n)) \approx \mathbb{Z}/(2). \tag{7.6.21}$$

Thus

$$n \geq 3 \Longrightarrow \pi_1(\mathrm{Spin}(n)) = 0. \tag{7.6.22}$$

Alternatively, using

$$\mathrm{Spin}(n+1)/\,\mathrm{Spin}(n) = S^n \tag{7.6.23}$$

in place of (7.6.18) yields analogues of (7.6.19)–(7.6.20), including

$$n \geq 3 \Longrightarrow \pi_1(\mathrm{Spin}(n)) \approx \pi_1(\mathrm{Spin}(n+1)), \tag{7.6.24}$$

and then (7.6.22) follows from

$$\pi_1(\mathrm{Spin}(3)) = \pi_1(\mathrm{SU}(2)) = \pi_1(S^3) = 0. \tag{7.6.25}$$

Chapter 8

$SO(n)$, harmonic functions, and analysis on spheres

This chapter examines fruitful interactions of the study of the action of $SO(n)$ on functions on $\mathbb{R}^n$ and on S^{n-1} with the study of harmonic functions on domains in $\mathbb{R}^n$ and of a class of special functions on S^{n-1} known as spherical harmonics.

Section 8.1 provides basic material on harmonic functions on a domain $\Omega \subset \mathbb{R}^n$, i.e., solutions to $\Delta u = 0$ on Ω, where $\Delta = \partial_1^2 + \cdots + \partial_n^2$ is the Laplace operator. The relevance of $SO(n)$ arises because Δ commutes with the action of $SO(n)$ on functions on $\mathbb{R}^n$ given by $u(x) \mapsto u(gx)$. We use this to give a proof of the mean value property,

$$u(p) = \mathrm{Avg}_{\partial B_r(p)}\, u, \tag{8.0.1}$$

for a harmonic function u on Ω, given $\overline{B_r(p)} \subset \Omega$. The proof given in Proposition 8.1.1 differs from standard proofs in avoiding the use of the divergence theorem. From (8.0.1) we proceed to the maximum principle for harmonic functions, hence uniqueness of solutions to the Dirichlet problem

$$\Delta u = 0 \ \text{ on } \ \Omega, \quad u = f \text{ on } \partial\Omega, \quad u \in C^2(\Omega) \cap C(\overline{\Omega}), \tag{8.0.2}$$

when $\Omega \subset \mathbb{R}^n$ is a bounded domain and $f \in C(\partial\Omega)$. We next examine existence of solutions to (8.0.2) when $\Omega = B^n$ is the unit ball. We start with

$n = 2$, producing the solution

$$u(re^{i\theta}) = \sum_{k=-\infty}^{\infty} r^{|k|} \hat{f}(k) e^{ik\theta}, \tag{8.0.3}$$

where $\hat{f}(k)$ are the Fourier coefficients of f,

$$\hat{f}(k) = \frac{1}{2\pi} \int_{\mathbb{T}} f(\theta) e^{-ik\theta} \, d\theta. \tag{8.0.4}$$

Plugging (8.0.4) into (8.0.3) and summing a geometric series yields the formula

$$u(re^{i\theta}) = \frac{1 - r^2}{2\pi} \int_{\mathbb{T}} \frac{f(\varphi)}{1 - 2r\cos(\theta - \varphi) + r^2} \, d\varphi, \tag{8.0.5}$$

and a change of variable yields

$$u(x) = \mathrm{PI}\, f(x) = \frac{1 - |x|^2}{2\pi} \int_{S^1} \frac{f(y)}{|x - y|^2} \, ds(y). \tag{8.0.6}$$

We take this as a cue to try to solve (8.0.2) for $\Omega = B^n$ as

$$\mathrm{PI}\, f(x) = \frac{1 - |x|^2}{A_{n-1}} \int_{S^{n-1}} \frac{f(y)}{|x - y|^n} \, dS(y), \tag{8.0.7}$$

and verify that this works. We end §8.1 with some consequences of this Poisson integral formula, including the C^∞ nature of all harmonic functions, and a removable singularity theorem.

The formula (8.0.3) works because we know the eigenfunctions $e^{ik\theta}$ of the Laplace operator ∂_θ^2 on S^1, yielding the harmonic functions $r^k e^{ik\theta} = z^k$ and $r^k e^{-ik\theta} = \overline{z}^k$, $k \in \mathbb{N}$. This leads to the formula (8.0.6). The theory of spherical harmonics on S^{n-1} for $n \geq 3$ continues this connection, but the formulas tend to flow in the opposite direction, since we now have the Poisson integral formula (8.0.7) and need to discover what the spherical harmonics are. We take this up in §8.2, seeking harmonic functions on $\mathbb{R}^n$ of the form

$$u(r\omega) = \varphi(r) g(\omega), \quad r \in \mathbb{R}^+, \ \omega \in S^{n-1}. \tag{8.0.8}$$

This is done via the formula for the Laplace operator in spherical polar coordinates,

$$\Delta u(r\omega) = \partial_r^2 u + \frac{n-1}{r} \partial_r u + \frac{1}{r^2} \Delta_S u, \tag{8.0.9}$$

where Δ_S is the Laplace-Beltrami operator on S^{n-1}. We obtain

$$\begin{gathered}\varphi(r) = r^k, \quad \Delta_S g = -\lambda_k^2 g, \\ k \in \mathbb{Z}^+, \quad \lambda_k^2 = k^2 + (n-2)k.\end{gathered} \tag{8.0.10}$$

This gives rise to the spaces

$$V_k = \{g \in C^\infty(S^{n-1}) : \Delta_S g = -\lambda_k^2 g\}, \tag{8.0.11}$$

the eigenspaces of Δ_S. The fact that u in (8.0.8) is smooth on B^n is seen to imply that such u is a harmonic polynomial on $\mathbb{R}^n$, homogeneous of degree k, i.e., an element of $\mathcal{H}_k$. We have an isomorphism,

$$\tau : \mathcal{H}_k \xrightarrow{\approx} V_k, \quad \tau u = u\big|_{S^{n-1}}. \tag{8.0.12}$$

In light of this, there is an incisive result, Proposition 8.2.3, which says that, if $\mathcal{P}_k = \mathcal{P}_k(\mathbb{R}^n)$ is the space of polynomials on $\mathbb{R}^n$, homogeneous of degree k, then there is a direct sum decomposition

$$\mathcal{P}_k = \mathcal{H}_k \oplus |x|^2\mathcal{H}_{k-2} \oplus \cdots \oplus |x|^{2j}\mathcal{H}_{k-2j}, \tag{8.0.13}$$

with $k \in \{2j, 2j+1\}$. Consequences of this include the computation

$$\dim V_k = \binom{k+n-2}{k} + \binom{k+n-3}{k-1}, \tag{8.0.14}$$

and the algebraic result that

$$g_j \in V_j,\ g_k \in V_k \Longrightarrow g_j g_k \in \bigoplus_{\ell=|j-k|}^{j+k} V_\ell. \tag{8.0.15}$$

One consequence of (8.0.15) is that the Stone-Weierstrass theorem applies to show that the space of finite linear combinations of eigenfunctions of Δ_S is dense in $C(S^{n-1})$, a result that also follows from the general theory of the Laplace operator on a compact Riemannian manifold, via the theory of elliptic differential operators. As a consequence, we have that, if
(8.0.16)

$$E_k : L^2(S^{n-1}) \longrightarrow V_k \ \text{ is the orthogonal projection,} \quad S_N = \sum_{k=0}^N E_k,$$

then

$$f \in L^2(S^{n-1}) \Longrightarrow S_N f \to f \text{ in } L^2\text{-norm}, \tag{8.0.17}$$

as $N \to \infty$. In Proposition 8.2.12, we show that if

$$(-\Delta_S + 1)^m f \in L^2(S^{n-1}), \quad 2m > \frac{n-1}{2}, \tag{8.0.18}$$

then

$$S_N f \longrightarrow f, \ \text{ uniformly, on } S^{n-1}. \tag{8.0.19}$$

In §8.3 we continue our study of the relation between the Dirichlet problem on B^n and spherical functions. We show in Proposition 8.3.2 that, for all $f \in C(S^{n-1})$,

$$\mathrm{PI}\, f(r\omega) = \sum_{k=0}^{\infty} r^k E_k f(\omega). \tag{8.0.20}$$

Combining this with the Poisson integral formula (8.0.7) yields the identity

$$\sum_{k=0}^{\infty} r^k E_k f(\omega) = \frac{1-r^2}{A_{n-1}} \int\limits_{S^{n-1}} \frac{f(y)}{(1-2r\omega\cdot y+r^2)^{n/2}}\, dS(y), \tag{8.0.21}$$

which in turn yields

$$E_k f(\omega) = \int\limits_{S^{n-1}} E_k(\omega, y) f(y)\, dS(y), \tag{8.0.22}$$

with

$$E_k(\omega, y) = \frac{2\nu_k}{(n-2)A_{n-1}}\, C_k^{(n-2)/2}(\omega\cdot y), \quad \nu_k = k + \frac{n-2}{2}. \tag{8.0.23}$$

Here the functions $C_k^{\alpha}(t)$ are Gegenbauer polynomials, defined by (8.3.15)–(8.3.18). They are polynomials of degree k in t. A special case, arising when $n = 3$, is the class of Legendre polynomials,

$$P_k(t) = C_k^{1/2}(t). \tag{8.0.24}$$

We see that

$$\dim V_k = \int\limits_{S^{n-1}} E_k(y, y)\, dS(y), \tag{8.0.25}$$

and that inserting (8.0.23) produces a formula for $\dim V_k$ in terms of $C_k^{(n-2)/2}(1)$, leading to another proof of (8.0.14).

For $n \geq 3$, we define a zonal function on S^{n-1} as a function that is invariant under the action of $SO(n-1)$, acting on $\mathbb{R}^n$ by fixing the vector $e = (0, \dots, 0, 1)$. We denote by $\mathcal{Z}(S^{n-1})$ the class of zonal functions on S^{n-1}, and by $\mathcal{Z}_k = V_k \cap \mathcal{Z}(S^{n-1})$ the class of zonal harmonics. An example is

$$Z_k(\omega) = C_k^{(n-1)/2}(\omega\cdot e). \tag{8.0.26}$$

A key result of §8.4 is that, for each $k \in \mathbb{Z}^+$,

$$\mathcal{Z}_k = \mathrm{Span}(Z_k). \tag{8.0.27}$$

Equivalently, $\dim \mathcal{Z}_k = 1$. We normalize to define the spherical harmonic

$$Y_k^0(\omega) = \|Z_k\|_{L^2}^{-1} Z_k(\omega). \tag{8.0.28}$$

In §8.4 we also look at

$$\Pi f(\omega) = \int_{SO(n-1)} f(R^{-1}\omega)\, dR, \tag{8.0.29}$$

see that $\Pi : V_k \to V_k$, denote the restriction by Π_k, and show that

$$\Pi_k : V_k \longrightarrow \mathcal{Z}_k \ \text{ is an orthogonal projection.} \tag{8.0.30}$$

One consequence is that

$$f \in V_k,\ f \perp Y_k^0 \Longrightarrow f(e) = 0, \tag{8.0.31}$$

from which in turn we obtain a second demonstration that

$$|Y_k^0(e)|^2 = \frac{\dim V_k}{A_{n-1}}, \tag{8.0.32}$$

complementing a demonstration via (8.0.25).

Section 8.5 concentrates on the action of $SO(n)$ on functions on S^{n-1}, given by

$$\pi(R)f(\omega) = f(R^{-1}\omega). \tag{8.0.33}$$

This action commutes with Δ_S, so it maps each eigenspace V_k to itself, and we have representations

$$\pi_k : SO(n) \longrightarrow \mathcal{L}(V_k). \tag{8.0.34}$$

The first key result of §8.5, Proposition 8.5.1, says that, for each $k \in \mathbb{Z}^+$,

$$\pi_k \ \text{ is irreducible.} \tag{8.0.35}$$

The key ingredient in the proof is (8.0.27).

For $n = 3$, the dimension count $\dim V_k = 2k + 1$ implies that each irreducible representation of $SO(3)$ is equivalent to exactly one π_k, i.e., each irreducible representation of $SO(3)$ is contained, exactly once, in $L^2(S^2)$. For $n \ge 4$, this is no longer the case. We look at the irreducible representations of $SO(4)$ and describe precisely which ones are contained in $L^2(S^3)$. In general, we characterize those representations of $SO(n)$ contained in $L^2(S^{n-1})$ as *class one* representations.

The remainder of §8.5 is devoted to applying the structure of the irreducible representations of $SO(3)$, as developed in Chapter 4, and partially recalled here, to specifying the structure of the representations π_k in this case. The analysis leads to a specification of an orthonormal basis

$$Y_k^\ell, \quad -k \le \ell \le k, \tag{8.0.36}$$

of V_k, obtained from Y_k^0 by repeatedly applying a certain "ladder operator" L_+, for $\ell > 0$, and taking complex conjugates (for $\ell < 0$). To wit, we obtain for $1 \le \ell \le k$ (if $k \ge 1$)

$$Y_k^\ell(\omega) = \alpha_{k\ell}(\omega_1 + i\omega_2)^\ell P_k^{(\ell)}(\omega_3), \quad Y_k^{-\ell}(\omega) = \overline{Y_k^\ell(\omega)}, \tag{8.0.37}$$

with coefficients $\alpha_{k\ell}$ obtainable from (8.5.60).

In §8.7 we derive some formulas for the characters $\chi_k(g) = \operatorname{Tr} \pi_k(g)$ of the representations π_k of $SO(n)$ on V_k. We start with the identity

$$\sum_{k=0}^\infty r^k \pi_k(g) E_k f(x) = \operatorname{PI} f(rg^{-1}x), \tag{8.0.38}$$

together with the Poisson integral formula (8.0.7) and proceed to derive the formula

$$\sum_{k=0}^\infty r^k \chi_k(g) = \frac{1-r^2}{A_{n-1}} \int\limits_{S^{n-1}} \frac{dS(y)}{(1 - 2ry \cdot gy + r^2)^{n/2}}. \tag{8.0.39}$$

Taking $g = I$ leads to a third proof of the dimension formula (8.0.14).

This chapter ends with a couple of appendices. Appendix 8.A computes that

$$\dim \mathcal{P}_k(\mathbb{R}^n) = \binom{n+k-1}{k}, \tag{8.0.40}$$

which is used in the proof that (8.0.13) leads to (8.0.14). The derivation of (8.0.40) involves some legerdemain with multi-variable power series.

Appendix 8.B establishes that if M is a homogeneous space, on which a compact Lie group G acts transitively as a group of isometries, and if $V \subset C(M)$ is a finite-dimensional space invariant under such an action, then, for any orthonormal basis $\{u_k\}$ of V,

$$\sum_k |u_k(x)|^2 = \frac{\dim V}{\operatorname{Vol}(M)}, \quad \forall\, x \in M. \tag{8.0.41}$$

As a consequence,

$$\|f\|_{L^\infty}^2 \le \frac{\dim V}{\operatorname{Vol}(M)} \|f\|_{L^2}^2, \quad \forall\, f \in V. \tag{8.0.42}$$

These results, specialized to $V = V_k \subset L^2(S^{n-1})$, lead to both (8.0.18)–(8.0.19) and (8.0.32). We have seen fit to cast some results on spherical harmonics in a broader context, in Appendix 8.B. See Chapter 9 for further extensions of results of this chapter, particularly to analysis of functions on rank-one symmetric spaces.

8.1. Harmonic functions

The Laplace operator Δ is defined on functions on an open set $\Omega \subset \mathbb{R}^n$ by

$$\Delta u(x) = \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2}. \tag{8.1.1}$$

Given $u \in C^2(\Omega)$, we say

$$u \text{ is harmonic on } \Omega \Leftrightarrow \Delta u = 0 \text{ on } \Omega. \tag{8.1.2}$$

As seen in §A.8, we have for $u_g(x) = u(gx)$ that

$$g \in SO(n) \Longrightarrow \Delta u_g = (\Delta u)_g. \tag{8.1.3}$$

Section A.8 also shows that, in spherical polar coordinates

$$x = r\omega, \quad r \in (0, \infty),\ \omega \in S^{n-1}, \tag{8.1.4}$$

we have

$$\Delta u(r\omega) = \partial_r^2 u + \frac{n-1}{r}\partial_r u + \frac{1}{r^2}\Delta_S u, \tag{8.1.5}$$

where Δ_S is the Laplace-Beltrami operator on S^{n-1}. In particular,

$$u(x) = f(|x|) \Longrightarrow \Delta u(x) = f''(|x|) + \frac{n-1}{|x|} f'(|x|). \tag{8.1.6}$$

This equation allows us to classify all the radial functions that are harmonic. They must satisfy

$$f''(r) + \frac{n-1}{r} f'(r) = 0, \tag{8.1.7}$$

or, with $g(r) = f'(r)$,

$$r g'(r) + (n-1) g(r) = 0. \tag{8.1.8}$$

This is an Euler equation, whose general solution is

$$g(r) = A r^{1-n}. \tag{8.1.9}$$

Hence

$$\begin{aligned} f(r) = \frac{A}{2-n} r^{2-n} &+ B, \quad \text{if } \ n \geq 3, \\ A \log r &+ B, \quad \text{if } \ n = 2. \end{aligned} \tag{8.1.10}$$

Thus a radial harmonic function v has the form

$$\begin{aligned} v(x) = C|x|^{2-n} &+ B, \quad n \geq 3, \\ C \log |x| &+ B, \quad n = 2. \end{aligned} \tag{8.1.11}$$

The following important result is known as the *mean value property* of harmonic functions. This is often proved using Green's theorem, but we will get it from a symmetry argument.

Proposition 8.1.1. *If $\Omega \subset \mathbb{R}^n$ is open, $u \in C^2(\Omega)$ is harmonic, and $\overline{B_r(p)} \subset \Omega$, then*

$$u(p) = \operatorname{Avg}_{\partial B_r(p)} u. \tag{8.1.12}$$

Proof. Translating, we can arrange that $p = 0$, so $u \in C(\overline{B_r(0)})$ is harmonic on $B_r(0)$, and we claim that

$$u(0) = \operatorname{Avg}_{\partial B_r(0)} u. \tag{8.1.13}$$

Let us set

$$v(x) = \int\limits_{SO(n)} u(gx)\, dg. \tag{8.1.14}$$

Then v is a radial function, and

$$v \equiv \operatorname{Avg}_{\partial B_r(0)} u \quad \text{on} \quad \partial B_r(0). \tag{8.1.15}$$

By (8.1.3), v is harmonic on $B_r(0)$. Hence v satisfies (8.1.11). On the other hand, (8.1.14) implies v is bounded as $|x| \to 0$, so we conclude that $v(x) = B$, constant on $B_r(0)$. This constant must be $v|_{\partial B_r(0)} = \operatorname{Avg}_{\partial B_r(0)} u$. Meanwhile, $u(0) = v(0)$, so we have (8.1.13). $\square$

Here is a variant of the mean value property. We leave the deduction from (8.1.12) as an exercise to the reader.

Corollary 8.1.2. *In the setting of Proposition 8.1.1,*

$$u(p) = \operatorname{Avg}_{B_r(p)} u. \tag{8.1.16}$$

The mean value property leads to the following strong maximum principle.

Proposition 8.1.3. *Assume $\Omega \subset \mathbb{R}^n$ is open and connected. Let $u : \Omega \to \mathbb{R}$ be harmonic, and assume there exists $p \in \Omega$ such that*

$$u(p) \ge u(x), \quad \forall\, x \in \Omega. \tag{8.1.17}$$

Then u is constant.

Proof. Let $\mathcal{O} = \{x \in \Omega : u(x) = u(p)\}$. Clearly $\mathcal{O}$ is a closed subset of Ω, and $p \in \mathcal{O}$. Let $q \in \mathcal{O}$, so there exists $\varepsilon > 0$ such that $\overline{B_\varepsilon(q)} \subset \Omega$. Applying corollary 8.1.2 to $B_\varepsilon(q)$, we get $u(x) = u(q)$ for all $x \in B_\varepsilon(q)$, so $B_\varepsilon(q) \subset \mathcal{O}$, hence $\mathcal{O}$ is open. Since Ω is connected, we have $\mathcal{O} = \Omega$. $\square$

The following is a straightforward consequence of Proposition 8.1.3.

Corollary 8.1.4. *Assume $\Omega \subset \mathbb{R}^n$ is bounded and open. If $u : \overline{\Omega} \to \mathbb{C}$ is continuous on $\overline{\Omega}$ and harmonic on Ω, then*

$$\sup_{\Omega} |u(x)| = \sup_{\partial\Omega} |u(x)|. \tag{8.1.18}$$

Corollary 8.1.4 implies uniqueness for solutions to the Dirichlet problem, which is the following. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, and take $f \in C(\partial\Omega)$. The Dirichlet problem is to find $u \in C(\overline{\Omega}) \cap C^2(\Omega)$ such that

$$\Delta u = 0 \text{ on } \Omega, \quad u\big|_{\partial\Omega} = f. \tag{8.1.19}$$

This is treated in many PDE books, such as [**39**], Chapter 5. Here we restrict attention to $\Omega = B^n = \{x \in \mathbb{R}^n : |x| < 1\}$, and establish existence by producing an integral formula for the solution.

We start with the case $n = 2$, where B^n is the unit disk,

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}, \tag{8.1.20}$$

and seek u satisfying

$$\Delta u = 0 \text{ on } D, \quad u = f \text{ on } \partial D, \tag{8.1.21}$$

given $f \in C(\partial D)$. In a precursor to arguments that will arise in §8.2, we bring in the Fourier coefficients,

$$\hat{f}(k) = \frac{1}{2\pi} \int\limits_{\mathbb{T}} f(\theta) e^{-ik\theta} \, d\theta, \tag{8.1.22}$$

where we identify $\mathbb{R}^2$ and $\mathbb{C}$, and $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ with $\partial D = S^1$, via $\theta \mapsto e^{i\theta}$. We claim that the solution to ((8.1.21) is given by

$$u(re^{i\theta}) = \sum_k \hat{f}(k) r^{|k|} e^{ik\theta}. \tag{8.1.23}$$

Note that $re^{i\theta} = z \Rightarrow re^{-i\theta} = \overline{z}$, so (8.1.23) yields

$$u(re^{i\theta}) = \sum_{k=0}^{\infty} \hat{f}(k) z^k + \sum_{k=1}^{\infty} \hat{f}(-k) \overline{z}^k. \tag{8.1.24}$$

Since $|\hat{f}(k)| \le \sup |f|$ for all k, we see that each power series converges for $|z| < 1$, so u is the sum of a holomorphic function and a conjugate holomorphic function, hence is harmonic on D. To conclude that u solves (8.1.21), we establish the following.

Proposition 8.1.5. *If $f \in C(\mathbb{T})$ and u is given by* (8.1.23)*, then*

$$u(re^{i\theta}) \longrightarrow f(\theta), \quad \text{as } r \nearrow 1, \tag{8.1.25}$$

uniformly in θ.

Proof. Inserting the integral formula (8.1.22) for $\hat{f}(k)$, we have, for $r < 1$,

$$\begin{aligned} u(re^{i\theta}) &= \frac{1}{2\pi} \sum_k \int\limits_{\mathbb{T}} f(\varphi) e^{ik(\theta-\varphi)} \, d\varphi \\ &= \int_{-\pi}^{\pi} f(\theta - \varphi) p_r(\varphi) \, d\varphi, \end{aligned} \tag{8.1.26}$$

where

$$
\begin{aligned}
p_r(\varphi) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r^{|k|} e^{ik\varphi} \\
&= \frac{1}{2\pi} \Bigl[\sum_{k=0}^{\infty} r^k e^{ik\varphi} + \sum_{k=1}^{\infty} r^k e^{-ik\varphi} \Bigr] \\
&= \frac{1}{2\pi} \frac{1-r^2}{1-2r\cos\varphi + r^2},
\end{aligned}
\tag{8.1.27}
$$

the last identity obtained by summing two geometric series.

Let us examine $p_r(\varphi)$. It is clear that the numerator and denominator of the last fraction in (8.1.27) are positive, so $p_r(\varphi) > 0$ for each $r \in [0,1),\ \varphi \in \mathbb{T}$. As $r \nearrow 1$, the numerator tends to 0, and the denominator tends to a nonzero limit, except at $\varphi = 0$. Since we have

$$
\int_{-\pi}^{\pi} p_r(\varphi)\, d\varphi = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_k r^{|k|} e^{ik\varphi}\, d\varphi = 1,
\tag{8.1.28}
$$

we see that, for r close to 1, $p_r(\varphi)$, as a function of φ, is highly peaked near $\varphi = 0$, and small elsewhere. It follows that, for each $\delta \in (0,\pi)$,

$$
\int\limits_{|\varphi|\le\delta} p_r(\varphi)\, d\varphi = 1 - \varepsilon(r,\delta),
\tag{8.1.29}
$$

with $\varepsilon(r,\delta) \to 0$ as $r \nearrow 1$. Now we break (8.1.26) into three pieces:

$$
\begin{aligned}
u(re^{i\theta}) = f(\theta) &\int_{-\delta}^{\delta} p_r(\varphi)\, d\varphi \\
&+ \int_{-\delta}^{\delta} \bigl[f(\theta-\varphi) - f(\theta) \bigr] p_r(\varphi)\, d\varphi \\
&+ \int\limits_{\delta\le|\varphi|\le\pi} f(\theta-\varphi) p_r(\varphi)\, d\varphi \\
= I + II &+ III.
\end{aligned}
\tag{8.1.30}
$$

We have

$$
\begin{aligned}
I &= f(\theta)(1-\varepsilon(r,\delta)), \\
|II| &\le \sup_{|\varphi|\le\delta} |f(\theta-\varphi) - f(\theta)|, \\
|III| &\le \varepsilon(r,\delta) \sup |f|.
\end{aligned}
\tag{8.1.31}
$$

These estimates yield (8.1.25). ∎

The integral formula defining the solution to the Dirichlet problem (8.1.21) is called the Poisson integral. We rewrite it as

$$u(re^{i\theta}) = \frac{1-r^2}{2\pi} \int_{\mathbb{T}} \frac{f(\varphi)}{1-2r\cos(\theta-\varphi)+r^2}\, d\varphi. \tag{8.1.32}$$

A change of variable gives, for $x \in \mathbb{R}^2$, $|x| < 1$,

$$u(x) = \frac{1-|x|^2}{2\pi} \int_{S^1} \frac{f(y)}{|x-y|^2}\, ds(y), \tag{8.1.33}$$

where $ds(y)$ denotes arclength.

Moving from dimension 2 to dimension $n \geq 3$, we are motivated to try a formula for the solution to the Dirichlet problem

$$\Delta u = 0 \text{ on } B^n, \quad u\big|_{S^{n-1}} = f, \tag{8.1.34}$$

on the unit ball $B^n \subset \mathbb{R}^n$, of the form

$$u(x) = C_n(1-|x|^2) \int_{S^{n-1}} \frac{f(y)}{|x-y|^n}\, dS(y). \tag{8.1.35}$$

We will show that this works, and along the way calculate the constant C_n. First we will show that, for each $f \in C(S^{n-1})$, the function u is harmonic on B^n. This is a consequence of the following.

Lemma 8.1.6. *For a given $y \in S^{n-1}$ (i.e., $|y| = 1$), set*

$$v(x) = (1-|x|^2)|x-y|^{-n}. \tag{8.1.36}$$

Then v is harmonic on $\mathbb{R}^n \setminus \{y\}$.

Proof. It suffices to show that $w(x) = v(x+y)$ is harmonic on $\mathbb{R}^n \setminus 0$. Since $1-|x+y|^2 = -(2x\cdot y + |x|^2)$ provided $|y| = 1$, we have

$$-w(x) = 2(y\cdot x)|x|^{-n} + |x|^{2-n}. \tag{8.1.37}$$

That $|x|^{2-n}$ is harmonic on $\mathbb{R}^n \setminus 0$ has already arisen in (8.1.11). Now applying $\partial/\partial x_j$ to a smooth harmonic function on an open set in $\mathbb{R}^n$ gives another, so the following are harmonic on $\mathbb{R}^n \setminus 0$:

$$w_j(x) = \frac{\partial}{\partial x_j}|x|^{2-n} = (2-n)x_j|x|^{-n}. \tag{8.1.38}$$

For $n = 2$ we take instead

$$\frac{\partial}{\partial x_j} \log|x| = x_j|x|^{-2}. \tag{8.1.39}$$

Thus the first term on the right side of (8.1.37) is a linear combination of these functions, so the lemma is proved. $\square$

To justify (8.1.35), it remains to show that if u is given by this formula and C_n is chosen correctly, then $u = f$ on S^{n-1}. Note that if we write $x = r\omega$, $\omega \in S^{n-1}$, then (8.1.35) yields

$$u(r\omega) = \int_{S^{n-1}} p_r(\omega, y) f(y)\, dS(y), \tag{8.1.40}$$

where

$$p_r(\omega, y) = C_n(1 - r^2)|r\omega - y|^{-n}. \tag{8.1.41}$$

It is clear that

$$p_r(\omega, y) \longrightarrow 0, \quad \text{as} \quad r \nearrow 1, \quad \text{if} \quad \omega \neq y, \tag{8.1.42}$$

the convergence being uniform on each compact subset of $\{(\omega, y) \in S^{n-1} \times S^{n-1} : \omega \neq y\}$. We claim that

$$\int_{S^{n-1}} p_r(\omega, y)\, dS(y) = C_n', \tag{8.1.43}$$

a constant independent of r and ω. The independence of ω follows by rotational symmetry. Thus we can integrate over ω. But Lemma 8.1.6 implies that

$$p_r(x, y) = C_n(1 - |rx|^2)|rx - y|^{-n} \tag{8.1.44}$$

is harmonic in x, for $|x| < 1/r$, so the mean value property for harmonic functions gives

$$\frac{1}{A_{n-1}} \int_{S^{n-1}} p_r(\omega, y)\, dS(\omega) = C_n, \tag{8.1.45}$$

for all $r < 1$, $y \in S^{n-1}$. This implies (8.1.43), with $C_n' = C_n A_{n-1}$.

Thus, in view of (8.1.42), $p_r(\omega, y)$ is highly peaked near $\omega = y \in S^{n-1}$, as $r \nearrow 1$. Hence an argument parallel to that used in the proof of Proposition 8.1.5 showd that the limit of (8.1.40) as $r \nearrow 1$ is equal to $C_n A_{n-1} f(\omega)$, for each $f \in C(S^{n-1})$. This justifies the formula (8.1.35) and fixes the constant: $C_n = 1/A_{n-1}$. We have proved most of the following.

Proposition 8.1.7. *Given $f \in C(S^{n-1})$, the solution in $C(\overline{B}^n) \cap C^2(B^n)$ to the Dirichlet problem* (8.1.34) *is given by the Poisson integral*

$$u(x) = \frac{1 - |x|^2}{A_{n-1}} \int_{S^{n-1}} \frac{f(y)}{|x - y|^n}\, dS(y). \tag{8.1.46}$$

Proof. It remains to establish uniqueness. In fact, the difference v of two such solutions would be harmonic on B^n and vanish on ∂B^n. Hence the maximum principle, Corollary 8.1.4, yields $v \equiv 0$ on B^n. □

Another way to write the conclusion (8.1.46) of Proposition 8.1.7 is

$$u(r\omega) = \frac{1-r^2}{A_{n-1}} \int\limits_{S^{n-1}} \frac{f(y)}{(1-2r\omega\cdot y+r^2)^{n/2}}\, dS(y), \tag{8.1.47}$$

for $\omega \in S^{n-1},\ 0 \le r < 1$.

Having Proposition 8.1.7, we can apply translations and scaling to solve the Dirichlet problem

$$\Delta u = 0 \text{ on } B_r(p), \quad u = f \text{ on } \partial B_r(p), \tag{8.1.48}$$

on arbitrary balls $B_r(p) = \{x \in \mathbb{R}^n : |x-p| < r\}$, with analogous Poisson integral formulas. We denote the solution to (8.1.48) by

$$u = \mathrm{PI}\, f. \tag{8.1.49}$$

We discuss some applications of Proposition 8.1.7 to the behavior of harmonic functions. Here is one.

Proposition 8.1.8. *Let $\Omega \subset \mathbb{R}^n$ be open. Assume $u \in C^2(\Omega)$ is harmonic. Then*

$$u \in C^\infty(\Omega), \tag{8.1.50}$$

and each function $u_\alpha = \partial^\alpha u$ is harmonic on Ω.

Proof. Let $B_r(p)$ be a ball whose closure is contained in Ω. Translating and scaling, we may as well suppose the ball is B^n. Apply (8.1.46), with $f = u|_{\partial B^n}$. Then apply ∂^α to the right side of (8.1.46). □

Here is an offshoot.

Proposition 8.1.9. *Let $\Omega \subset \mathbb{R}^n$ be open. Assume $u_\nu \in C^\infty(\Omega)$ are harmonic and $u_\nu \to u$, uniformly on compact subsets of Ω. Then*

$$u \in C^\infty(\Omega),\ \partial^\alpha u_\nu \to \partial^\alpha u \textit{ uniformly on compact subsets of } \Omega, \tag{8.1.51}$$

and u is harmonic on Ω.

Proof. It suffices to show that if $x_0 \in \overline{B_R(x_0)} \subset \Omega$ and $0 < \rho < R$, then

$$\begin{aligned} &u_\nu \to u \text{ uniformly on } \partial B_R(x_0) \\ &\Rightarrow \partial^\alpha u_\nu \text{ uniformly Cauchy on } B_\rho(x_0). \end{aligned} \tag{8.1.52}$$

Translating and dilating, we can assume $x_0 = 0$ and $R = 1$, and apply Proposition 8.1.7, to write

$$u_\mu(x) - u_\nu(x) = \frac{1-|x|^2}{A_{n-1}} \int\limits_{S^{n-1}} \frac{u_\mu(y)-u_\nu(y)}{|x-y|^n}\, dS(y). \tag{8.1.53}$$

We can apply ∂^α to the right side to get (8.1.52). □

The next result is a removable singularity theorem. Take $B = B_1(0)$.

Proposition 8.1.10. *Assume $u \in C^2(B \setminus 0) \cap C(\overline{B} \setminus 0)$ is harmonic on B and bounded, i.e., there exists $M < \infty$ such that*

$$|u(x)| \le M, \quad \forall\, x \in \overline{B} \setminus 0. \tag{8.1.54}$$

Then u can be extended (in a unique fashion) to be harmonic on all of B.

Proof. Let $f = u|_{\partial B} \in C(\partial B)$, and set

$$v = \mathrm{PI}\, f, \quad v \in C(\overline{B}) \cap C^2(B). \tag{8.1.55}$$

We claim $v = u$ on $B \setminus 0$. To see this, consider $w = u - v$ on $B \setminus 0$. We have $w \in C(\overline{B} \setminus 0) \cap C^2(B \setminus 0)$, $\Delta w = 0$ on $B \setminus 0$, and $w = 0$ on ∂B. Also $|w| \le 2M$ on $\overline{B} \setminus 0$. We claim this implies $w \equiv 0$. To prove this, we can assume w is real valued.Now bring in the function $H \in C(\overline{B} \setminus 0) \cap C^2(B \setminus 0)$, given by

$$\begin{aligned} H(x) &= |x|^{2-n} - 1, \quad \text{if } n \ge 3, \\ &\log \frac{1}{|x|}, \quad \text{if } n = 2. \end{aligned} \tag{8.1.56}$$

We see that H is harmonic on $B \setminus 0$, $H \ge 0$ on $B \setminus 0$, $H = 0$ on ∂B, and $H(x) \to +\infty$ as $x \to 0$. Hence, for each $\varepsilon > 0$, there exists $\delta_0 > 0$ such that

$$\varepsilon H - w \ge 0 \text{ on } \partial B_\delta(0), \quad \forall\, \delta \in (0, \delta_0]. \tag{8.1.57}$$

The maximum principle implies that

$$\varepsilon H - w \ge 0 \tag{8.1.58}$$

on $B \setminus B_\delta(0)$. Taking $\delta \searrow 0$ yields (8.1.58) on $B \setminus 0$. Then taking $\varepsilon \searrow 0$ yields

$$w \le 0 \text{ on } B \setminus 0. \tag{8.1.59}$$

A similar argument gives $w \ge 0$ on $B \setminus 0$, hence $w \equiv 0$, and the proof is complete. $\square$

8.2. Spherical harmonics

The Laplace-Beltrami operator Δ_S on S^{n-1} gives rise to an orthonormal basis of $L^2(S^{n-1})$ consisting of eigenfunctions of Δ_S. These functions are called spherical harmonics. Our approach to studying these spherical harmonics will exploit the connection they have with the Dirichlet problem on the ball,

$$\Delta u = 0 \text{ on } B^n, \quad u = f \text{ on } S^{n-1}. \tag{8.2.1}$$

This connection was previewed (for $n = 2$) in (8.1.21)–(8.1.27). Use of (8.1.22) derives from the fact that

$$\{e^{ik\theta} : k \in \mathbb{Z}\} \text{ is an orthonormal basis of } L^2(\mathbb{T}, d\theta/2\pi), \tag{8.2.2}$$

consisting of eigenfunctions of the Laplace operator on $\mathbb{T}$, i.e., $\partial_\theta^2 e^{ik\theta} = -k^2 e^{ik\theta}$. In §8.1 we used this to find a formula for the Poisson integral.

For $n \geq 3$, we go in the opposite direction. We have in hand the formula (8.1.46) for the Poisson integral, and we will use this as a tool to investigate the eigenfunctions of Δ_S. To start, we look for harmonic functions on B^n in the form

$$u(r\omega) = \varphi(r)g(\omega). \tag{8.2.3}$$

To get this, we use the formula (8.1.5) for Δ in spherical polar coordinates. In case u is given by (8.2.3), we have

$$\Delta u(r\omega) = \Big\{\varphi''(r) + \frac{n-1}{r}\varphi'(r)\Big\}g(\omega) + \frac{1}{r^2}\Delta_S g(\omega). \tag{8.2.4}$$

Thus (8.2.3) defines a harmonic function on $B^n \setminus 0$ if and only if there is a constant μ such that

$$\Delta_S g = \mu g \text{ on } S^{n-1}, \tag{8.2.5}$$

and

$$\varphi''(r) + \frac{n-1}{r}\varphi'(r) + \frac{\mu}{r^2}\varphi(r) = 0, \tag{8.2.6}$$

for $0 < r < 1$. Recall that Proposition 8.1.8 implies the harmonic function u is C^∞ on B^n. Hence, in (8.2.3), we must have $g \in C^\infty(S^{n-1})$. If g satisfies (8.2.5), we say g is an eigenfunction of Δ_S, with eigenvalue μ. Note that

$$\mu \int\limits_{S^{n-1}} |g|^2 \, dS = \int\limits_{S^{n-1}} (\Delta_S g)\overline{g} \, dS = -\int\limits_{S^{n-1}} |\nabla g|^2 \, dS, \tag{8.2.7}$$

so $\mu \leq 0$. Say $\mu = -\lambda^2$. Now the equation (8.2.6) is an Euler equation, whose solutions are linear combinations of

$$\varphi_\pm(r) = r^{k_\pm}, \tag{8.2.8}$$

where $k = k_\pm$ satisfies the equation

$$k(k-1) + (n-1)k - \lambda^2 = 0, \tag{8.2.9}$$

with roots

$$k_\pm = -\frac{n-2}{2} \pm \frac{1}{2}\sqrt{(n-2)^2 + 4\lambda^2}. \tag{8.2.10}$$

The smoothness of u on B^n requires that the exponent in (8.2.8) be positive, so we need

$$\varphi(r) = r^k, \quad k = -\frac{n-2}{2} + \frac{1}{2}\sqrt{(n-2)^2 + 4\lambda^2}. \tag{8.2.11}$$

Furthermore, such smoothness requires

$$k \in \mathbb{Z}^+ = \{0, 1, 2, 3, \dots\}. \tag{8.2.12}$$

Then (8.2.10) yields the eigenvalue $\mu = -\lambda_k^2$, with

$$\begin{aligned} \lambda_k^2 &= k^2 + (n-2)k \\ &= \Big(k + \frac{n-2}{2}\Big)^2 - \Big(\frac{n-2}{2}\Big)^2. \end{aligned} \tag{8.2.13}$$

Let us set

$$V_k = \{g \in C^\infty(S^{n-1}) : \Delta_S g = -\lambda_k^2 g\}. \tag{8.2.14}$$

We see that if $k \in \mathbb{Z}^+$ and $g \in V_k$, then $u(r\omega) = r^k g(\omega)$ is harmonic on $B_n \setminus 0$ and bounded. It follows from Proposition 8.1.10 that $\{0\}$ is a removable singularity, and u extends to be harmonic on all of B^n, and furthermore $u \in C^\infty(B^n)$. It is homogeneous of degree k, so $\partial_x^\alpha u(x)$ is homogeneous of degree $k - |\alpha|$. In particular, it is homogeneous of degree 0 for $|\alpha| = k$. Being smooth, it must be constant, so

$$|\alpha| = k \Longrightarrow \partial_x^\alpha u = c_\alpha, \quad \text{const.} \tag{8.2.15}$$

Hence $u(x)$ is a *polynomial* in x, homogeneous of degree k. Let us set

$$\begin{aligned} \mathcal{H}_k = {} & \text{space of harmonic polynomials on } \mathbb{R}^n \\ & \text{homogeneous of degree } k. \end{aligned} \tag{8.2.16}$$

The polynomials can have complex coefficients. If we need to specify n, we write $\mathcal{H}_k(\mathbb{R}^n)$. We have the following.

Proposition 8.2.1. *The map*

$$\tau : \mathcal{H}_k \longrightarrow V_k, \quad \tau u = u\big|_{S^{n-1}}, \tag{8.2.17}$$

is an isomorphism.

Next, we have the following important orthogonality result.

Proposition 8.2.2. *Assume* $g_j \in V_j,\ g_k \in V_k,\ j \neq k$. *Then*

$$(g_j, g_k)_{L^2(S^{n-1})} = \int_{S^{n-1}} g_j \overline{g_k}\, dS = 0. \tag{8.2.18}$$

Proof. For such g_j, g_k,

$$\begin{aligned} -\lambda_j^2 (g_j, g_k)_{L^2} &= (\Delta_S g_j, g_k)_{L^2} \\ &= (g_j, \Delta_S g_k)_{L^2} \\ &= -\lambda_k^2 (g_j, g_k)_{L^2}. \end{aligned} \tag{8.2.19}$$

since $j \neq k \Rightarrow \lambda_j \neq \lambda_k$, we have (8.2.18). ☐

The following result provides valuable information on the spaces $\mathcal{H}_k$. Let

$$\mathcal{P}_k = \text{ space of polynomials on } \mathbb{R}^n, \text{ homogeneous of degree } k. \tag{8.2.20}$$

As in (8.2.16), we allow complex coefficients, and if we need to specify n, we write $\mathcal{P}_k(\mathbb{R}^n)$.

Proposition 8.2.3. *For all* $k \in \mathbb{Z}^+$, *we have the direct sum decomposition*

$$\mathcal{P}_k = \mathcal{H}_k \oplus |x|^2 \mathcal{H}_{k-2} \oplus \cdots \oplus |x|^{2j} \mathcal{H}_{k-2j}, \tag{8.2.21}$$

with $k \in \{2j, 2j+1\}$.

Proof. By Proposition 8.2.2, the various summands on the right side of (8.2.21) are mutually orthogonal with respect to the L^2-inner product on S^{n-1}, so the sum on the right is direct. It remains to do a dimension count.

We do this by induction on k. Note that $\mathcal{P}_0 = \mathcal{H}_0$ and $\mathcal{P}_1 = \mathcal{H}_1$, so (8.2.21) is clear for $k = 0, 1$. Now assume the analogue of (8.2.21) holds for $\mathcal{P}_\ell$, for all $\ell < k$. Given this, the right side of (8.2.21) is equal to

$$\mathcal{H}_k \oplus |x|^2 \mathcal{P}_{k-2} = \mathcal{Q}_k, \tag{8.2.22}$$

and we need to show this has the same dimension as $\mathcal{P}_k$. The direct sum condition in (8.2.22) implies

$$\dim \mathcal{H}_k + \dim \mathcal{P}_{k-2} = \dim \mathcal{Q}_k \le \dim \mathcal{P}_k. \tag{8.2.23}$$

Now consider

$$\Delta : \mathcal{P}_k \longrightarrow \mathcal{P}_{k-2}. \tag{8.2.24}$$

The null space is $\mathcal{N}(\Delta) = \mathcal{H}_k$, so the fundamental theorem of linear algebra implies

$$\begin{aligned} \dim \mathcal{P}_k &= \dim \mathcal{N}(\Delta) + \dim \mathcal{R}(\Delta) \\ &\le \dim \mathcal{H}_k + \dim \mathcal{P}_{k-2}. \end{aligned} \tag{8.2.25}$$

Comparison with (8.2.23) gives $\dim \mathcal{P}_k = \dim \mathcal{Q}_k$ and finishes the proof. (It also shows that Δ in (8.2.24) is surjective.) $\square$

We can apply Proposition 8.2.3, and its corollary

$$\mathcal{P}_k = \mathcal{H}_k \oplus |x|^2 \mathcal{P}_{k-2}, \tag{8.2.26}$$

to compute $\dim \mathcal{H}_k$ (hence $\dim V_k$). Let us set

$$d_k(n) = \dim \mathcal{P}_k(\mathbb{R}^n). \tag{8.2.27}$$

Then (8.2.26) implies

$$\dim \mathcal{H}_k = d_k(n) - d_{k-2}(n). \tag{8.2.28}$$

As shown in Appendix 8.A,

$$d_k(n) = \binom{n+k-1}{k}. \tag{8.2.29}$$

Hence we have $\dim \mathcal{H}_k$ as a difference of two binomial coefficients. We prefer to express $\dim \mathcal{H}_k$ as a *sum* of two binomial coefficients, so we proceed. Let us set

$$\mathcal{P}^k(\mathbb{R}^n) = \text{ space of polynomials on } \mathbb{R}^n \text{ of degree} \le k. \tag{8.2.30}$$

We see that

$$d_k(n) = \dim \mathcal{P}^k(\mathbb{R}^{n-1}) = d_k(n-1) + d_{k-1}(n-1) + \cdots + d_0(n-1), \tag{8.2.31}$$

and similarly

$$d_{k-1}(n) = d_{k-1}(n-1) + \cdots + d_0(n-1), \tag{8.2.32}$$

hence

$$d_k(n) - d_{k-1}(n) = d_k(n-1). \tag{8.2.33}$$

Similarly

$$d_{k-1}(n) - d_{k-2}(n) = d_{k-1}(n-1), \tag{8.2.34}$$

so

$$d_k(n) - d_{k-2}(n) = d_k(n-1) + d_{k-1}(n-1). \tag{8.2.35}$$

Hence (8.2.28) gives

$$\dim \mathcal{H}_k = d_k(n-1) + d_{k-1}(n-1). \tag{8.2.36}$$

Using the formula (8.2.29), we record our conclusion as follows.

Proposition 8.2.4. *On S^{n-1},*

$$\dim V_k = \binom{k+n-2}{k} + \binom{k+n-3}{k-1}. \tag{8.2.37}$$

For example,

$$\text{On } S^2, \ \dim V_k = 2k+1. \tag{8.2.38}$$

Returning from dimension counts to other consequences of Proposition 8.2.3, we have the following important algebraic result.

Proposition 8.2.5. *Given $g_j \in V_j$ and $g_k \in V_k$, the product satisfies*

$$g_j g_k \in \bigoplus_{\ell=0}^{j+k} V_\ell. \tag{8.2.39}$$

Proof. The functions g_j and g_k extend to elements of $\mathcal{P}_j$ and $\mathcal{P}_k$, so $g_j g_k$ extends to an element of $\mathcal{P}_{j+k}$. Then (8.2.39) follows from (8.2.21), applied to $\mathcal{P}_{j+k}$. □

Before we discuss our main use of Proposition 8.2.5, we linger to savor the following refinement.

Proposition 8.2.6. *In the setting of Proposition 8.2.5, actually*

$$g_j g_k \in \bigoplus_{\ell=|j-k|}^{j+k} V_\ell. \tag{8.2.40}$$

Proof. Suppose $0 \le j < k$ and $m < k - j$, so $k > j + m$. We need to show that

$$f \in V_m \Longrightarrow f \perp g_j g_k. \tag{8.2.41}$$

This follows from the implications

$$\begin{aligned} f \in V_m, \ g_j \in V_j &\Longrightarrow f g_j \in \bigoplus_{\ell=0}^{j+m} V_\ell \\ &\Longrightarrow f g_j \perp V_k, \end{aligned} \tag{8.2.42}$$

the first implication holding by Proposition 8.2.5. □

Proposition 8.2.5 implies the following important density result.

Proposition 8.2.7. *The space*

$$\mathcal{V} = \bigcup_{\ell>0} \bigoplus_{k=0}^{\ell} V_k \tag{8.2.43}$$

of finite linear combinations of eigenfunctions of Δ_S is dense in $C(S^{n-1})$.

Proof. We see that (8.2.43) is an algebra of continuous functions on S^{n-1} Clearly it separates points (since V_1 does that) and it is closed under complex conjugates, so the Stone-Weierstrass theorem applies, to yield denseness in $C(S^{n-1})$. ☐

REMARK. General results from PDE theory (cf. [**39**], Chapters 5 and 8) also yield this denseness, in the more general setting of the eigenfunctions of the Laplace-Beltrami operator on a general compact Riemannian manifold. However, the argument given above is a bit more elementary, and in any case the special structure behind (8.2.39) and its refinement (8.2.40) are of great intrinsic interest.

The following is a convenient restatement of Proposition 8.2.7.

Corollary 8.2.8. *For each $k \in \mathbb{Z}^+$, let*

$$\{Y_k^\ell : \ell \in \Sigma_k\} \tag{8.2.44}$$

be an orthonormal basis of V_k, where Σ_k is an index set, of cardinality $\dim V_k$. *Then*

$$\{Y_k^\ell : k \geq 0, \ell \in \Sigma_k\} \tag{8.2.45}$$

is an orthonormal set of functions on S^{n-1} whose linear span is dense in $C(S^{n-1})$, hence an orthonormal basis of $L^2(S^{n-1})$.

Expansions of functions in terms of such an orthonormal basis are called spherical harmonic expansions. In detail, given $f \in L^2(S^{n-1})$, set

$$\hat{f}(k,\ell) = \int_{S^{n-1}} f(y)\overline{Y_k^\ell(y)}\, dS(y). \tag{8.2.46}$$

Then set

$$E_k f(y) = \sum_{\ell \in \Sigma_k} \hat{f}(k,\ell) Y_k^\ell(y), \tag{8.2.47}$$

yielding

$$E_k : L^2(S^{n-1}) \longrightarrow V_k. \tag{8.2.48}$$

The map E_k is an orthogonal projection of $L^2(S^{n-1})$ onto V_k, satisfying

$$\begin{aligned} &E_k f = f, \quad \forall\, f \in V_k, \\ &f - E_k f \perp V_k, \quad \forall\, f \in L^2(S^{n-1}), \end{aligned} \tag{8.2.49}$$

in the sense that

$$(f - E_k f, g)_{L^2} = 0, \quad \forall\, g \in V_k. \tag{8.2.50}$$

The properties (8.2.48)–(8.2.49) uniquely characterize E_k. As such, E_k is independent of the choice of orthonormal basis of V_k. We set

$$\begin{aligned} S_N f &= \sum_{k=0}^{N} E_k f \\ &= \sum_{k=0}^{N} \sum_{\ell\in\Sigma_k} \hat{f}(k,\ell) Y_k^\ell. \end{aligned} \tag{8.2.51}$$

Given Corollary 8.2.8, standard Hilbert space theory yields the following.

Proposition 8.2.9. *We have*

$$\begin{gathered} f \in L^2(S^{n-1}) \Longrightarrow S_N f \to f \ \text{ in } L^2\text{-norm, and} \\ \sum_{k=0}^{\infty} \sum_{\ell\in\Sigma_k} |\hat{f}(k,\ell)|^2 = \|f\|_{L^2}^2. \end{gathered} \tag{8.2.52}$$

It is of interest to give conditions on f that imply $S_N f \to f$ uniformly on S^{n-1}. This study is complicated by the fact that, for $n \ge 3$, the eigenfunctions Y_k^ℓ are not uniformly bounded (as we will see later). Somewhat compensating for this is the following interesting result.

Proposition 8.2.10. *For each $k \in \mathbb{Z}^+$, if $\{Y_k^\ell : \ell \in \Sigma_k\}$ is an orthonormal basis of V_k, then*

$$\sum_{\ell\in\Sigma_k} |Y_k^\ell(\omega)|^2 = \frac{\dim V_k}{A_{n-1}}, \quad \forall\, \omega \in S^{n-1}. \tag{8.2.53}$$

To prove this, we bring in the action of $SO(n)$ on $L^2(S^{n-1})$,

$$\pi(R)f(x) = f(R^{-1}x), \quad R \in SO(n). \tag{8.2.54}$$

Since R acts on S^{n-1} as an isometry, this action commutes with Δ_S, so $\pi(R) : V_k \to V_k$ for each k. We set

$$\pi_k(R) : V_k \longrightarrow V_k, \quad \pi_k(R) = \pi(R)\Big|_{V_k}. \tag{8.2.55}$$

Now $SO(n)$ acts transitively on S^{n-1}, so S^{n-1} is a homogeneous space. It turns out that Proposition 8.2.10 is a special case of a general result about finite dimensional subspaces of $C(M)$ invariant under the action of a transitive isometry group of M. See Proposition 8.B.1.

A companion to Proposition 8.B.1 is Corollary 8.B.2, which yields the following.

Corollary 8.2.11. *In the setting of Proposition 8.2.10,*

$$\|f\|_{L^\infty} \le \Big(\frac{\dim V_k}{A_{n-1}}\Big)^{1/2} \|f\|_{L^2}, \quad \forall\, f \in V_k. \tag{8.2.56}$$

Consequently,

$$\|E_k f\|_{L^\infty} \le \Big(\frac{\dim V_k}{A_{n-1}}\Big)^{1/2} \|E_k f\|_{L^2}, \quad \forall\, f \in L^2(S^{n-1}). \tag{8.2.57}$$

Let us set

$$D_k = \dim V_k, \quad L = 1 - \Delta_S, \quad \langle \lambda_k \rangle = (1 + \lambda_k^2)^{1/2}. \tag{8.2.58}$$

Then, for $m \in \mathbb{R}$,

$$E_k L^m f = L^m E_k f = (1 + \lambda_k^2)^m E_k f = \langle \lambda_k \rangle^{2m} E_k f. \tag{8.2.59}$$

We deduce from (8.2.57) that

$$A_{n-1}^{1/2} \|E_k f\|_{L^\infty} \le D_k^{1/2} \langle \lambda_k \rangle^{-2m} \|E_k L^m f\|_{L^2}. \tag{8.2.60}$$

Using this estimate, we will prove the following.

Proposition 8.2.12. *Let $f \in L^2(S^{n-1})$. Assume*

$$L^m f \in L^2(S^{n-1}), \quad 2m > \frac{n-1}{2}. \tag{8.2.61}$$

Then

$$\sum_k \|E_k f\|_{L^\infty} \le C \|L^m f\|_{L^2}. \tag{8.2.62}$$

Hence, as $N \to \infty$,

$$S_N f \longrightarrow f \quad \text{uniformly on} \quad S^{n-1}. \tag{8.2.63}$$

Proof. We apply (8.2.60) and Cauchy's inequality to write

$$\begin{aligned} A_{n-1}^{1/2} \sum_k \|E_k f\|_{L^\infty} &\le \sum_k D_k^{1/2} \langle \lambda_k \rangle^{-2m} \|E_k L^m f\|_{L^2} \\ &\le \Big(\sum_k D_k \langle \lambda_k \rangle^{-4m}\Big)^{1/2} \Big(\sum_k \|E_k L^m f\|_{L^2}^2\Big)^{1/2} \\ &= B_m \|L^m f\|_{L^2}, \end{aligned} \tag{8.2.64}$$

where

$$B_m^2 = \sum_k D_k \langle \lambda_k \rangle^{-4m}. \tag{8.2.65}$$

Recall that λ_k^2 is given by (8.2.13) and D_k by (8.2.37). It follows that

$$\langle \lambda_k \rangle^{-4m} D_k \le c_n (1 + k)^{-4m+n-2}. \tag{8.2.66}$$

Hence

$$4m > n - 1 \Longrightarrow B_m < \infty, \tag{8.2.67}$$

and we have (8.2.62). Finally,

$$\begin{aligned}\|f - S_N f\|_{L^\infty} &\le \sum_{k>N} \|E_k(f - S_N f)\|_{L^\infty} \\ &\le C\|(I - S_N)L^m f\|_{L^2},\end{aligned} \tag{8.2.68}$$

and we have (8.2.63). $\square$

8.3. The Poisson integral and spherical harmonic formulas

We return to the connection between the Dirichlet problem and spherical harmonic expansions, and examine

$$\begin{aligned} u(r\omega) &= \sum_{k=0}^{\infty} r^k E_k f(\omega) \\ &= \sum_{k=0}^{\infty} \sum_{\ell\in\Sigma_k} \hat{f}(k,\ell) r^k Y_k^\ell(\omega). \end{aligned} \tag{8.3.1}$$

Recall that each term $r^k Y_k^\ell(\omega) = h_k^\ell(r\omega)$ is a harmonic polynomial (homogeneous of degree k) on $\mathbb{R}^n$. We have from (8.2.37) and (8.2.57) that, for all $\omega \in S^{n-1}$,

$$\begin{aligned} &\sum_{k=0}^{\infty} r^k |E_k f(\omega)| \\ &\le c_n \sum_{k=0}^{\infty} (1+k)^{(n-2)/2} r^k \|E_k f\|_{L^2} \\ &\le c_n \Bigl(\sum_{k=0} (1+k)^{n-2} r^{2k}\Bigr)^{1/2} \|f\|_{L^2}, \end{aligned} \tag{8.3.2}$$

the last estimate by Cauchy's inequality. It follows that whenever $f \in L^2(S^{n-1})$, the series (8.3.1) converges uniformly on all balls $\overline{B}_R = \{x \in \mathbb{R}^n : |x| \le R\}$, for each $R < 1$, to a function $u \in C(B^n)$. Of course, each partial sum

$$\mathcal{S}_N f(r\omega) = \sum_{k=0}^{N} r^k E_k f(\omega) \tag{8.3.3}$$

is harmonic, being a finite sum of harmonic polynomials. Applying Proposition 8.1.9, we have the following.

Proposition 8.3.1. *Given $f \in L^2(S^{n-1})$, the function u defined by* (8.3.1) *is harmonic on B^n, and the sequence $\mathcal{S}_N f(r\omega)$ of harmonic polynomials satisfies*

$$\partial^\alpha \mathcal{S}_N f(x) \longrightarrow \partial^\alpha u(x), \tag{8.3.4}$$

uniformly on each ball $\overline{B_\rho(0)}$, for each $\rho < 1$.

We now want to show that the solution to the Dirichlet problem (8.1.34), with $f \in C(S^{n-1})$, has the representation (8.3.1). Let us fix some notation, and denote by

$$\mathrm{PI} : C(S^{n-1}) \longrightarrow \{u \in C(\overline{B}^n) \cap C^\infty(B) : \Delta u = 0 \text{ on } B^n\} \tag{8.3.5}$$

the solution to (8.1.34) given by Proposition 8.1.7, i.e., the Poisson integral

$$\mathrm{PI}\, f(x) = \frac{1-|x|^2}{A_{n-1}} \int\limits_{S^{n-1}} \frac{f(y)}{|x-y|^n}\, dS(y). \tag{8.3.6}$$

Here is the result.

Proposition 8.3.2. *For all $f \in C(S^{n-1})$,*

$$\mathrm{PI}\, f(r\omega) = \sum_{k=0}^{\infty} r^k E_k f(\omega), \tag{8.3.7}$$

for $r\omega \in B^n$.

Proof. Denote the right side of (8.3.7) by $\widetilde{\mathrm{PI}} f(r\omega)$. As seen in (8.3.1)–(8.3.3) and Proposition 8.1.9, we have

$$\widetilde{\mathrm{PI}} : L^2(S^{n-1}) \longrightarrow \{u \in C^\infty(B^n) : \Delta u = 0 \text{ on } B^n\}, \tag{8.3.8}$$

and

$$|\widetilde{\mathrm{PI}} f(r\omega)| \le c_n \Big(\sum_{k=0}^{\infty} (1+k)^{n-2} r^{2k} \Big)^{1/2} \|f\|_{L^2}. \tag{8.3.9}$$

We want to show that

$$\mathrm{PI}\, f(r\omega) = \widetilde{\mathrm{PI}} f(r\omega), \quad \forall\, r\omega \in B^n, \tag{8.3.10}$$

for all $f \in C(S^{n-1})$. It is clear that if f is a finite sum of eigenfunctions, i.e., if $f \in \mathcal{V}$, defined in (8.2.43), then $\widetilde{\mathrm{PI}} f$ is a smooth solution to (8.1.34), so by the uniqueness part of Proposition 8.1.7, (8.3.10) holds for all $f \in \mathcal{V}$. As seen in Proposition 8.2.7, $\mathcal{V}$ is dense in $C(S^{n-1})$. Thus, given $f \in C(S^{n-1})$, there exist $f_\nu \in \mathcal{V}$ such that

$$f_\nu \to f \text{ uniformly on } C(S^{n-1}), \text{ hence in } L^2\text{-norm}. \tag{8.3.11}$$

We have

$$\mathrm{PI}\, f_\nu(r\omega) = \widetilde{\mathrm{PI}} f_\nu(r\omega), \tag{8.3.12}$$

for all $r\omega \in B^n$. Furthermore, as $\nu \to \infty$,

$$\begin{aligned} \mathrm{PI}\, f_\nu(r\omega) &\longrightarrow \mathrm{PI}\, f(r\omega), \\ \widetilde{\mathrm{PI}} f_\nu(r\omega) &\longrightarrow \widetilde{\mathrm{PI}} f(r\omega), \end{aligned} \tag{8.3.13}$$

the latter result by (8.3.9), applied to $f - f_\nu$. Together, (8.3.12)–(8.3.13) give (8.3.10), hence (8.3.7). $\square$

Combining Proposition 8.3.2 and (8.1.47), we have, for $f \in C(S^{n-1})$,

$$\sum_{k=0}^{\infty} r^k E_k f(\omega) = \frac{1-r^2}{A_{n-1}} \int\limits_{S^{n-1}} \frac{f(y)}{(1-2r\omega\cdot y + r^2)^{n/2}}\, dS(y). \tag{8.3.14}$$

We are consequently motivated to expand the integrand on the right side of (8.3.14) in powers of r. The following "generating function identity,"

$$(1-2tr+r^2)^{-\alpha}=\sum_{k=0}^{\infty}C_k^{\alpha}(t)r^k, \tag{8.3.15}$$

defines a class of special functions known as Gegenbauer polynomials. To compute these, we can use the identity

$$(1-z)^{-\alpha}=\sum_{j=0}^{\infty}\binom{j+\alpha-1}{j}z^j, \tag{8.3.16}$$

with $z=r(2t-r)$, to write the left side of (8.3.15) as

$$\begin{aligned}&\sum_{j=0}^{\infty}\binom{j+\alpha-1}{j}r^j(2t-r)^j\\ &=\sum_{j=0}^{\infty}\sum_{\ell=0}^{j}\binom{j+\alpha-1}{j}\binom{j}{\ell}(-1)^{\ell}r^{j+\ell}(2t)^{j-\ell}\\ &=\sum_{k=0}^{\infty}\sum_{\ell=0}^{[k/2]}(-1)^{\ell}\binom{k-\ell+\alpha-1}{k-\ell}\binom{k-\ell}{\ell}(2t)^{k-2\ell}r^k.\end{aligned} \tag{8.3.17}$$

Hence

$$C_k^{\alpha}(t)=\sum_{\ell=0}^{[k/2]}(-1)^{\ell}\binom{k-\ell+\alpha-1}{k-\ell}\binom{k-\ell}{\ell}(2t)^{k-2\ell}. \tag{8.3.18}$$

We have

$$\mathrm{PI}\, f(r\omega)=\frac{1-r^2}{A_{n-1}}\sum_{k=0}^{\infty}r^k\int_{S^{n-1}}C_k^{n/2}(\omega\cdot y)f(y)\,dS(y), \tag{8.3.19}$$

for $0\le r<1$. Comparison with the left side of (8.3.14) gives

$$E_kf(\omega)=\frac{1}{A_{n-1}}\int_{S^{n-1}}\Big[C_k^{n/2}(\omega\cdot y)-C_{k-2}^{n/2}(\omega\cdot y)\Big]f(y)\,dS(y), \tag{8.3.20}$$

provided we make the convention that $C_k^{\alpha}(t)=0$ for $k<0$. Summing (8.3.20) over $0\le k\le N$ yields

$$S_Nf(\omega)=\frac{1}{A_{n-1}}\int_{S^{n-1}}\Big[C_N^{n/2}(\omega\cdot y)+C_{N-1}^{n/2}(\omega\cdot y)\Big]f(y)\,dS(y). \tag{8.3.21}$$

We seek an alternative formula for E_k. For its derivation, it is convenient to temporarily replace the exponent k in r^k by ν_k, given by

$$\nu_k^2=\lambda_k^2+\Big(\frac{n-2}{2}\Big)^2,\quad \nu_k=k+\frac{n-2}{2}. \tag{8.3.22}$$

Multiplying (8.3.14) by $r^{(n-2)/2}$ and making the change of variable $r = e^{-s}$ yields

$$\sum_{k=0}^{\infty} e^{-\nu_k s} E_k f(\omega) = \frac{2}{A_{n-1}} \int_{S^{n-1}} \frac{\sinh s}{(2\cosh s - 2\omega \cdot y)^{n/2}} f(y)\, dS(y). \tag{8.3.23}$$

Now integrating over $s \in [s_1, \infty)$ and taking $r = e^{-s_1}$ (and dividing by $r^{(n-2)/2}$) gives
(8.3.24)

$$\sum_{k=0}^{\infty} \nu_k^{-1} r^k E_k f(\omega) = \frac{2}{(n-2)A_{n-1}} \int_{S^{n-1}} \frac{f(y)}{(1 - 2r\omega \cdot y + r^2)^{(n-2)/2}}\, dS(y).$$

We again apply the generating formula identity (8.3.15), this time with $\alpha = (n-2)/2$, and compare coefficients of r^k, to get

$$E_k f(\omega) = \frac{2\nu_k}{(n-2)A_{n-1}} \int_{S^{n-1}} C_k^{(n-2)/2}(\omega \cdot y) f(y)\, dS(y). \tag{8.3.25}$$

In the classical case $S^2 \subset \mathbb{R}^3$, these Gegenbauer polynomials specialize to *Legendre polynomials*

$$P_k(t) = C_k^{1/2}(t). \tag{8.3.26}$$

Since $A_2 = 4\pi$ and $\nu_k = k + 1/2$ in this case, we get

$$E_k f(\omega) = \frac{2k+1}{4\pi} \int_{S^2} P_k(\omega \cdot y) f(y)\, dS(y). \tag{8.3.27}$$

We denote the integral kernel of the projection E_k by $E_k(\omega, y)$, so

$$E_k f(\omega) = \int_{S^{n-1}} E_k(\omega, y) f(y)\, dS(y). \tag{8.3.28}$$

Then the content of (8.3.25) is that

$$E_k(\omega, y) = \frac{2\nu_k}{(n-2)A_{n-1}} C_k^{(n-2)/2}(\omega \cdot y). \tag{8.3.29}$$

The following is a useful general identity.

Proposition 8.3.3. *If $\{Y_k^\ell : \ell \in \Sigma_k\}$ is an orthonormal basis of V_k, then*

$$E_k(\omega, y) = \sum_{\ell \in \Sigma_k} Y_k^\ell(\omega) \overline{Y_k^\ell(y)}. \tag{8.3.30}$$

Proof. Denote the right side of (8.3.30) by $F_k(\omega, y)$, and set

$$F_k f(\omega) = \int_{S^{n-n}} F_k(\omega, y) f(y)\, dS(y). \tag{8.3.31}$$

We see that $F_k Y_k^\ell = Y_k^\ell$ for all $\ell \in \Sigma_k$, and that $F_k f = 0$ for $f \perp V_k$, so indeed $F_k = E_k$. $\square$

If we set $\omega = y$ in (8.3.30) and integrate over S^{n-1}, we get

$$\int_{S^{n-1}} E_k(y,y)\,dS(y) = \dim V_k. \tag{8.3.32}$$

Since $\omega = y \in S^{n-1} \Rightarrow \omega \cdot y = 1$, we deduce from (8.3.29) that

$$\dim V_k = \frac{2\nu_k}{n-2} C_k^{(n-2)/2}(1). \tag{8.3.33}$$

On the other hand, setting $t = 1$ in (8.3.15), obtaining $(1-r)^{-2\alpha}$, we have

$$C_k^\alpha(1) = \binom{k+2\alpha-1}{k}, \tag{8.3.34}$$

so (8.3.33) implies

$$\dim V_k = \frac{2k+n-2}{n-2}\binom{k+n-3}{k}, \tag{8.3.35}$$

which, as one can check (writing the numerator in the fraction above as $k + (k+n-2)$), agrees with (8.2.37). Specializing to $n = 3$, we have

$$P_k(1) = C_k^{1/2}(1) = 1, \tag{8.3.36}$$

and hence

$$\text{on } S^2, \quad \dim V_k = 2\nu_k = 2k+1, \tag{8.3.37}$$

in agreement with (8.2.38).

The following identity is a useful complement to (8.3.32).

Proposition 8.3.4. *For each $k \in \mathbb{Z}^+$,*

$$\int_{S^{n-1}}\int_{S^{n-1}} |E_k(\omega,y)|^2\,dS(\omega)\,dS(y) = \dim V_k. \tag{8.3.38}$$

Proof. From (8.3.30), we have

$$|E_k(\omega,y)|^2 = \sum_{\ell,m\in\Sigma_k} Y_k^\ell(\omega)\overline{Y_k^\ell(y)}\,\overline{Y_k^m(\omega)}Y_k^m(y). \tag{8.3.39}$$

Integrating over $S^{n-1} \times S^{n-1}$ and using orthonormality of $\{Y_k^\ell : \ell \in \Sigma_k\}$ gives

$$\begin{aligned} \iint |E_k(\omega,y)|^2\,dS(\omega)\,dS(y) &= \sum_{\ell,m\in\Sigma_k} \delta_{\ell m} \\ &= \sum_{\ell\in\Sigma_k} 1 = \dim V_k. \end{aligned} \tag{8.3.40}$$

□

To proceed, we would like to apply the formula (8.3.29) to (8.3.38). The following general comments are useful. Each $T \in SO(n)$ acts on S^{n-1}, and we have, for each $y \in S^{n-1}$,

$$\begin{aligned} \int_{S^{n-1}} F(\omega \cdot y)\, dS(\omega) &= \int_{S^{n-1}} F(T\omega \cdot Ty)\, dS(\omega) \\ &= \int_{S^{n-1}} F(\omega \cdot Ty)\, dS(\omega), \end{aligned} \tag{8.3.41}$$

the last identity because T is an isometry on S^{n-1}, and hence perserves volumes. It follows that the integral on the left side of (8.3.41) is independent of $y \in S^{n-1}$. We can fix $e \in S^{n-1}$, and obtain

$$\int_{S^{n-1}} \int_{S^{n-1}} F(\omega \cdot y)\, dS(\omega)\, dS(y) = A_{n-1} \int_{S^{n-1}} F(\omega \cdot e)\, dS(\omega). \tag{8.3.42}$$

It follows that, with

$$\gamma_{nk} = \frac{2k+n-2}{(n-2)A_{n-1}}, \tag{8.3.43}$$

we have

$$\begin{aligned} &\int_{S^{n-1}} \int_{S^{n-1}} |E_k(\omega, y)|^2\, dS(\omega)\, dS(y) \\ &= \gamma_{nk}^2 \int_{S^{n-1}} \int_{S^{n-1}} C_k^{(n-2)/2}(\omega \cdot y)^2\, dS(\omega)\, dS(y) \\ &= \gamma_{nk}^2 A_{n-1} \int_{S^{n-1}} C_k^{(n-2)/2}(\omega \cdot e)^2\, dS(\omega), \end{aligned} \tag{8.3.44}$$

and this is equal to $\dim V_k$.

Let us specialize this to $n = 3$, i.e., to analysis on $S^2 \subset \mathbb{R}^3$. Then we have the Legendre polynomials P_k, given by (8.3.36), and (8.3.44) yields

$$\frac{1}{4\pi} \int_{S^2} P_k(\omega \cdot e)^2\, dS(\omega) = \frac{1}{2k+1}. \tag{8.3.45}$$

The identities obtained above from (8.3.38) can be generalized. In fact, from (8.3.30) (plus the fact that each $E_j(\omega, y)$ is real valued) we have

$$\int_{S^{n-1}} E_j(\omega, z) E_k(z, y)\, dS(z) = \delta_{jk} E_k(\omega, y), \quad \forall\, \omega, y \in S^{n-1}. \tag{8.3.46}$$

Using (8.3.29), we can rewrite this as

$$\gamma_{nk} \int_{S^{n-1}} C_j^{(n-2)/2}(\omega \cdot z) C_k^{(n-2)/2}(z \cdot y)\, dS(z) = \delta_{jk} C_k^{(n-2)/2}(\omega \cdot y). \tag{8.3.47}$$

In particular, for $n = 3$,

$$\frac{2k+1}{4\pi} \int_{S^2} P_j(\omega \cdot z) P_k(z \cdot y)\, dS(z) = \delta_{jk} P_k(\omega \cdot y). \tag{8.3.48}$$

Note that taking $j = k$ and $\omega = y = e \in S^2$ gives (8.3.45), since $P_k(1) = 1$.

8.4. Zonal functions

A *zonal function* on S^{n-1} is a continuous function of the form

$$f(\omega) = \varphi(\omega \cdot e), \tag{8.4.1}$$

where $e = (0, \dots, 0, 1)$ (which we might call the "north pole"). We write $f \in \mathcal{Z}(S^{n-1})$. Provided $n \geq 3$, an equivalent condition is the following. Consider $SO(n-1)$ as a subgroup of $SO(n)$ consisting of rotations about the x_n-axis:

$$SO(n-1) = \{R \in SO(n) : Re = e\}. \tag{8.4.2}$$

Then, given $f \in C(S^{n-1})$,

$$f \in \mathcal{Z}(S^{n-1}) \Longleftrightarrow f(R\omega) = f(\omega), \quad \forall\, R \in SO(n-1),\ \omega \in S^{n-1}. \tag{8.4.3}$$

For $n = 2$, $SO(1)$ consists only of the identity transformation, and this equivalence fails. In the rest of this section, we will assume $n \geq 3$.

We define the space of *zonal harmonics*

$$\mathcal{Z}_k = V_k \cap \mathcal{Z}(S^{n-1}). \tag{8.4.4}$$

We have seen examples of elements of $\mathcal{Z}_k$, namely

$$Z_k(\omega) = C_k^{(n-2)/2}(\omega \cdot e). \tag{8.4.5}$$

The following complement to Proposition 8.2.7 is of key importance.

Proposition 8.4.1. *The linear span of* $\{Z_k : k \in \mathbb{Z}^+\}$ *is dense in* $\mathcal{Z}(S^{n-1})$.

Proof. From (8.3.18) we see that $C_k^\alpha(t)$ is a polynomial in t of degree k, whose leading term is

$$\binom{k+\alpha-1}{k}(2t)^k. \tag{8.4.6}$$

Thus the linear span of $\{C_k^\alpha : k \in \mathbb{Z}^+\}$ is the space of all polynomials in t, which, by the Weierstrass approximation theorem, is dense in $C([-1,1])$, so we have the proposition. $\square$

Here is an important corollary of Proposition 8.4.1.

Proposition 8.4.2. *For each* $k \in \mathbb{Z}^+$,

$$\mathcal{Z}_k = \operatorname{Span}(Z_k). \tag{8.4.7}$$

In particular, $\dim \mathcal{Z}_k = 1$.

Proof. Suppose $f \in \mathcal{Z}_k$ and $f \perp Z_k$, i.e., $(f, Z_k)_{L^2} = 0$. Clearly $(f, Z_j)_{L^2} = 0$ for all $j \neq k$, so $(f, g)_{L^2} = 0$ for all $g \in \operatorname{Span}\{Z_j\}$, hence, by Proposition 8.4.1, for all $g \in \mathcal{Z}_k$. Taking $g = f$ yields $\int |f|^2\, dS = 0$, hence $f = 0$. $\square$

To proceed, let us set

$$Y_k^0(\omega) = \|Z_k\|_{L^2}^{-1} Z_k(\omega), \tag{8.4.8}$$

so $\{Y_k^0 : k \in \mathbb{Z}^+\}$ is an orthonormal set of functions in $\mathcal{Z}(S^{n-1})$. As observed in (8.2.55), the action of $SO(n)$ on $C(S^{n-1})$ given by

$$\pi(R)f(\omega) = f(R^{-1}\omega) \tag{8.4.9}$$

preserve each space V_k. Hence it commutes with E_k. Thus the characterization (8.4.3) implies

$$E_k : \mathcal{Z}(S^{n-1}) \longrightarrow \mathcal{Z}_k. \tag{8.4.10}$$

By (8.4.7), the identity (8.2.47) specializes to

$$\begin{aligned} f \in \mathcal{Z}(S^{n-1}) \Rightarrow E_k f(\omega) &= (f, Y_k^0)_{L^2} Y_k^0(\omega) \\ &= \hat{f}(k,0) Y_k^0(\omega). \end{aligned} \tag{8.4.11}$$

Hence Proposition 8.2.9 specializes to the following.

Proposition 8.4.3. *If $f \in \mathcal{Z}(S^{n-1})$, then*

$$S_N f(\omega) = \sum_{k=0}^{N} \hat{f}(k,0) Y_k^0(\omega) \tag{8.4.12}$$

has the property that

$$S_N f \longrightarrow f \ \text{ in } L^2\text{-norm}, \tag{8.4.13}$$

and

$$\sum_{k=0}^{\infty} |\hat{f}(k,0)|^2 = \|f\|_{L^2}^2. \tag{8.4.14}$$

If we specialize to $n = 3$, then (8.3.45) yields

$$Y_k^0(\omega) = \Bigl(\frac{2k+1}{4\pi}\Bigr)^{1/2} P_k(\omega \cdot e), \tag{8.4.15}$$

and we have, for $f(\omega) = \varphi(\omega \cdot e)$,

$$\varphi(\omega \cdot e) = \sum_{k=0}^{\infty} \varphi_k P_k(\omega \cdot e), \tag{8.4.16}$$

with

$$\begin{aligned} \varphi_k &= \frac{2k+1}{4\pi} \int\limits_{S^2} \varphi(\omega \cdot e) P_k(\omega \cdot e)\, dS(\omega) \\ &= \Bigl(k + \frac{1}{2}\Bigr) \int_{-1}^{1} \varphi(t) P_k(t)\, dt, \end{aligned} \tag{8.4.17}$$

a result known as the Funk-Hecke theorem.

We turn to a consideration of the transformation

$$\Pi : C(S^{n-1}) \longrightarrow \mathcal{Z}(S^{n-1}), \tag{8.4.18}$$

given by

$$\Pi f(\omega) = \int\limits_{SO(n-1)} f(R^{-1}\omega)\, dR, \tag{8.4.19}$$

where dR denotes Haar measure on $SO(n-1)$. Since $\pi(R)$ commutes with E_k, so does Π, and we have

$$\Pi_k : V_k \longrightarrow \mathcal{Z}_k, \quad \Pi_k = \Pi\Big|_{V_k}. \tag{8.4.20}$$

Clearly

$$f \in \mathcal{Z}_k \Longrightarrow \Pi_k f = f. \tag{8.4.21}$$

Also, with

$$\mathcal{Z}_k^{\perp} = \{f \in V_k : (f, Z_k)_{L^2} = 0\}, \tag{8.4.22}$$

we have

$$f \in \mathcal{Z}_k^{\perp} \Longrightarrow f_R \in \mathcal{Z}_k^{\perp}, \quad \forall\, R \in SO(n-1), \tag{8.4.23}$$

where $f_R(\omega) = f(R^{-1}\omega)$, so

$$\Pi_k : \mathcal{Z}_k^{\perp} \longrightarrow \mathcal{Z}_k^{\perp} \cap \mathcal{Z}_k = 0. \tag{8.4.24}$$

To summarize:

Proposition 8.4.4. *The map Π_k, given by* (8.4.19)–(8.4.20), *is the orthogonal projection of V_k onto $\mathcal{Z}_k$, hence, for $f \in V_k$,*

$$\Pi_k f(\omega) = (f, Y_k^0)_{L^2} Y_k^0(\omega). \tag{8.4.25}$$

Incidentally, note from (8.4.19) that, for all $f \in V_k$,

$$\Pi_k f(e) = f(e). \tag{8.4.26}$$

Thus we have:

Corollary 8.4.5. *for $f \in V_k$,*

$$f \perp Y_k^0 \Longrightarrow f(e) = 0. \tag{8.4.27}$$

In particular, if $\{Y_k^\ell : \ell \in \Sigma_k\}$ is an orthonormal basis of V_k,

$$\ell \neq 0 \Longrightarrow Y_k^\ell(e) = 0. \tag{8.4.28}$$

Taking into account the identity (8.2.53), we have:

Corollary 8.4.6. *for Y_k^0 given by* (8.4.5) *and* (8.4.8), *we have*

$$|Y_k^0(e)|^2 = \frac{\dim V_k}{A_{n-1}}. \tag{8.4.29}$$

This gives a sharp illustration of the fact, mentioned above Proposition 8.2.10, that the eigenfunctions Y_k^ℓ are not uniformly bounded. For $n = 3$, (8.4.29) specializes to

$$Y_k^0(e) = \Bigl(\frac{2k+1}{4\pi}\Bigr)^{1/2}, \quad \text{on } S^2, \tag{8.4.30}$$

which we have already seen in (8.4.15).

8.5. $SO(n)$ actions on the spaces V_k of spherical harmonics

As we have seen, the rotation group $SO(n)$ acts on functions on S^{n-1} by

$$\pi(R)f(\omega) = f(R^{-1}\omega). \tag{8.5.1}$$

This action works on various function spaces, including $C^\infty(S^{n-1})$, $C(S^{n-1})$, and $L^2(S^{n-1})$. Since the transformation R is an isometry on S^{n-1}, this action on $C^\infty(S^{n-1})$ commutes with the Laplace operator Δ_S, so one has each eigenspace V_k invariant, yielding

$$\pi_k : SO(n) \longrightarrow \mathcal{L}(V_k). \tag{8.5.2}$$

One can check that the map π_k is continuous, in fact smooth. Also, for $R_j \in SO(n)$,

$$\pi_k(R_1R_2) = \pi_k(R_1)\pi_k(R_2), \quad \pi_k(I) = I, \tag{8.5.3}$$

so π_k is a representation of $SO(n)$ on V_k. Note also that $\pi(R)$ and hence each $\pi_k(R)$ preserves the L^2-norm.

Recall that if V is a finite-dimensional inner product space, and

$$\rho : SO(n) \longrightarrow \mathcal{L}(V) \tag{8.5.4}$$

is a smooth representation by transformations that preserve the norm, then ρ is a unitary representation of $SO(n)$. Also recall that a linear subspace $W \subset V$ is said to be invariant under the representation ρ provided

$$\rho(R) : W \longrightarrow W, \quad \forall\, R \in SO(n). \tag{8.5.5}$$

When this holds and ρ is unitary, we also have

$$\rho(R) : W^\perp \longrightarrow W^\perp, \quad \forall\, R \in SO(n), \tag{8.5.6}$$

where $W^\perp = \{v \in V : (v, w) = 0,\ \forall\, w \in W\}$. We recall the proof. If $w \in W,\ v \in W^\perp$, $R \in SO(n)$, then

$$(w, \pi(R)v) = (\pi(R)^*w, v) = (\pi(R^{-1})w, v), \tag{8.5.7}$$

which vanishes if (8.5.5) holds.

Recall also that a unitary representation ρ is said to be irreducible if V has no proper invariant linear subspaces. As just seen, if W is a proper invariant subspace, then $V = W \oplus W^\perp$ and ρ acts on each factor. If $\dim V < \infty$, an inductive procedure decomposes

$$V = W_0 \oplus \cdots \oplus W_M \tag{8.5.8}$$

into factors W_j, on each of which $SO(n)$ acts irreducibly.

With these notions in hand, we can make the following important statement about the representations π_k defined by (8.5.1)–(8.5.2).

Proposition 8.5.1. *For each $k \in \mathbb{Z}^+$, the representation π_k of $SO(n)$ on the eigenspace V_k is irreducible.*

Proof. Suppose $W \subset V_k$ is invariant under π_k and $f \in W$, $f \neq 0$. Then $f(\omega_0) \neq 0$ for some $\omega_0 \in S^{n-1}$. Pick $R_0 \in SO(n)$ such that $R_0\omega_0 = e$. Then $g = \pi(R_0)f \in W$ and $g(e) \neq 0$. Applying Π_k, defined by (8.4.20), we have $h = \Pi_k g \in W \cap \mathcal{Z}_k$, and $h(e) = g(e) \neq 0$. Thanks to Proposition 8.4.2, this implies $Z_k \in W$.

If W is not all of V_k, then $W^\perp \subset V_k$ is a nonzero invariant subspace, and the same argument implies $Z_k \in W^\perp$. Contradiction. □

Taking into account that, when $n = 3$, $\dim V_k = 2k + 1$, and recalling the classification of the irreducible unitary representations of $SO(3)$ given in the first section of Chapter 4, we see that each irreducible representation of $SO(3)$ is contained, exactly once, in the action of $SO(3)$ on $L^2(S^2)$.

For $n \geq 4$, not all irreducible representations of $SO(n)$ are contained in the action on $L^2(S^{n-1})$. For example, let us take $n = 4$ and recall the analysis of the irreducible representations of $SO(4)$ given in the first section of Chapter 4, making use of the two-fold covering

$$SU(2) \times SU(2) \longrightarrow SO(4), \tag{8.5.9}$$

arising from the action of $SU(2) \times SU(2)$ by isometries on $SU(2) \approx S^3$,

$$(g_1, g_2) \cdot x = g_1 x g_2^{-1}, \quad g_j \in SU(2),\ x \in SU(2) \approx S^3. \tag{8.5.10}$$

In such a case, the irreducible representations of $SO(4)$ are given by

$$\begin{aligned} &\gamma_{k\ell}(g) = D_{k/2}(g_1) \otimes D_{\ell/2}(g_2), \quad \text{for} \quad g = (g_1, g_2), \\ &\text{with } k, \ell \in \mathbb{Z}^+,\ k + \ell \text{ even}, \end{aligned} \tag{8.5.11}$$

where $D_{k/2}$ is the irreducible representation of $SU(2)$ on $\mathbb{C}^{k+1}$. Now an orthonormal basis of $L^2(S^3) \approx L^2(SU(2))$ consisting of eigenfunctions of the Laplace-Beltrami operator is formed as follows (by the Peter-Weyl theorem):

$$\sqrt{k+1}\, u_k^{ij}, \quad k \in \mathbb{Z}^+, \quad i, j \in \{1, \dots, k+1\}, \tag{8.5.12}$$

where u_k^{ij} are the matrix entries of $D_{k/2}$. Consequently, the irreducible representations of $SO(4)$ contained in the action on $L^2(S^3)$ are those of the form

$$\gamma_{kk}, \quad k \in \mathbb{Z}^+. \tag{8.5.13}$$

In connection with this, note that

$$S^3 = SO(4)/SO(3) = SU(2) \times SU(2)/K, \tag{8.5.14}$$

where $K \approx SU(2)$ is given by

$$K = \{(g, g) : g \in SU(2)\}, \tag{8.5.15}$$

so the zonal functions on S^3 are the scalar multiples of

$$\chi_{k/2} = \operatorname{Tr} D_{k/2}. \tag{8.5.16}$$

For $k \neq \ell$, the representations $\gamma_{k\ell}$ do not have a nonzero element of $\mathbb{C}^{k+1} \otimes \mathbb{C}^{\ell+1}$ that is invariant under the action of K.

Generally, we say an irreducible unitary representation ρ of $SO(n)$ on V is a *class one* representation provided V contains a nonzero element that is fixed by $\rho(g)$ for all $g \in SO(n-1)$. We have the following result.

Proposition 8.5.2. *Each class one representation of $SO(n)$ is contained, exactly once, in the action on $L^2(S^{n-1})$.*

See Proposition 9.2.8 for a proof, in a more general setting.

Returning to the case $n = 3$, we note that the representations π_k of $SO(3)$ on V_k are equivalent to the representations D_k produced in Chapter 4. We aim to use the analysis of the structure of D_k to produce an explicit orthonormal basis of the space V_k os spherical harmonics on S^2. To this end, it will be convenient to retrace some steps taken in Chapter 4.

To get started, we select the following basis of Skew(3):
(8.5.17)
$$A_1 = \begin{pmatrix} 0 & -1 & \\ 1 & 0 & \\ & & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & & -1 \\ & 0 & \\ 1 & & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & & \\ & 0 & -1 \\ & 1 & 0 \end{pmatrix}.$$

Note that the families of transformations e^{tA_j} are groups of rotations about the x_{3-j}-axis, each periodic in t of period 2π. A straightforward calculation yields the following commutator identities,

$$[A_1, A_2] = A_3, \quad [A_2, A_3] = A_1, \quad [A_3, A_1] = A_2. \tag{8.5.18}$$

Now suppose $\sigma : \mathrm{Skew}(3) \to \mathcal{L}(V)$ is a skew-adjoint representation. Set

$$L_j = \sigma(A_j) \in \mathcal{L}(V), \quad L_j^* = -L_j. \tag{8.5.19}$$

The identities (8.5.15) yield

$$[L_1, L_2] = L_3, \quad [L_2, L_3] = L_1, \quad [L_3, L_1] = L_2. \tag{8.5.20}$$

One key to understanding the structure of this representation is provided by

$$M = L_1^2 + L_2^2 + L_3^2 \in \mathcal{L}(V). \tag{8.5.21}$$

Note that

$$M = M^*, \quad (Mv, v) = -\sum_{j=1}^{3} \|L_j v\|^2 \leq 0, \ \forall\, v \in V, \tag{8.5.22}$$

so V has an orthonormal basis of eigenvalues of M, and all its eigenvalues are ≤ 0. Now, using the general identity

$$[X, Y^2] = [X, Y]Y + Y[X, Y], \quad X, Y \in \mathcal{L}(V), \tag{8.5.23}$$

we can verify that

$$[L_j, M] = 0, \quad \forall\, j. \tag{8.5.24}$$

It follows that each eigenspace of M is invariant under L_1, L_2, and L_3. Consequently, by Schur's lemma,

$$M = -\lambda^2 I, \quad \text{for some } \ \lambda \in [0, \infty]. \tag{8.5.25}$$

To proceed, we diagonalize L_1 on V. Set

$$W_\mu = \{v \in V : L_1 v = i\mu v\}, \quad V = \bigoplus_{i\mu \in \operatorname{Spec} L_1} W_\mu. \tag{8.5.26}$$

The structure of σ is determined by how L_2 and L_3 behave on W_μ. It is convenient to set

$$L_\pm = L_2 \mp i L_3. \tag{8.5.27}$$

The following key identity forllows directly from (8.5.20):

$$[L_1, L_\pm] = \pm i L_\pm. \tag{8.5.28}$$

This leads to the following.

Lemma 8.5.3. *For each* μ,

$$L_\pm : W_\mu \longrightarrow W_{\mu \pm 1}. \tag{8.5.29}$$

In particular, if $i\mu \in \operatorname{Spec} L_1$*, then either* $L_+ = 0$ *on* W_μ *or* $i(\mu+1) \in \operatorname{Spec} L_1$*, and also either* $L_- = 0$ *on* W_μ *or* $i(\mu - 1) \in \operatorname{Spec} L_1$.

Proof. If $v \in W_\mu$, then (8.5.28) implies

$$L_1 L_\pm v = L_\pm L_1 v \pm i L_\pm v = i(\mu \pm 1) L_\pm v, \tag{8.5.30}$$

and this identity yields the lemma. □

Remark. Because of (8.5.29), one calls $L_\pm$ ladder operators.

The following gives more precise information.

Proposition 8.5.4. *If* σ *is an irreducible skew-adjoint representation of* Skew(3) *on* V*, then* $\operatorname{Spec} L_1$ *must consist of a sequence*

$$\operatorname{Spec} L_1 = i\{\mu_0, \mu_0 + 1, \dots, \mu_0 + \ell = \mu_1\}, \tag{8.5.31}$$

with

$$L_+ : W_{\mu_0 + j} \xrightarrow{\approx} W_{\mu_0 + j + 1}, \quad \textit{for} \ \ 0 \le j \le \ell - 1, \tag{8.5.32}$$

and

$$L_- : W_{\mu_1 - j} \xrightarrow{\approx} W_{\mu_1 - j - 1}, \quad \textit{for} \ \ 0 \le j \le \ell - 1. \tag{8.5.33}$$

Proof. A computation gives

$$\begin{aligned} L_-L_+ &= L_2^2 + L_3^2 + i[L_3, L_2] \\ &= -\lambda^2 - L_1^2 - iL_1, \end{aligned} \tag{8.5.34}$$

on V, and similarly

$$L_+L_- = -\lambda^2 - L_1^2 + iL_1 \tag{8.5.35}$$

on V. Hence

$$\begin{aligned} L_-L_+ &= \mu(\mu+1) - \lambda^2, \quad \text{on} \ \ W_\mu, \\ L_+L_- &= \mu(\mu-1) - \lambda^2, \quad \text{on} \ \ W_\mu. \end{aligned} \tag{8.5.36}$$

Also, since L_2 and L_3 are skew-adjoint,

$$L_+ = -L_-^*, \tag{8.5.37}$$

so

$$L_+L_- = -L_-^*L_-, \quad L_-L_+ = -L_+^*L_+. \tag{8.5.38}$$

Hence, we have the identity of null spaces,

$$\mathcal{N}(L_+) = \mathcal{N}(L_-L_+), \quad \mathcal{N}(L_-) = \mathcal{N}(L_+L_-). \tag{8.5.39}$$

These observations establish (8.5.31)–(8.5.33). □

In the setting of Proposition 8.5.4, we see that, if $v \in W_{\mu_0}$ is nonzero, then

$$\operatorname{Span}\{v, L_+v, \dots, L_+^{\mu_1-\mu_0}v\} \tag{8.5.40}$$

is invariant under L_1, L_+, and L_-, hence under the representation σ, so it must be all of V, if σ is irreducible. This implies

$$\dim W_\mu = 1, \quad \text{for } i\mu \in \operatorname{Spec} L_1, \ \mu_0 \le \mu \le \mu_1. \tag{8.5.41}$$

From (8.5.36) we see that

$$\mu_1(\mu_1+1) = \lambda^2 = \mu_0(\mu_0-1), \tag{8.5.42}$$

hence

$$\begin{aligned} \mu_1 - \mu_0 = \ell \Longrightarrow \mu_0 &= -\frac{\ell}{2}, \quad \mu_1 = \frac{\ell}{2}, \\ \dim V &= \ell + 1, \quad \text{and} \\ \lambda^2 &= \frac{\ell(\ell+2)}{4}. \end{aligned} \tag{8.5.43}$$

The vectors in the basis (8.5.40) of V are mutually orthogonal, since the various eigenspaces of L_1 are, but they do not form an orthonormal basis. To nail down the structure of the action of Skew(3) on V, we have the following.

Proposition 8.5.5. *Assume σ is an irreducible skew-adjoint representation of* Skew(3) *on V, $i\mu \in \operatorname{Spec} L_1$, and $w_\mu \in W_\mu$. Then the norms of the vectors $L_\pm w_\mu \in W_{\mu\pm1}$ are given by*

$$\begin{aligned}\|L_+ w_\mu\|^2 &= [\lambda^2 - \mu(\mu+1)]\,\|w_\mu\|^2,\\ \|L_- w_\mu\|^2 &= [\lambda^2 - \mu(\mu-1)]\,\|w_\mu\|^2.\end{aligned} \tag{8.5.44}$$

Proof. Using (8.5.36) and (8.5.37), we have

$$\begin{aligned}\|L_+ w_\mu\|^2 &= (L_+^* L_+ w_\mu, w_\mu)\\ &= -(L_- L_+ w_\mu, w_\mu)\\ &= [\lambda^2 - \mu(\mu+1)]\,\|w_\mu\|^2.\end{aligned} \tag{8.5.45}$$

The computation of $\|L_- w_\mu\|^2$ is similar. ☐

This leads to the following explicit description of the action of Skew(3) on V.

Proposition 8.5.6. *Let V be a complex inner product space of dimension $\ell+1$, $\ell \in \mathbb{Z}^+$. If σ is an irreducible skew-adjoint representation of* Skew(3) *on V, then V has an orthonormal basis*

$$v_\mu, \quad \mu = -\frac{\ell}{2} + j, \quad j \in \{0, \dots, \ell\}, \tag{8.5.46}$$

with respect to which

$$\begin{aligned}L_1 v_\mu &= i\mu v_\mu,\\ L_+ v_\mu &= \sqrt{\lambda^2 - \mu(\mu+1)}\,v_{\mu+1},\\ L_- v_\mu &= \sqrt{\lambda^2 - \mu(\mu-1)}\,v_{\mu-1},\end{aligned} \tag{8.5.47}$$

where

$$\lambda^2 = \frac{\ell(\ell+2)}{4}. \tag{8.5.48}$$

Consequently, for each $\ell \in \mathbb{Z}^+$, there is, up to equivalence, just one such representation of Skew(3).

Among the representations of Skew(3) arising in Proposition 8.5.6 are the derived representations σ_k associated to the representations π_k of $SO(3)$ on $V_k \subset C(S^2)$, with $\dim V_k = 2k+1$. Thus we get "half" of the representations described in Proposition 8.5.6, those for which ℓ is even. As seen in Chapter 4, the other half arise from irreducible unitary representations of $SU(2)$.

We now concentrate on the representations π_k of $SO(3)$ and derived representations σ_k, on the eigenspaces $V_k \subset C^\infty(S^2)$ of Δ_S, and see how

Proposition 8.5.6 yields further specific formulas. First note that, since $\dim V_k = 2k+1$, (8.5.48) gives

$$\lambda = k(k+1) = \lambda_k, \tag{8.5.49}$$

with λ_k as in (8.2.13), with $n=3$. Equivalently,

$$L_1^2 + L_2^2 + L_3^2 = \Delta_S. \tag{8.5.50}$$

Next, we have an orthonormal basis of V_k of the form

$$Y_k^\ell, \quad -l \le \ell \le k, \tag{8.5.51}$$

satisfying

$$L_1 Y_k^\ell = i\ell Y_k^\ell. \tag{8.5.52}$$

For $\ell = 0$, this identifies Y_k^0 as a zonal function. Following (8.4.15), we take

$$Y_k^0(\omega) = \Bigl(\frac{2k+1}{4\pi}\Bigr)^{1/2} P_k(\omega \cdot e), \tag{8.5.53}$$

where $e = (0,0,1)$ and $P_k(t)$ are the Legendre polynomials. For each $\ell \in \{-k,\dots,0,\dots,k\}$, (8.5.52) says

$$Y_k^\ell(e^{tA_1}\omega) = e^{i\ell t} Y_k^\ell(\omega), \tag{8.5.54}$$

with A_1 given by (8.5.17). Writing a general element $f \in V_k$ as $f = \sum_j c_j Y_k^j$, we deduce the following result.

Proposition 8.5.7. *Given $f \in V_k$, $\ell \in \{-k,\dots,0,\dots,k\}$, we have*

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-i\ell t} f(e^{tA_1}\omega)\,dt = (f, Y_k^\ell)_{L^2}\, Y_k^\ell(\omega). \tag{8.5.55}$$

Elements of V_k that one could plug into (8.5.55) include

$$f_k^y(\omega) = P_k(\omega\cdot y), \quad y \in S^2, \tag{8.5.56}$$

and

$$g_k^c(\omega) = \Bigl(\sum_j c_j\omega_j\Bigr)^k, \quad c_j \in \mathbb{C},\ \sum_j c_j^2 = 0, \tag{8.5.57}$$

since then $(\sum c_j x_j)^2$ is a harmonic polynomial, homogeneous of degree k. A particular case of this is

$$g_k(\omega) = (\omega_1 + i\omega_2)^k, \tag{8.5.58}$$

which satisfies $L_1 g_k = ikg_k$, implying

$$Y_k^k(\omega) = \alpha_k(\omega_1 + i\omega_2)^k, \tag{8.5.59}$$

for some constant α_k.

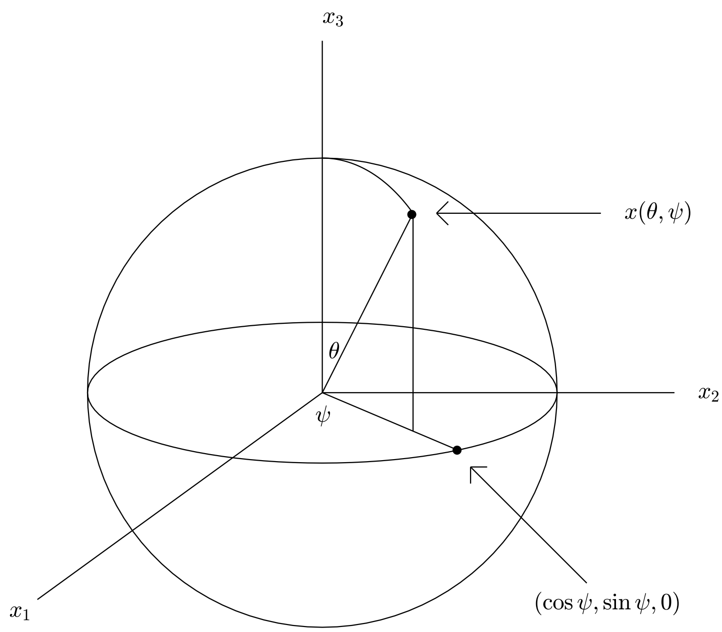

Figure 8.5.1. Spherical coordinates $x(\theta, \psi) = (\sin\theta\cos\psi, \sin\theta\sin\psi, \cos\theta)$

A more direct path to explicit formulas for Y_k^ℓ is found by applying (8.5.47), to get

$$L_+ Y_k^\ell(\omega) = \sqrt{k(k+1) - \ell(\ell+1)}\, Y_k^{\ell+1}(\omega). \tag{8.5.60}$$

We can start at $\ell = 0$ with (8.5.53) and apply this iteratively to obtain formulas for $Y_k^\ell(\omega)$, for $1 \le \ell \le k$. For $-k \le \ell \le -1$, we could work similarly with iterates of L_-, or we could just take

$$Y_k^{-\ell}(\omega) = \overline{Y_k^\ell(\omega)}, \tag{8.5.61}$$

noting that $L_1 \overline{f} = \overline{L_1 f}$.

To apply (8.5.60) explicitly, we use *spherical coordinates* (θ, ψ), defined by

$$\begin{gathered} x(\theta, \psi) = (\sin\theta\cos\psi, \sin\theta\sin\psi, \cos\theta), \\ 0 \le \theta \le \pi, \quad 0 \le \psi \le 2\pi, \end{gathered} \tag{8.5.62}$$

so $\theta = 0$ defines the "north pole," $(0, 0, 1) = e$, and $\theta = \pi$ defines the south pole. See Figure 8.5.1. In these coordinates, we have

$$L_1 = \frac{\partial}{\partial\psi}, \quad L_\pm = ie^{\pm i\psi}\Big[\pm\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\psi}\Big]. \tag{8.5.63}$$

Also, in these coordinates (8.5.53) takes the form

$$Y_k^0(\theta,\psi) = \Big(\frac{2k+1}{4\pi}\Big)^{1/2} P_k(\cos\theta). \tag{8.5.64}$$

To start the iteration (8.5.60) at $\ell = 0$, we have the general formula

$$L_+ g(\theta) = ie^{i\psi} g'(\theta), \tag{8.5.65}$$

hence

$$L_+ G(\cos\theta) = -ie^{i\psi}(\sin\theta)G'(\cos\theta). \tag{8.5.66}$$

More generally, we calculate

$$\begin{aligned} &L_+\Big(e^{i\ell\psi}\sin^\ell\theta\, G_\ell(\cos\theta)\Big) \\ &= -ie^{i(\ell+1)\psi}\sin^{\ell+1}\theta\, G_\ell'(\cos\theta). \end{aligned} \tag{8.5.67}$$

Hence, inductively, we obtain the formula

$$Y_k^\ell(\theta,\psi) = \alpha_{k\ell}e^{i\ell\psi}\sin^\ell\theta\, P_k^{(\ell)}(\cos\theta), \quad 0 \le \ell \le k, \tag{8.5.68}$$

with constants $\alpha_{k\ell}$ obtainable via (8.5.60). Recall that $P_k(t)$ is a polynomial in t of degree k, so $P_k^{(k)}(t)$ is constant, so (8.5.68) gives

$$Y_k^k(\theta,\psi) = \alpha_k e^{ik\psi} sin^k\theta, \tag{8.5.69}$$

which recovers (8.5.59), since, by (8.5.62),

$$e^{i\psi}\sin\theta = \omega_1 + i\omega_2. \tag{8.5.70}$$

In light of this identity, we see that another way to write (8.5.68) is

$$Y_k^\ell(\omega) = \alpha_{k\ell}(\omega_1 + i\omega_2)^\ell P_k^{(\ell)}(\omega_3), \quad 0 \le \ell \le k. \tag{8.5.71}$$

We record the conclusion.

Proposition 8.5.8. *Each eigenspace $V_k \subset C^\infty(S^2)$ of the Laplace operator Δ_S on S^2 has an orthonormal basis $\{Y_k^\ell : -k \le \ell \le k\}$ of the form*

$$Y_k^0(\omega) = \Big(\frac{2k+1}{4\pi}\Big)^{1/2} P_k(\omega_3), \tag{8.5.72}$$

and, for $1 \le \ell \le k$ (if $k \ge 1$),

$$\begin{aligned} Y_k^\ell(\omega) &= \alpha_{k\ell}(\omega_1 + i\omega_2)^\ell P_k^{(\ell)}(\omega_3), \\ Y_k^{-\ell}(\omega) &= \alpha_{k\ell}(\omega_1 - i\omega_2)^\ell P_k^{(\ell)}(\omega_3), \end{aligned} \tag{8.5.73}$$

with coefficients $\alpha_{k\ell} \in (0,\infty)$ obtainable from (8.5.60).

8.6. $SO(n)$ actions on V_k, continued

Our next order of business is to fit the representations π_k of $SO(n)$ on $V_k \subset L^2(S^{n-1})$ into the classification of irreducible unitary representations of $SO(n)$, as presented in §§7.1–7.2. We recall the classification.

Suppose $n = 2\ell$ or $2\ell+1$. We denote by E_j ($1 \le j \le \ell$) the generators of 2π-periodic rotations in the (x_{2j-1}, x_{2j})-plane in $\mathbb{R}^n$. Then $\{E_j : 1 \le j \le \ell\}$ forms a basis of the Lie algebra $\mathfrak{h}$ of a maximal torus $\mathbb{T} \subset SO(n)$. In (7.2.64)–(7.2.65), we describe the irreducible unitary representations of $SO(n)$ as having highest weights that are non-negative integer combinations of certain ℓ-tuples,

$$(1,0,\dots,0),\ (1,1,0,\dots,0),\dots,(1,\dots,1,1),\ \text{and}\ (1,\dots,1,-1), \tag{8.6.1}$$

when $n = 2\ell$, for a representation on $\Lambda^\mu \mathbb{C}^{2\ell}$, $1 \le \mu \le \ell - 1$, and on $\Lambda^\ell_\pm \mathbb{C}^{2\ell}$, and

$$(1,0,\dots,0),\ (1,1,0,\dots,0),\dots,(1,\dots,1,1), \tag{8.6.2}$$

when $n = 2\ell + 1$, for a representation on $\Lambda^\mu \mathbb{C}^{2\ell+1}$, $1 \le \mu \le \ell$. We denote the representations with highest weights given by (8.6.1) and (8.6.2) by

$$D_{(1,0,\dots,0)},\ D_{(1,1,0,\dots,0)},\dots,D_{(1,\dots,1,1)},\ \text{and}\ D_{(1,\dots,1,-1)}, \tag{8.6.3}$$

and

$$D_{(1,0,\dots,0)},\ D_{(1,1,0,\dots,0)},\dots,D_{(1,\dots,1,1)}, \tag{8.6.4}$$

respectively. Then the general irreducible unitary representation of $SO(n)$ is denoted

$$D_{(d_1,\dots,d_\ell)}, \quad n = 2\ell \text{ or } 2\ell+1, \tag{8.6.5}$$

where the indices satisfy

$$\begin{aligned} d_1 \ge \cdots \ge d_{\ell-1} \ge |d_\ell|, &\quad \text{if } n = 2\ell, \\ d_1 \ge \cdots \ge d_\ell \ge 0, &\quad \text{if } n = 2\ell+1. \end{aligned} \tag{8.6.6}$$

In case our representation of $SO(n)$ is π_k on $V_k \subset L^2(S^{n-1})$, we see that the highest weight vector is

$$(x_1 + ix_2)^k, \tag{8.6.7}$$

with weight

$$(k, 0, \dots, 0). \tag{8.6.8}$$

This yields the following.

Proposition 8.6.1. *The representation π_k of $SO(n)$ on $V_k \subset L^2(S^{n-1})$ is equivalent to*

$$D_{(k,0,\dots,0)}. \tag{8.6.9}$$

8.7. Characters of the representations π_k of $SO(n)$ on V_k

To study the characters

$$\chi_k(g) = \operatorname{Tr} \pi_k(g), \quad g \in SO(n), \tag{8.7.1}$$

we will use the identity

$$\begin{aligned} \sum_{k=0}^{\infty} r^k \pi_k(g) E_k f(x) &= \operatorname{PI} f(rg^{-1}x) \\ &= \frac{1-r^2}{A_{n-1}} \int_{S^{n-1}} \frac{f(y)}{|rx - gy|^n} \, dS(y) \\ &= \int_{S^{n-1}} K_n(r, g, x, y) f(y) \, dS(y), \end{aligned} \tag{8.7.2}$$

to write

$$\begin{aligned} \sum_{k=0}^{\infty} r^k \chi_k(g) &= \int_{S^{n-1}} K_n(r, g, y, y) \, dS(y) \\ &= \frac{1-r^2}{A_{n-1}} \int_{S^{n-1}} \frac{dS(y)}{|gy - ry|^n}, \end{aligned} \tag{8.7.3}$$

or, alternatively,

$$\sum_{k=0}^{\infty} r^k \chi_k(g) = \frac{1-r^2}{A_{n-1}} \int_{S^{n-1}} \frac{dS(y)}{(1 - 2ry \cdot gy + r^2)^{n/2}}. \tag{8.7.4}$$

Note that taking $g = I$ in (8.7.4) yields

$$\begin{aligned} \sum_{k=0}^{\infty} r^k \dim V_k &= \frac{1-r^2}{(1-r)^n} = \frac{1+r}{(1-r)^{n-1}} \\ &= \sum_{k=0}^{\infty} \binom{k+n-2}{k} r^k + \sum_{k=0}^{\infty} \binom{k+n-2}{k} r^{k+1} \\ &= \sum_{k=0}^{\infty} \binom{k+n-2}{k} r^k + \sum_{k=1}^{\infty} \binom{k+n-3}{k-1} r^k, \end{aligned} \tag{8.7.5}$$

providing a further alternative proof of the dimension formula (8.2.37).

An alternative approach to these characters is to deduce from (8.3.28)–(8.3.29) that

$$\chi_k(g) = \frac{2\nu_k}{(n-2)A_{n-1}} \int_{S^{n-1}} C_k^{(n-2)/2}(y \cdot gy) \, dS(y). \tag{8.7.6}$$

8.A. Dimension of $\mathcal{P}_k(\mathbb{R}^n)$

Let $\mathcal{P}_k(\mathbb{R}^n)$ denote the space of polynomials on $\mathbb{R}^n$ that are homogeneous of degree k. We want to compute its dimension. We start with the identity

$$d_k(n) := \dim \mathcal{P}_k(\mathbb{R}^n) = \sum_{|\alpha|=k} 1, \tag{8.A.1}$$

where $\alpha = (\alpha_1, \dots, \alpha_n),\ \alpha_\nu \geq 0,\ |\alpha| = \alpha_1 + \cdots + \alpha_n$. This is the value at $x_1 = \cdots = x_n = 1$ of

$$p_k(x) = \sum_{|\alpha|=k} x^\alpha, \tag{8.A.2}$$

where $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. To attack (8.A.2), we look at the infinite series

$$\begin{aligned}\sum_{\alpha\geq 0} x^\alpha &= \sum_{\alpha_1\geq 0} \cdots \sum_{\alpha_n \geq 0} x_1^{\alpha_1} \cdots x_n^{\alpha_n} \\ &= (1-x_1)^{-1} \cdots (1-x_n)^{-1}.\end{aligned} \tag{8.A.3}$$

This series converges provided each $|x_j| < 1$. This does not immediately lead to an evaluation of (8.A.2) at $x_j = 1$, but let us consider

$$\sum_{\alpha\geq 0} t^{|\alpha|} x^\alpha = \sum_{\alpha \geq 0} (tx)^\alpha = (1-tx_1)^{-1} \cdots (1-tx_n)^{-1}, \tag{8.A.4}$$

valid for $|x_j| \leq 1$ as long as $|t| < 1$. We deduce from (8.A.2) and (8.A.4) that

$$\sum_{k\geq 0} t^k p_k(x) = (1-tx_1)^{-1} \cdots (1-tx_n)^{-1}, \tag{8.A.5}$$

for $|x_j| \leq 1,\ |t| < 1$. Consequently,

$$\sum_{k\geq 0} d_k(n) t^k = (1-t)^{-n}, \quad \text{for} \quad |t| < 1. \tag{8.A.6}$$

If we denote the right side of (8.A.6) by $f_n(t)$, then repeated differentiation gives

$$f_n^{(k)}(t) = n(n+1) \cdots (n+k-1)\,(1-t)^{-n-k}, \tag{8.A.7}$$

hence

$$(1-t)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} t^k. \tag{8.A.8}$$

Comparison with (8.A.6) gives the following conclusion.

Proposition 8.A.1. *For $n \geq 1,\ k \geq 0$,*

$$\dim \mathcal{P}_k(\mathbb{R}^n) = \binom{n+k-1}{k}. \tag{8.A.9}$$

Given that

$$\mathcal{P}_k(\mathbb{R}^n) \approx \mathcal{P}^k(\mathbb{R}^{n-1}), \tag{8.A.10}$$

where $\mathcal{P}^k(\mathbb{R}^n)$ denotes the space of polynomials on $\mathbb{R}^n$ of degree $\leq k$, we also get:

Corollary 8.A.2. *For $n \geq 1,\ k \geq 0$,*

$$\dim \mathcal{P}^k(\mathbb{R}^n) = \binom{n+k}{k}. \tag{8.A.11}$$

Remark. The formulas (8.A.2)–(8.A.5) hold for complex x as well as real x. For $x_j = e^{i\theta_j}$, one gets the characters of an important family of representations of the unitary group $U(n)$. See §§4.3 and 4.13.

8.B. Invariant function spaces on a compact homogeneous space

Let M be a compact Riemannian manifold. Assume its isometry group G acts transitively on M. We say M is a homogeneous space. The group G acts on functions on M, by

$$\pi(g)f(x) = f(g^{-1}x), \quad x \in M, \ g \in G. \tag{8.B.1}$$

This action is unitary on $L^2(M)$. Let $V \subset C(M)$ be a linear space of functions satisfying

$$\dim V < \infty, \quad \pi(g) : V \to V, \ \forall\, g \in G. \tag{8.B.2}$$

The L^2 inner product gives V the structure of an inner product space. We aim to prove the following.

Proposition 8.B.1. *Let $\{u_k : k \in S\}$ be an orthonormal basis of V. Then*

$$\sum_{k \in S} |u_k(x)|^2 = \frac{\dim V}{\mathrm{Vol}(M)}, \quad \forall\, x \in M. \tag{8.B.3}$$

Proof. Since $\pi(g)$ is unitary, (8.B.2) implies $\pi(g) : V^\perp \to V^\perp$. Hence, if

$$E : L^2(M) \longrightarrow V \tag{8.B.4}$$

is the orthogonal projection, we have

$$\pi(g)E = E\pi(g). \tag{8.B.5}$$

Note that, for $f \in L^2(M)$,

$$Ef(x) = \int_M E(x,y) f(y)\, dV(y), \tag{8.B.6}$$

with

$$E(x,y) = \sum_k u_k(x) \overline{u_k(y)}. \tag{8.B.7}$$

Now

$$\pi(g)Ef(x) = \int_M E(g^{-1}x, y) f(y)\, dV(y), \tag{8.B.8}$$

and

$$\begin{aligned} E\pi(g)f(x) &= \int_M E(x,y) f(g^{-1}y)\, dV(y) \\ &= \int_M E(x, gy) f(y)\, dV(y), \end{aligned} \tag{8.B.9}$$

the latter identity because the G-action preserves volumes. Hence

$$E(g^{-1}x, y) = E(x, gy), \quad \forall\, x, y \in M,\ g \in G. \tag{8.B.10}$$

Taking $x = gy$ yields

$$E(y, y) = E(gy, gy), \quad \forall\, g \in G,\ y \in M. \tag{8.B.11}$$

Transitivity of the G-action then implies that $E(y, y)$ is independent of y. Using (8.B.7), we obtain

$$\sum_k |u_k(x)|^2 = E(x, x) = A^2, \quad \text{independent of } x. \tag{8.B.12}$$

Integrating both sides over $x \in M$ gives

$$A^2 \operatorname{Vol}(M) = \sum_k \int_M |u_k(x)|^2 \, dV(x) = \dim V, \tag{8.B.13}$$

and we have (8.B.3). □

As a corollary, we have the following estimate.

Corollary 8.B.2. *In the setting of Proposition 8.B.1, if $f \in V$, then*

$$\sup_x |f(x)| \le \Bigl(\frac{\dim V}{\operatorname{Vol}(M)}\Bigr)^{1/2} \|f\|_{L^2}. \tag{8.B.14}$$

Proof. With $\hat{f}(k) = (u_k, f)$, we have, for each $x \in M$,

$$\begin{aligned} |f(x)| &= \Bigl|\sum_{k \in S} \hat{f}(k) u_k(x)\Bigr| \\ &\le \Bigl(\sum_{k \in S} |\hat{f}(k)|^2\Bigr)^{1/2} \Bigl(\sum_{k \in S} |u_k(x)|^2\Bigr)^{1/2}, \end{aligned} \tag{8.B.15}$$

yielding (8.B.14), as a consequence of (8.B.3). □

Chapter 9

Representations of compact groups on eigenspaces of Δ

Let M be a compact Riemannian manifold, with Laplace-Beltrami operator Δ. General results from elliptic PDE (cf. [**39**], Chapter 5) imply that $L^2(M)$ has an orthonormal basis of eigenfunctions of Δ. More precisely, there exist $\lambda_k \in [0,\infty)$, $\lambda_k \nearrow \infty$, such that for each k,

$$V_{\lambda_k} = \{u \in C^\infty(M) : \Delta u = -\lambda_k^2 u\} \tag{9.0.1}$$

is nonzero and finite dimensional, and if we pick an orthonormal basis of each V_{λ_k}, then their union is an orthonormal subset of $L^2(M)$ whose span is dense in $L^2(M)$ (in fact, dense in $C^\infty(M)$). We say $\lambda_k \in \operatorname{Spec}\Lambda$, where $\Lambda = \sqrt{-\Delta}$, so $\Lambda u = \lambda_k u$ for $u \in V_{\lambda_k}$.

Suppose G is a compact Lie group of isometries of M. We obtain a unitary representation π of G on $L^2(M)$:

$$\pi(g)u(x) = u(g^{-1}x). \tag{9.0.2}$$

Then

$$\pi(g)\Delta u = \Delta \pi(g)u, \quad \forall\, g \in G,\ u \in C^\infty(M), \tag{9.0.3}$$

and hence $\pi(g) : V_\lambda \to V_\lambda$ for each $\lambda \in \operatorname{Spec}\Lambda$. We obtain the representations

$$\pi_\lambda(g) : V_\lambda \longrightarrow V_\lambda, \quad \pi_\lambda(g) = \pi(g)\big|_{V_\lambda}. \tag{9.0.4}$$

We investigate these representations in various situations, and see how results on them illuminate analysis on M.

In §9.1 we take M to be a homogeneous space, so G acts transitively on M. Picking $p_0 \in M$, we have

$$M \approx G/K, \quad K = \{g \in G : gp_0 = p_0\}. \tag{9.0.5}$$

In such cases, $d\pi(\Delta_G) = \mathcal{L}$ is an elliptic self-adjoint operator on M. It is of interest to know whether

$$\mathcal{L} = \alpha\Delta, \tag{9.0.6}$$

for some constant α, where Δ is the Laplace-Beltrami operator on M. Proposition 9.1.1 gives a sufficient condition for this to hold.

We take a look at K-invariant elements of a finite-dimensional space $V \subset C^\infty(M)$ that is invariant under π, setting

$$\mathcal{Z}(V) = \{z \in V : \pi(k)z = z, \ \forall\, k \in K\}. \tag{9.0.7}$$

We note that

$$V \neq 0 \Longrightarrow \mathcal{Z}(V) \neq 0, \tag{9.0.8}$$

and that

$$\dim \mathcal{Z}(V) = 1 \Longrightarrow \pi \ \text{ is irreducible on } V. \tag{9.0.9}$$

In §9.2 we concentrate on a special class of compact homogeneous spaces, known as compact rank-one symmetric spaces. For such M,

$$K \text{ acts transitively on } S_{p_0}M, \tag{9.0.10}$$

where $S_{p_0}M$ denotes the unit sphere in $T_{p_0}M$. Leading examples are the spheres $S^{n-1} = SO(n)/SO(n-1)$. Other examples are given in (9.2.2). One goal here is to generalize certain results of Chapter 8. For example, with V_λ as in (9.0.4) and $\mathcal{Z}_\lambda = \mathcal{Z}(V_\lambda)$, we have

$$\begin{aligned} \lambda \in \operatorname{Spec} \Lambda &\Longrightarrow \dim \mathcal{Z}_\lambda = 1 \\ &\Longrightarrow \pi_\lambda \text{ is irreducible on } V_\lambda, \end{aligned} \tag{9.0.11}$$

whenever M is a compact, rank-one symmetric space. If we take

$$z_\lambda \in \mathcal{Z}_\lambda, \quad z_\lambda(p_0) > 0, \quad \|z_\lambda\|_{L^2} = 1, \tag{9.0.12}$$

then, extending (8.3.25), we have that the orthogonal projection of $L^2(M)$ onto V_λ is given by

$$E_\lambda v(x) = z_\lambda(p_0) \int\limits_M v(y) z_\lambda(g^{-1}y)\, dV(y), \quad x = gK. \tag{9.0.13}$$

We also extend from S^{n-1} to general compact, rank-one symmetric M the dimension identity

$$\dim V_\lambda = z_\lambda(p_0)^2 \operatorname{Vol}(M), \tag{9.0.14}$$

and the characterization of irreducible representations of G that are contained in $L^2(M)$ as class one representations.

In §9.3 we move away from homogeneous spaces, and instead look at cases where G is a *finite group* of isometries of M. We show that if $\pi : G \to \mathcal{L}(L^2(M))$ is injective and G is noncommutative, then infinitely many of the spaces V_λ have dimension > 1. We use heat equation techniques to establish the following asymptotic result. For an irreducible representation ρ of G (on a space of dimension $d(\rho)$), let $V_{\rho,\lambda}$ denote the subspace of V_λ on which G acts like copies of ρ. Set

$$W_R = \bigoplus_{\lambda \le R} V_\lambda, \quad W_{\rho,R} = \bigoplus_{\lambda \le R} V_{\rho,\lambda}. \tag{9.0.15}$$

Then, for each ρ,

$$\lim_{R\to\infty} \frac{\dim W_{\rho,R}}{\dim W_R} = \frac{d(\rho)^2}{o(g)}. \tag{9.0.16}$$

9.1. Homogeneous spaces

Take M as in the introduction to this chapter, let G be a group of isometries of M. We say M is a homogeneous space if G acts transitively on M. In such a case, we can pick $p_0 \in M$ and obtain

$$M \approx G/K, \quad K = \{g \in M : gp_0 = p_0\}. \tag{9.1.1}$$

It follows from Frobenius's' theorem that if G has a bi-invariant metric tensor and if

$$\{X_k : 1 \leq j \leq \gamma\} \tag{9.1.2}$$

is an orthonormal basis of $\mathfrak{g}$, then

$$L_j = d\pi(X_j),\ 1 \leq j \leq \gamma \text{ spans } T_xM, \text{ for each } x \in M. \tag{9.1.3}$$

In such a case $\sum X_j^2 = \Delta_G$ is the Laplace-Beltrami operator on G, and

$$\begin{aligned} d\pi(\Delta_G) = \mathcal{L} &\text{ is a negative semidefinite, self-adjoint} \\ &\text{second-order elliptic differential operator on } M. \end{aligned} \tag{9.1.4}$$

Sometimes $\mathcal{L} = \Delta$ (the Laplace-Beltrami operator on M), at least up to a scalar multiple, but not always, as we will see. On the other hand, each operator L_j, and hence $\mathcal{L}$, is a limit of elements of the algebra of operators on $C^\infty(M)$ given by $\{\pi(g) : g \in G\}$ (a property that Δ might or might not possess). We deduce from (9.0.3) that

$$\mathcal{L}\Delta u = \Delta\mathcal{L}u, \quad \forall\, u \in C^\infty(M). \tag{9.1.5}$$

In particular, with V_λ given by (9.0.1), we have

$$\mathcal{L} : V_\lambda \longrightarrow V_\lambda, \quad \forall\, \lambda \in \operatorname{Spec} \Lambda. \tag{9.1.6}$$

The following result identifies a class of homogeneous spaces for which Δ and $\mathcal{L}$ are essentially the same. Note that, by (9.1.1), K acts on $T_{p_0}M$.

Proposition 9.1.1. *Assume that*

$$\begin{aligned} T_{p_0}M &\text{ has just one } K\text{-invariant inner product,} \\ &\text{up to a scalar multiple.} \end{aligned} \tag{9.1.7}$$

Then there exists $\alpha > 0$ such that $\mathcal{L} = \alpha\Delta$.

Proof. If (9.1.6) holds, there exists $\alpha > 0$ such that $\mathcal{L}$ and $\alpha\Delta$ have the same principal symbol at p_0, hence at each $x \in M$ (since they both commute with π), so

$$\mathcal{L} - \alpha\Delta = X \ \text{ is a first-order differential operator on } M. \tag{9.1.8}$$

This operator is real and annihilates constants, so X is a real vector field. This implies $X + X^*$ is a zero-order operator. But $\mathcal{L}$ and Δ are self-adjoint, so $X = X^*$, and hence $X = 0$. ∎

If $M = G/K$ is a compact homogeneous space, and if V is a finite-dimensional subspace of $C^\infty(M)$ that is invariant under π, set

$$\mathcal{Z}(V) = \{z \in V : \pi(k)z = z,\ \forall\, k \in K\}. \tag{9.1.9}$$

The following is a useful observation.

Proposition 9.1.2. *If $V \neq 0$, then $\mathcal{Z}(V) \neq 0$.*

Proof. Take a nonzero $v \in V$, say $v(x_0) \neq 0$. Then take $g \in G$ such that $gx_0 = p_0$. It follows that $w(x) = v(g^{-1}x)$ belongs to V and satisfies $w(p_0) \neq 0$, so

$$z = \int\limits_K \pi(k)w\, dk \in \mathcal{Z}(V), \quad \text{and} \quad z(p_0) = w(p_0) \neq 0. \tag{9.1.10}$$

□

This yields a useful method for establishing irreducibility.

Corollary 9.1.3. *Take π as in* (9.0.2)*. If $\pi : V \to V$ and* $\dim \mathcal{Z}(V) = 1$*, then π is irreducible on V.*

9.2. Rank-one symmetric spaces

Let M be a compact, connected Riemannian manifold, G a transitive group of isometries, $p_0 \in M$, $K = \{g \in G : gp_0 = p_0\}$. We say M is a rank-one symmetric space if

$$K \text{ acts transitively on the unit sphere } S_{p_0}M \text{ in } T_{p_0}M. \tag{9.2.1}$$

Note that, in such a case, the hypothesis (9.1.7) of Proposition 9.1.1 holds. Furthermore, if (9.2.1) holds, then also K_0, the connected component of the identity in K, acts transitively on $S_{p_0}M$, provided $\dim M \geq 2$. Clearly $S^{n-1} = SO(n)/SO(n-1)$ is a rank-one symmetric space. Other examples include

$$\begin{aligned} &G = SU(n),\ K = SU(n-1), \quad M = \mathbb{CP}^{n-1}, \\ &G = Sp(n),\ K = Sp(n-1), \quad M = \mathbb{HP}^{n-1}, \\ &G = F_4,\ K = \mathrm{Spin}(9), \quad M = \mathbb{C}a\mathbb{P}^2, \end{aligned} \tag{9.2.2}$$

the last one called the Cayley projective plane. As in §9.1, we set

$$\pi(g)u(x) = u(g^{-1}x). \tag{9.2.3}$$

If $V \subset C^\infty(M)$ is a finite dimensional space invariant under π, we define $\mathcal{Z}(V)$ as in (9.1.10). In particular, we have the eigenspace V_λ of the Laplace operator, as in (9.0.1), and we set

$$\mathcal{Z}_\lambda = \{z \in V_\lambda : \pi(k)z = z,\ \forall\, k \in K\}. \tag{9.2.4}$$

Proposition 9.1.2 applies:

$$V_\lambda \neq 0 \Longrightarrow \mathcal{Z}_\lambda \neq 0. \tag{9.2.5}$$

The following result generalizes Proposition 8.4.2, but the proof here, making use of basic results of PDE, is completely different from the proof given there.

Proposition 9.2.1. *If M is a rank-one symmetric space, then*

$$\lambda \in \mathrm{Spec}\,\Lambda \Longrightarrow \dim \mathcal{Z}_\lambda = 1. \tag{9.2.6}$$

To prove this, we start with a lemma.

Lemma 9.2.2. *In the setting of Proposition 9.2.1, there exists $\delta = \delta(\lambda) > 0$ such that for each nonzero $u \in \mathcal{Z}_\lambda$,*

$$u(x) \neq 0, \quad \forall\, x \in B_\delta(p_0) \setminus p_0. \tag{9.2.7}$$

Proof. Using the variational characterization of the fundamental frequency for the Dirichlet problem on the domain $\partial B_\varepsilon(p_0)$, you can pick $\delta = \delta(\lambda) > 0$ so small that

$$\begin{aligned}(\Delta + \lambda^2)u = 0 \text{ on } B_\delta(p_0),\ u = 0 \text{ on } \partial B_\varepsilon(p_0),\ \varepsilon < \delta \\ \Longrightarrow u \equiv 0 \text{ on } B_\varepsilon(p_0).\end{aligned} \tag{9.2.8}$$

Now if $u \in \mathcal{Z}_\lambda$ and $u(y) = 0$ for some $y \in B_\delta(p_0)$, say $\operatorname{dist}(y, p_0) = \varepsilon < \delta$, then $u = 0$ on $\partial B_\varepsilon(p_0)$, so (9.2.8) implies $u \equiv 0$ on $B_\varepsilon(p_0)$. Then the Holmgren uniqueness theorem implies $u \equiv 0$ on M. ☐

Proof of Proposition 9.2.1. Suppose real valued $v, w \in \mathcal{Z}_\lambda$. Take δ as in Lemma 9.2.2, $r \in (0, \delta)$. There exist $a, b \in \mathbb{R}$, not both zero, such that

$$av + bw = 0 \text{ on } \partial B_r(p_0). \tag{9.2.9}$$

Then Lemma 9.2.2 implies

$$av + bw \equiv 0 \text{ on } M, \tag{9.2.10}$$

so v and w are linearly dependent. This proves (9.2.6). ☐

As a direct consequence of Proposition 9.2.1 and Corollary 9.1.3, we have the following generalization of Proposition 8.5.1.

Corollary 9.2.3. *In the setting of Proposition 9.2.1,*

$$\lambda \in \operatorname{Spec} \Lambda \Longrightarrow \pi \textit{ acts irreducibly on } V_\lambda. \tag{9.2.11}$$

Let z_λ span $\mathcal{Z}_\lambda$, and arrange that

$$z_\lambda(p_0) > 0, \quad \|z_\lambda\|_{L^2} = 1. \tag{9.2.12}$$

We aim to establish a key relation between $z_\lambda(p_0)$ and $\dim V_\lambda$. The following result gets us started.

Proposition 9.2.4. *If $\lambda \in \operatorname{Spec} \Lambda$, then*

$$v \in V_\lambda,\ v \perp z_\lambda \Longrightarrow v(p_0) = 0. \tag{9.2.13}$$

Proof. Given such v,

$$w(x) = \int_K v(k^{-1}x)\, dk \Rightarrow w \in V_\lambda,\ w \perp \mathcal{Z}_\lambda, \quad \text{and} \quad w(p_0) = v(p_0). \tag{9.2.14}$$

But also $w \in \mathcal{Z}_\lambda$, so $w = 0$, hence $v(p_0) = 0$. ☐

Now for our identity.

Proposition 9.2.5. *The element $z_\lambda \in \mathcal{Z}_\lambda$ satisfying* (9.2.12) *also satisfies*

$$z_\lambda(p_0)^2 = \frac{\dim V_\lambda}{\operatorname{Vol}(M)}. \tag{9.2.15}$$

Proof. Let $\{z_\lambda, v_j\}$ be an orthonormal basis of V_λ. Then, as a special case of Proposition 8.B.1,

$$z_\lambda(x)^2 + \sum_j |v_j(x)|^2 = \frac{\dim V_\lambda}{\mathrm{Vol}(M)}, \quad \forall\, x \in M. \tag{9.2.16}$$

Setting $x = p_0$ and applying Proposition 9.2.4 yields (9.2.15). $\square$

Remark. The identity (9.2.15) generalizes (8.4.29), which treats the case $M = S^{n-1}$.

To proceed, it is useful to consider the map

$$A_\lambda : V_\lambda \longrightarrow L^2(M), \quad A_\lambda(v)(x) = (v, \pi(g)z_\lambda)_{L^2},\ x = gK. \tag{9.2.17}$$

We see that, for $h \in G$,

$$\begin{aligned} \pi(h)A_\lambda(v)(gK) &= A_\lambda(v)(h^{-1}gK) \\ &= (v, \pi(h^{-1}g)z_\lambda) \\ &= (\pi(h)v, \pi(g)z_\lambda) \\ &= A_\lambda(\pi(h)v)(x). \end{aligned} \tag{9.2.18}$$

Consequently

$$\pi(h)A_\lambda = A_\lambda \pi(h), \quad \forall\, h \in G. \tag{9.2.19}$$

Denote the image of A_λ by W_λ, so

$$A_\lambda : V_\lambda \longrightarrow W_\lambda, \tag{9.2.20}$$

By (9.2.19), $\pi(h) : W_\lambda \to W_\lambda$ for all $h \in G$. Hence

$$\begin{aligned} &B_\lambda = A_\lambda^* A_\lambda : V_\lambda \longrightarrow V_\lambda, \quad \text{and} \\ &\pi(h)^{-1} B_\lambda \pi(h) = B_\lambda, \quad \forall\, h \in G, \end{aligned} \tag{9.2.21}$$

so by Schur's lemma B_λ is a scalar multiple of the identity on V_λ. It is not zero since $A_\lambda z_\lambda(p_0) = (z_\lambda, z_\lambda) = 1$, so A_λ is an isomorphism in (9.2.20), unitary up to a scalar, and π acts irreducibly on W_λ. Since the various irreducible spaces V_μ are mutually inequivalent, this forces $W_\lambda = V_\lambda$. Hence

$$A_\lambda : V_\lambda \longrightarrow V_\lambda, \quad \pi_\lambda(h)A_\lambda = A_\lambda \pi_\lambda(h),\ \forall\, h \in G, \tag{9.2.22}$$

so, by Schur's lemma, A_λ is a scalar multiple of the identity on V_λ,

$$A_\lambda v = c_\lambda v, \quad \forall\, v \in V_\lambda. \tag{9.2.23}$$

Taking $v = z_\lambda,\ x = p_0$ yields

$$c_\lambda z_\lambda(p_0) = (z_\lambda, z_\lambda) = 1, \tag{9.2.24}$$

and we have the following conclusion.

Proposition 9.2.6. *If $M = G/K$ is a rank-one symmetric space, $\lambda \in \operatorname{Spec}\Lambda$, and $z_\lambda \in \mathcal{Z}_\lambda$ satisfies* (9.2.16), *then*

$$v(x) = z_\lambda(p_0) \int_M v(y) z_\lambda(g^{-1}y)\, dV(y), \quad \forall\, v \in V_\lambda,\ x = gK. \tag{9.2.25}$$

The formula (9.2.25) extends to one for the orthogonal projection of $L^2(M)$ onto V_λ. Compare (8.3.25).

Proposition 9.2.7. *If M is a rank-one symmetric space, the orthogonal projection of $L^2(M)$ onto V_λ is given by*

$$E_\lambda v(x) = z_\lambda(p_0) \int_M v(y) z_\lambda(g^{-1}y)\, dV(y), \quad x = gK. \tag{9.2.26}$$

Proof. The operator E_λ defined by (9.2.26) satisfies $E_\lambda v = v$ for all $v \in V_\lambda$, by Proposition 9.2.6, and $E_\lambda v = 0$ for $v \perp V_\lambda$, since then $(v, \pi(g) z_\lambda) = 0$, $\forall\, g$. $\square$

Note that we can write

$$\begin{gathered} E_\lambda v(x) = \int_M E_\lambda(x,y) v(y)\, dV(y), \\ E_\lambda(x,y) = z_\lambda(p_0) z_\lambda(g^{-1}y), \quad x = gK. \end{gathered} \tag{9.2.27}$$

We have

$$E_\lambda(gK, hK) = z_\lambda(p_0) z_\lambda(Kg^{-1}hK), \tag{9.2.28}$$

hence

$$E_\lambda(x,x) = z_\lambda(p_0)^2, \tag{9.2.29}$$

so

$$\operatorname{Tr} E_\lambda = z_\lambda(p_0)^2 \operatorname{Vol}(M), \tag{9.2.30}$$

recovering (9.2.15), since $\operatorname{Tr} E_\lambda = \dim V_\lambda$. Compare the computation (8.3.32)–(8.3.35).

Another interesting result arises from considering the following variant of (9.2.17). Suppose ρ is an irreducible unitary representation of G on an inner product space V. Assume V contains a K-invariant element, satisfying

$$z \in V, \quad \rho(k) z = z,\ \forall\, k \in K, \quad \|z\|_V = 1. \tag{9.2.31}$$

Now consider

$$A : V \longrightarrow L^2(M), \quad Av(x) = (v, \pi(g) z)_V,\ x = gK. \tag{9.2.32}$$

Parallel to (9.2.18)–(9.2.19), we have

$$\pi(h)A = A\rho(h), \quad \forall\, h \in G. \tag{9.2.33}$$

Denoting the range of A by $W \subset L^2(M)$, we have analogues of (9.2.20)–(9.2.21), yielding that

$$A : V \longrightarrow W \subset L^2(M), \quad \pi(h) : W \longrightarrow W,\ \forall\, h \in G, \tag{9.2.34}$$

A is an isomorphism, unitary up to a scalar, and π acts irreducibly on W, hence $W = V_\lambda$ for some $\lambda \in \operatorname{Spec}\Lambda$, so

$$\begin{aligned} &A : V \longrightarrow W,\ \text{isomorphically}, \\ &\pi_\lambda(h)A = A\rho(h), \quad \forall\, h \in G, \end{aligned} \tag{9.2.35}$$

and the representation ρ of G on V is equivalent to π_λ on V_λ. An irreducible unitary representation ρ on V for which there is a $z \in V$ satisfying (9.2.31) is called a *class one representation* of G. We have the following conclusion.

Proposition 9.2.8. *If $M = G/K$ is a rank-one symmetric space, each class one irreducible unitary representation of G is contained, exactly once, in $L^2(M)$.*

9.3. Finite symmetry group actions on eigenspaces

The presence of a noncommutative finite group G of symmetries of an elliptic, self-adjoint differential operator L, with discrete spectrum, can force the existence of an infinite number of multiple eigenvalues of L. We show that such a phenomenon occurs rather generally, due to the fact that some irreducible representation ρ of G of degree > 1 must be contained in infinitely many eigenspaces of L. We then show that, under mild assumptions, the relative frequency of occurrence of each irreducible representation ρ of G in the sum of all the eigenspaces with eigenvalues $\leq R$ tends as $R \to \infty$ to a limit equal to the relative frequency that ρ occurs in the regular representation of G. We also study an example where the multiplicities can get arbitrarily large.

Inevitability of multiple eigenvalues

Let G be a finite group of measure-preserving transformations of a nonatomic, σ-finite measure space $(X, \mathcal{B}, \mu)$. Thus G has a unitary representation on $H = L^2(X, \mu)$, given by $U(g)f(x) = f(g^{-1}x)$. We make the hypothesis that

$$U : G \longrightarrow \mathcal{U}(H) \text{ is injective.} \tag{9.3.1}$$

Our first result is the following.

Proposition 9.3.1. *Let K be a compact self-adjoint operator on $H = L^2(X, \mu)$. Assume also that K is injective. Assume G is noncommutative, and that* (9.3.1) *holds. If K commutes with $U(g)$ for all $g \in G$, then K has infinitely many multiple eigenvalues.*

Proof. As a preliminary comment, we note that, if H_0 is the closed linear span of all the 1-dimensional eigenspaces of K, then G acts on H_0, and the restriction of U, given by $V(g) = U(g)\big|_{H_0}$, has the property that $G/\ker V$ is commutative. Thus H_0 cannot equal H, since then (9.3.1) would imply G is commutative. The content of the proposition is that the orthogonal complement H_1 of H has infinite dimension.

To see this, we first note that there is an irreducible representation ρ of G, on a space V_ρ of dimension greater than 1, such that ρ is contained in U. This follows by the same sort of argument as above; if every irreducible representation of G contained in U were one-dimensional, then G would act as a commutative group of transformations of H, and this contradicts our hypotheses.

Now, given such ρ, consider the orthogonal projection P_ρ of H onto the subspace on which G acts as a sum of copies of ρ; it is given by

$$P_\rho = \frac{d(\rho)}{o(G)} \sum_{g\in G} \overline{\chi_\rho(g)}\, U(g), \tag{9.3.2}$$

where $\chi_\rho(g) = \mathrm{Tr}\ \rho(g)$ and $d(\rho) = \dim V_\rho$. In other words,

$$P_\rho f(x) = \frac{d(\rho)}{o(G)} \sum_{g\in G} \overline{\chi_\rho(g)}\, f(g^{-1}x). \tag{9.3.3}$$

We know that $P_\rho \neq 0$. The proof will be complete when we show that the rank of P_ρ is infinite, since no eigenspace of K can contain infinitely many copies of ρ.

Now, if P_ρ has finite rank, it would be a Hilbert-Schmidt operator, so there would exist $\varphi \in L^2(X \times X)$ such that

$$P_\rho f(x) = \int_X \varphi(x,y) f(y)\ d\mu(y). \tag{9.3.4}$$

However, as long as $P_\rho \neq 0$, (9.3.3) and (9.3.4) are incompatible if X has no atoms. In fact, we see that, for any $A \in \mathcal{B}$, the set

$$\widetilde{A} = \{x \in X : \text{not } \varphi(x,y) = 0 \text{ a.e. } y \in A\}$$

satisfies

$$\mu(\widetilde{A}) \le o(G)\, \mu(A). \tag{9.3.5}$$

If X has no atoms, this implies $\varphi = 0$ a.e. □

Example 1. Let X be the circle S^1 with its standard arc-length measure. Let $V \in C(S^1)$ be real valued. Then the differential operator

$$L = \frac{d^2}{d\theta^2} - V(\theta) \tag{9.3.6}$$

has compact resolvent. Take $K = (L - \lambda)^{-1}$ for some sufficiently large $\lambda \in (0, \infty)$. We deduce that:

Corollary 9.3.2. *If $V(\theta)$ is invariant under a rotation through $2\pi/\ell$, for some $\ell \ge 3$, and also invariant under a reflection, then L has infinitely many double eigenvalues.*

Of course, each eigenspace of L has dimension 1 or 2. The argument just recounted arose in a conversation of the author and E. Trubowitz, in 1975.

Example 2. Let X be an equilateral triangle in the plane. Consider the

Laplace operator Δ on X, with the Dirichlet (or Neumann) boundary condition. Then Δ is invariant under the group of isometries of X, a group isomorphic to S_3. It follows that Δ has infinitely many multiple eigenvalues.

M. Pinsky has shown that, in this case, Δ has eigenspaces of arbitrarily large dimension. We have a similar situation when X is a square. The number theoretic explanation that Δ has eigenspaces of arbitrarily high dimension is well known in that case.

Note that Example 2 can be extended to every regular polygon in $\mathbb{R}^2$, and also to every regular polyhedron in $\mathbb{R}^3$. I do not know if Δ has eigenspaces of arbitrarily high dimension in all these cases.

There are other variations of Example 2, to which Proposition 9.3.1 applies. For example, X could be a wriggly perturbation of an equilateral triangle, still having S_3 as a symmetry group.

To take a variation of Example 1, consider the action of S_5 on $\mathbb{R}^3$, as the group of isometries of the regular icosahedron. This also provides a group of isometries of the unit sphere S^2. One can then consider

$$L = \Delta - V, \tag{9.3.7}$$

where $V \in C(S^2)$ is real valued and invariant under this action of S_5.

Asymptotic density

Here we will specialize, as follows. We take X to be Ω, an open subset of some smooth Riemannian manifold M, such that $\overline{\Omega}$ is compact. We make the following geometrical hypothesis on the action of G on Ω, which implies (9.3.1):

$$g \neq e \Longrightarrow \mathrm{Vol}\{x \in \Omega : gx = x\} = 0. \tag{9.3.8}$$

We suppose L is a strongly elliptic, second order, negative semidefinite, differential operator on Ω, and that the action of U on $L^2(\Omega)$ commutes with the semigroup e^{tL}. Assume either that L has the Dirichlet boundary condition, or that it has some other coercive boundary condition, such as the Neumann boundary condition, and $\partial\Omega$ is sufficiently regular, so that the standard asymptotic analysis of the integral kernel $p(t, x, y)$ of e^{tL} is valid.

In such a case, e^{tL} is trace class for each $t > 0$. For each irreducible representation ρ of G, the operator P_ρ given by (9.3.2) commutes with e^{tL}, and we have the following two identities. On the one hand,

$$\mathrm{Tr} P_\rho e^{tL} = \sum (\dim V_{\rho,\lambda}) e^{-t\lambda^2}, \tag{9.3.9}$$

where the sum is over $\lambda \in \mathrm{Spec}\, \sqrt{-L}$ and $V_{\rho,\lambda}$ is the subspace of the λ^2-eigenspace V_λ of $-L$ on which U acts as a sum of copies of ρ. On the other hand, by (9.3.3),

$$\mathrm{Tr} P_\rho e^{tL} = \frac{d(\rho)}{o(G)} \sum_{g\in G} \overline{\chi_\rho(g)} \int_\Omega p(t, g^{-1}x, x)\, dV(x). \tag{9.3.10}$$

The asymptotic analysis of $p(t,x,y)$ alluded to above implies

$$\int_\Omega p(t,x,x)\, dV(x) = (\mathrm{Vol}\ \Omega)(4\pi t)^{-n/2} + o(t^{-n/2}), \tag{9.3.11}$$

as $t \searrow 0$. On the other hand, the behavior of $p(t,x,y)$ off the diagonal yields the following, in cases where (9.3.8) holds:

$$g \neq e \Longrightarrow \int_\Omega p(t, g^{-1}x, x)\, dV(x) = o(t^{-n/2}). \tag{9.3.12}$$

Hence, under these hypotheses, we have

$$\mathrm{Tr} P_\rho e^{tL} = \frac{d(\rho)^2}{o(G)} (\mathrm{Vol}\ \Omega)(4\pi t)^{-n/2} + o(t^{-n/2}), \quad t \searrow 0. \tag{9.3.13}$$

We are ready to prove the following.

Proposition 9.3.3. *For $R \in (0,\infty)$, set*

$$W_R = \bigoplus_{\lambda \le R} V_\lambda, \quad W_{\rho,R} = \bigoplus_{\lambda \le R} V_{\rho,\lambda}. \tag{9.3.14}$$

Then, for each irreducible representation ρ of G,

$$\lim_{R\to\infty} \frac{\dim W_{\rho,R}}{\dim W_R} = \frac{d(\rho)^2}{o(G)}. \tag{9.3.15}$$

Proof. The asymptotic behavior

$$\dim W_R = \frac{\mathrm{Vol}\ \Omega}{\Gamma(\frac{n}{2}+1)(4\pi)^{n/2}} R^n + o(R^n), \quad R \to \infty, \tag{9.3.16}$$

follows from (9.3.11), via Karamata's Tauberian theorem; cf. [**39**], Chapter 8. The same argument applied to (9.3.13) yields

$$\dim W_{\rho,R} = \frac{d(\rho)^2}{o(G)} \frac{\mathrm{Vol}\ \Omega}{\Gamma(\frac{n}{2}+1)(4\pi)^{n/2}} R^n + o(R^n), \quad R \to \infty, \tag{9.3.17}$$

and then (9.3.15) follows. $\square$

Note that Proposition 9.3.3 applies to all the examples mentioned in the early paragraphs of this section.

Regarding the right side of (9.3.15), we note that the subspace of $\ell^2(G)$ on which the regular representation of G acts like copies of ρ is a space of dimension $d(\rho)^2$.

D_4 acting on $\mathbb{T}^2$: high multiplicities

The dihedral group D_4 acts as a group of isometries of $\mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z}^2)$, hence as a unitary group on $L^2(\mathbb{T}^2)$, leaving invariant each eigenspace of the Laplace operator Δ.

Proposition 9.3.4. *Each irreducible representation ρ of D_4 has the property that there are eigenspaces of Δ containing arbitrarily many copies of ρ.*

To see this, first recall one way of showing that there are eigenspaces of Δ of arbitrarily high dimension. Namely,

$$\mathrm{Spec}(-\Delta) = \{j^2 + k^2 : j, k \in \mathbb{Z}\}, \tag{9.3.18}$$

and if $\nu = j^2 + k^2$, the dimension of the ν-eigenspace of $-\Delta$ is equal to the number of pairs $(j, k) \in \mathbb{Z} \times \mathbb{Z}$ such that $j^2 + k^2 = \nu$. Now, number theoretical constraints imply that the set of sums of two squares has mean density zero in $\mathbb{Z}^+$. On the other hand, the sum of the dimensions of the ν-eigenspaces of $-\Delta$, for $\nu \le R$, which is the number of integer lattice points within a disk of radius $\sqrt{R}$, behaves like πR as $R \to \infty$. It follows that some eigenspaces must have arbitrarily large dimension.

Now, by Proposition 9.3.3, the same argument extends to the parts of the eigenspaces of Δ on which D_4 acts like copies of ρ, so the proposition follows.

Chapter 10

The groups Sp(n) and their representations

The groups $Sp(n)$, introduced in §1.2, consist of matrices $A \in M(n, \mathbb{H})$ satisfying $A^*A = I$, where $\mathbb{H}$ denotes the algebra of quaternions. Here we study the structure and representations of these groups.

We begin in §10.1 with a treatment of quaternions, supplementing that given in §1.2. For the sake of continuity, we reproduce some of the results of §1.2. We provide a second proof of the important associativity property of multiplication on $\mathbb{H}$, different from that arising in §1.2. We relate this to results on automorphisms of $\mathbb{H}$, which we then extend to an action of $SO(3)$ as a group of automorphisms of $\mathbb{H}$. Conjugation produces an action of the group of unit quaternions $Sp(1)$ as automorphisms of $\mathbb{H}$, yielding a 2-to-1 covering homomorphism $Sp(1) \to SO(3)$, consistent with a natural isomorphism $Sp(1) \approx SU(2)$, established here, and a 2-to-1 covering $SU(2) \to SO(3)$ produced in §4.1.

Section 10.2 comprises a minicourse on quaternionic linear algebra, i.e., the study of $\mathbb{H}$-linear transformations on quaternionic vector spaces, with particular attention to quaternionic inner product spaces. This has notable differences from real and complex linear algebra, and is likely not so familiar to most readers. We present basic material on dimension and bases, particularly orthonormal bases. We define eigenvectors in this setting and show that if $V \neq 0$, each $T \in \mathcal{L}_{\mathbb{H}}(V)$ has an eigenvector. We define the adjoint T^* of $T \in \mathcal{L}_{\mathbb{H}}(V)$ when V is a quaternionic inner product space, and show that if either $T^* = T$ or $T^* = -T$, then V has an orthonormal basis of eigenvectors of T. We then define $Sp(V)$ to consist of $T \in \mathcal{L}_{\mathbb{H}}(V)$ such that

$T^*T = I$, and show that whenever $T \in Sp(V)$, V has an orthonormal basis of eigenvectors of T. In case $V = \mathbb{H}^n$, with the standard inner product, $Sp(V) = Sp(n)$, and the last result then bears on identifying a maximal torus in $Sp(n)$.

In §10.3 we consider roots, weights, and representations of $Sp(1)$ and $Sp(2)$. For $Sp(1)$ there is not much to do, since, as stated above, we have $\mathrm{Sp}(2) \approx SU(2)$, which was analyzed in §4.1. For $Sp(2)$, which has Lie algebra

$$\mathfrak{sp}(2) = \Big\{ \begin{pmatrix} u_1 & \xi \\ -\overline{\xi} & u_2 \end{pmatrix} : \xi \in \mathbb{H}, u_\ell \in \mathbb{R}^3 \Big\},$$

we take

$$\mathfrak{h} = \Big\{ \vartheta_{ab} = \begin{pmatrix} ai & \\ & bi \end{pmatrix} : a, b \in \mathbb{R} \Big\},$$

and find 8 roots and associated root vectors. We observe a similarity to the root diagram of $SO(5)$, treated in §7.1. This is no accident; in fact

$$Sp(2) \approx \mathrm{Spin}(5),$$

which allows us to describe the irreducible unitary representations of $Sp(2)$, using the results of Chapter 7. To derive this isomorphism, we consider the representation of $Sp(2)$ on $M(2, \mathbb{H})$,

$$\kappa : Sp(2) \to \mathcal{L}_{\mathbb{R}}(M(2, \mathbb{H})), \quad \kappa(A)X = AXA^*,$$

and decompose

$$M(2, \mathbb{H}) = \mathfrak{sp}(2) \oplus \{cI : c \in \mathbb{R}\} \oplus S_0^2(\mathbb{H}),$$

with

$$S_0^2(\mathbb{H}) = \Big\{ \begin{pmatrix} a & \overline{\xi} \\ \xi & -a \end{pmatrix} : a \in \mathbb{R},\ \xi \in \mathbb{H} \Big\},$$

obtaining from

$$\kappa_0 : Sp(2) \longrightarrow \mathcal{L}_{\mathbb{R}}(S_0^2(\mathbb{H}))$$

a two-fold covering homomorphism $\kappa_0 : Sp(2) \to SO(5)$. We also relate this covering homomorphism to the natural action of $Sp(2)$ on the quaternionic projective space $\mathbb{P}^1(\mathbb{H})$.

in §10.5 we analyze the roots of $Sp(n)$, for $n \geq 3$. We take

$$\mathfrak{h} = \{\vartheta_a : a \in \mathbb{R}^n\}, \quad \vartheta_a = \begin{pmatrix} a_1 i & & \\ & \ddots & \\ & & a_n i \end{pmatrix},$$

and show the roots consist of

$$\rho_\ell^{\pm}(a) = \pm 2a_\ell, \quad \rho_{\ell m}^{\alpha\beta}(a) = (-1)^\alpha a_\ell + (-1)^\beta a_m,$$

where, respectively, $1 \le \ell \le n$ and $1 \le \ell < m \le n$ and $\alpha, \beta \in \{0, 1\}$. With respect to a natural ordering, the positive roots are ρ_ℓ^+ and $\rho_{\ell m}^{0\beta}$, and the simple roots are

$$\rho_{\ell,\ell+1}^{01},\ 1 \le \ell \le n-1, \quad \text{and} \quad \rho_n^+.$$

Examining their inner products produces the Dynkin diagram for $Sp(n)$.

In §10.6, we examine the weights and representations of $Sp(n)$. We identify the possible weights of representations as the elements $\lambda \in \mathfrak{h}'$ satisfying

$$\lambda(a) = \lambda_1 a_1 + \cdots + \lambda_n a_n,$$

with $\lambda_\ell \in \mathbb{Z}$. We identify the dominant integral weights as those elements satisfying

$$\lambda_\ell \in \mathbb{Z}^+, \quad \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge 0.$$

We assign the label $D_{(\lambda_1,\ldots,\lambda_n)}$ to the irreducible unitary representation of $Sp(n)$ with highest weight $\lambda(a) = \lambda_1 a_1 + \cdots + \lambda_n a_n$ (unique up to equivalence). We consider the standard representation π of $Sp(n)$ on $\mathbb{H}^n$ (a complex vector space isomorphic to $\mathbb{C}^{2n}$), compute its weights, and identify it as

$$\pi = D_{(1,0,\ldots,0)}.$$

We extend the results of §10.3 on κ to the analogous representation of $Sp(n)$ on $M(n, \mathbb{H})$, which decomposes as

$$M(n, \mathbb{H}) = \mathfrak{sp}(n) \oplus \{cI : c \in \mathbb{R}\} \oplus S_0^2(\mathbb{H}^n),$$

giving the representations Ad and κ_0 on $\mathfrak{sp}(n)$ and $S_0^2(\mathbb{H}^n)$, and their complexifications. We examine their weights, and see that

$$\mathrm{Ad} = D_{(2,0,\ldots,0)}, \quad \kappa_0 = D_{(1,1,0,\ldots,0)}.$$

Furthermore, we show that, for each $k \in \{1, \ldots, n\}$

$$D_{(1,\ldots,1,\ldots0)} \text{ (k ones) occurs as a subrepresentation of } \Lambda^k\pi \text{ on } \Lambda^k_{\mathbb{C}}\mathbb{H}^n.$$

This verifies the theorem of the highest weight for $Sp(n)$ and hence provides a representation-theoretic proof that this group is simply connected.

10.1. Quaternions

The space $\mathbb{H}$ of quaternions is a four-dimensional real vector space, identified with $\mathbb{R}^4$, with basis elements $1, i, j, k$, the element 1 identified with the real number 1. Elements of $\mathbb{H}$ are represented as follows:

$$\xi = a + bi + cj + dk, \tag{10.1.1}$$

with $a, b, c, d \in \mathbb{R}$. We call a the real part of ξ ($a = \operatorname{Re}\xi$) and $bi + cj + dk$ the vector part (denoted $\operatorname{Im}\xi$). We also have a multiplication on $\mathbb{H}$, an $\mathbb{R}$-bilinear map $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$, such that $1 \cdot \xi = \xi \cdot 1 = \xi$, and otherwise governed by the rules

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik, \tag{10.1.2}$$

and

$$i^2 = j^2 = k^2 = -1. \tag{10.1.3}$$

Otherwise stated, if we write

$$\xi = a + u, \quad a \in \mathbb{R}, \quad u \in \mathbb{R}^3, \tag{10.1.4}$$

and similarly write $\eta = b + v,\ b \in \mathbb{R},\ v \in \mathbb{R}^3$, the product is given by

$$\xi\eta = (a + u)(b + v) = (ab - u \cdot v) + av + bu + u \times v. \tag{10.1.5}$$

Here $u \cdot v$ is the dot product in $\mathbb{R}^3$, and $u \times v$ is the cross product. The quantity $ab - u \cdot v$ is the real part of $\xi\eta$ and $av + bu + u \times v$ is the vector part. Note that

$$\xi\eta - \eta\xi = 2u \times v. \tag{10.1.6}$$

It is useful to take note of the following symmetries of $\mathbb{H}$.

Proposition 10.1.1. *Let $K : \mathbb{H} \to \mathbb{H}$ be an $\mathbb{R}$-linear transformation such that $K1 = 1$ and K cyclically permutes (i, j, k) (e.g., $Ki = j,\ Kj = k,\ Kk = i$). Then K preserves the product in $\mathbb{H}$, i.e.,*

$$K(\xi\eta) = K(\xi)K(\eta), \quad \forall\, \xi, \eta \in \mathbb{H}. \tag{10.1.7}$$

We say K is an automorphism of $\mathbb{H}$. We also have (10.1.7) *if K switches two of (i, j, k) and changes the sign of one (e.g., $Ki = i,\ Kj = k,\ Kk = -j$).*

Proof. This is straightforward from the multiplication rules (10.1.2)–(10.1.3). □

Using Proposition 10.1.1, we give a second proof of associativity.

Proposition 10.1.2. *Multiplication in $\mathbb{H}$ is associative, i.e.,*

$$\zeta(\xi\eta) = (\zeta\xi)\eta, \quad \forall\, \zeta, \xi, \eta \in \mathbb{H}. \tag{10.1.8}$$

Proof. Given the $\mathbb{R}$-bilinearity of the product, it suffices to check (10.1.8) when each ζ, ξ, and η is either $1, i, j$, or k. Since 1 is the multiplicative unit, the result (10.1.8) is easy when any factor is 1. Furthermore, one can use Proposition 10.1.1 to reduce the possibilities further; for example, one can take $\xi = i$. Then it suffices to show that

$$\zeta(i\eta) = (\zeta i)\eta, \quad \forall\, \zeta, \eta \in \{i, j, k\}.$$

Taking $\eta = i, j, k$, respectively, we have the task of verifying that

$$-\zeta = (\zeta i)i, \quad \zeta k = (\zeta i)j, \quad -\zeta j = (\zeta i)k, \tag{10.1.9}$$

for $\zeta \in \{i, j, k\}$. Note that, by Proposition 10.1.1, the third set of identities in (10.1.9) follows from the second. This leaves 6 straightforward calculations for the reader to check. $\square$

Remark. In the case that $\xi = u, \eta = v$, and $\zeta = w$ are purely vectorial, we have

(10.1.10)
$$\begin{aligned} w(uv) &= w(-u \cdot v + u \times v) = -(u \cdot v)w - w \cdot (u \times v) + w \times (u \times v), \\ (wu)v &= (-w \cdot u + w \times u)v = -(w \cdot u)v - (w \times u) \cdot v + (w \times u) \times v. \end{aligned}$$

Then the identity of the two left sides is equivalent to the pair of identities

$$w \cdot (u \times v) = (w \times u) \cdot v, \tag{10.1.11}$$

$$w \times (u \times v) - (w \times u) \times v = (u \cdot v)w - (w \cdot u)v. \tag{10.1.12}$$

The identity (10.1.12) also follows from the pair of identities

$$w \times (u \times v) - (w \times u) \times v = (v \times w) \times u, \tag{10.1.13}$$

and

$$(v \times w) \times u = (u \cdot v)w - (w \cdot u)v. \tag{10.1.14}$$

See the exercises below for more on this.

In addition to the product, we also have a conjugation operation on $\mathbb{H}$:

$$\overline{\xi} = a - bi - cj - dk = a - u. \tag{10.1.15}$$

A calculation gives

$$\xi\overline{\eta} = (ab + u \cdot v) - av + bu - u \times v. \tag{10.1.16}$$

In particular,

$$\mathrm{Re}(\xi\overline{\eta}) = \mathrm{Re}(\overline{\eta}\xi) = (\xi, \eta), \tag{10.1.17}$$

the right side denoting the Euclidean inner product on $\mathbb{R}^4$. Setting $\eta = \xi$ in (10.1.16) gives

$$\xi\overline{\xi} = |\xi|^2, \tag{10.1.18}$$

the Euclidean square-norm of ξ. In particular, whenever $\xi \in \mathbb{H}$ is nonzero, it has a multiplicative inverse,

$$\xi^{-1} = |\xi|^{-2}\overline{\xi}. \tag{10.1.19}$$

We say a ring $\mathcal{R}$ with unit 1 is a division ring if each nonzero $\xi \in \mathcal{R}$ has a multiplicative inverse. It follows from (10.1.19) that $\mathbb{H}$ is a division ring. It is not a field, since multiplication in $\mathbb{H}$ is not commutative. Sometimes $\mathbb{H}$ is called a "noncommutative field."

To continue with products and conjugation, a routine calculation gives

$$\overline{\xi\eta} = \overline{\eta}\,\overline{\xi}. \tag{10.1.20}$$

Hence, via the associative law,

$$|\xi\eta|^2 = (\xi\eta)(\overline{\xi\eta}) = \xi\eta\overline{\eta}\overline{\xi} = |\eta|^2\xi\overline{\xi} = |\xi|^2|\eta|^2, \tag{10.1.21}$$

or

$$|\xi\eta| = |\xi|\,|\eta|. \tag{10.1.22}$$

Note that $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$ sits in $\mathbb{H}$ as a commutative subring, for which the properties (10.1.18) and (10.1.22) are familiar.

Let us examine (10.1.22) when $\xi = u$ and $\eta = v$ are purely vectorial. We have

$$uv = -u \cdot v + u \times v. \tag{10.1.23}$$

Hence, directly,

$$|uv|^2 = (u \cdot v)^2 + |u \times v|^2, \tag{10.1.24}$$

while (10.1.22) implies

$$|uv|^2 = |u|^2|v|^2. \tag{10.1.25}$$

On the other hand, if θ is the angle between u and v in $\mathbb{R}^3$,

$$u \cdot v = |u|\,|v|\,\cos\theta.$$

Hence (10.1.24) implies

$$|u \times v|^2 = |u|^2|v|^2\sin^2\theta. \tag{10.1.26}$$

We next consider the set of unit quaternions:

$$Sp(1) = \{\xi \in \mathbb{H} : |\xi| = 1\}. \tag{10.1.27}$$

Using (10.1.19) and (10.1.22), we see that $Sp(1)$ is a group under multiplication. It sits in $\mathbb{R}^4$ as the unit sphere S^3. We compare $Sp(1)$ with the group $SU(2)$, consisting of 2×2 complex matrices of the form

$$U = \begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix}, \quad \xi, \eta \in \mathbb{C}, \quad |\xi|^2 + |\eta|^2 = 1. \tag{10.1.28}$$

The group $SU(2)$ is also in natural one-to-one correspondence with S^3. Furthermore, we have:

Proposition 10.1.3. *The groups $SU(2)$ and $Sp(1)$ are isomorphic under map $\sigma : SU(2) \to Sp(1)$ defined by*

$$\sigma(U) = \xi + j\eta, \tag{10.1.29}$$

for U as in (10.1.28).

Proof. The map (10.1.29) is clearly bijective. To see it is a homomorphism of groups, we calculate:

$$\begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix} \begin{pmatrix} \xi' & -\overline{\eta}' \\ \eta' & \overline{\xi}' \end{pmatrix} = \begin{pmatrix} \xi\xi' - \overline{\eta}\eta' & -\xi\overline{\eta}' - \overline{\eta}\overline{\xi}' \\ \eta\xi' + \overline{\xi}\eta' & -\eta\overline{\eta}' + \overline{\xi}\overline{\xi}' \end{pmatrix}, \tag{10.1.30}$$

given $\xi, \eta \in \mathbb{C}$. Noting that, for $a, b \in \mathbb{R}$, $j(a + bi) = (a - bi)j$, we have

$$\begin{aligned} (\xi + j\eta)(\xi' + j\eta') &= \xi\xi' + \xi j\eta' + j\eta\xi' + j\eta j\eta' \\ &= \xi\xi' - \overline{\eta}\eta' + j(\eta\xi' + \overline{\xi}\eta'). \end{aligned} \tag{10.1.31}$$

Comparison of (10.1.30) and (10.1.31) verifies that (10.1.29) is a homomorphism of groups. ☐

We next define the map

$$\pi : Sp(1) \longrightarrow \mathcal{L}(\mathbb{R}^3) \tag{10.1.32}$$

by

$$\pi(\xi)u = \xi u \xi^{-1} = \xi u \overline{\xi}, \quad \xi \in Sp(1),\ u \in \mathbb{R}^3 \subset \mathbb{H}. \tag{10.1.33}$$

To justify (10.1.32), we need to show that if u is purely vectorial, so is $\xi u \overline{\xi}$. In fact, by (10.1.20),

$$\zeta = \xi u \overline{\xi} \Longrightarrow \overline{\zeta} = \overline{\overline{\xi}}\,\overline{u}\,\overline{\xi} = -\xi u \overline{\xi} = -\zeta, \tag{10.1.34}$$

so that is indeed the case. By (10.1.22),

$$|\pi(\xi)u| = |\xi|\,|u|\,|\overline{\xi}| = |u|, \quad \forall\, u \in \mathbb{R}^3,\ \xi \in Sp(1),$$

so in fact

$$\pi : Sp(1) \longrightarrow SO(3), \tag{10.1.35}$$

and it follows easily from the definition (10.1.33) that if also $\zeta \in Sp(1)$, then $\pi(\xi\zeta) = \pi(\xi)\pi(\zeta)$, so (10.1.35) is a group homomorphism. It is readily verified that

$$\operatorname{Ker} \pi = \{\pm 1\}. \tag{10.1.36}$$

Note that we can extend (10.1.32) to

$$\pi : Sp(1) \longrightarrow \mathcal{L}(\mathbb{H}), \quad \pi(\xi)\eta = \xi\eta\overline{\xi}, \quad \xi \in Sp(1),\ \eta \in \mathbb{H}, \tag{10.1.37}$$

and again $\pi(\xi\zeta) = \pi(\xi)\pi(\zeta)$ for $\xi, \zeta \in Sp(1)$. Furthermore, each map $\pi(\xi)$ is a ring homomorphism, i.e.,

$$\pi(\xi)(\alpha\beta) = (\pi(\xi)\alpha)(\pi(\xi)\beta), \quad \alpha, \beta \in \mathbb{H},\ \xi \in Sp(1). \tag{10.1.38}$$

Since $\pi(\xi)$ is invertible, this is a group of ring automorphisms of $\mathbb{H}$.

Here is another presentation of this automorphism group. Define

$$\tilde{\pi} : SO(3) \longrightarrow \mathcal{L}(\mathbb{H}), \quad \tilde{\pi}(T)(a + u) = a + Tu, \tag{10.1.39}$$

given $a + u \in \mathbb{H},\ a \in \mathbb{R},\ u \in \mathbb{R}^3$. It is a consequence of the identity

$$T(u \times v) = Tu \times Tv, \text{ for } u, v \in \mathbb{R}^3,\ T \in SO(3),$$

(cf. Exercise 3 below) that

$$\tilde{\pi}(T)(\alpha\beta) = (\tilde{\pi}(T)\alpha)(\tilde{\pi}(T)\beta), \quad \alpha, \beta \in \mathbb{H},\ T \in SO(3). \tag{10.1.40}$$

Thus $SO(3)$ acts as a group of automorphisms of $\mathbb{H}$. (Note that Proposition 10.1.1 is a special case of this.) We claim this is the same group of automorphisms as described in (10.1.37)–(10.1.38), via (10.1.35). This is a consequence of the fact that π in (10.1.35) is surjective. We mention that the first automorphism K mentioned in Proposition 10.1.1 has the form (10.1.37) with

$$\xi = \frac{1}{2}(1 + i + j + k).$$

We return to a construction used in Proposition 10.1.3 and make some further observations. Thus, with $\alpha_\nu, \beta_\nu \in \mathbb{C}$, set

$$\xi = \alpha_1 + j\alpha_2, \quad \eta = \beta_1 + j\beta_2. \tag{10.1.41}$$

We have a $\mathbb{C}$-linear isomorphism

$$\gamma : \mathbb{H} \longrightarrow \mathbb{C}^2, \quad \gamma(\xi) = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \tag{10.1.42}$$

where $\mathbb{C}$ acts on $\mathbb{H}$ on the *right*. We set

$$\operatorname{Co} \xi = \alpha_1, \quad \operatorname{Sp} \xi = \alpha_2. \tag{10.1.43}$$

Note that

$$\begin{aligned} \overline{\eta}\xi &= (\overline{\beta}_1 - \overline{\beta}_2 j)(\alpha_1 + j\alpha_2) \\ &= \alpha_1\overline{\beta}_1 + \alpha_2\overline{\beta}_2 - j(\alpha_1\beta_2 - \alpha_2\beta_1). \end{aligned} \tag{10.1.44}$$

Hence

$$\begin{aligned} \mathrm{Co}(\overline{\eta}\xi) &= \alpha_1\overline{\beta}_1 + \alpha_2\overline{\beta}_2 \\ &= ((\gamma(\xi), \gamma(\eta)\,)), \end{aligned} \tag{10.1.45}$$

where $((\ ,\ ))$ denotes the standard Hermitian inner product on $\mathbb{C}^2$, and

$$\begin{aligned} \mathrm{Sp}(\overline{\eta}\xi) &= -(\alpha_1\beta_2 - \alpha_2\beta_1) \\ &= -\sigma(\gamma(\xi), \gamma(\eta)), \end{aligned} \tag{10.1.46}$$

where

$$\sigma((\alpha_1, \alpha_2), (\beta_1, \beta_2)) = \alpha_1\beta_2 - \alpha_2\beta_1 \tag{10.1.47}$$

defines σ as an antisymmetric, $\mathbb{C}$-bilinear form on $\mathbb{C}^2$, called the symplectic form.

Exercises

Exercises 1–8 deal with cross products of vectors in $\mathbb{R}^3$, and relate to the identities (10.1.10)–(10.1.14), which in turn relate to associativity of the product on $\mathbb{H}$.

1. If $u, v \in \mathbb{R}^3$, we define the cross product $u \times v = \Pi(u, v)$ to be the unique bilinear map $\Pi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ satisfying

$$u \times v = -v \times u, \quad \text{and}$$
$$i \times j = k, \quad j \times k = i, \quad k \times i = j,$$

where $\{i, j, k\}$ is the standard basis of $\mathbb{R}^3$. Show that, for all $u, v, w \in \mathbb{R}^3$,

$$w \cdot (u \times v) = \det \begin{pmatrix} w_1 & u_1 & v_1 \\ w_2 & u_2 & v_2 \\ w_3 & u_3 & v_3 \end{pmatrix}. \tag{10.1.48}$$

and show that this property uniquely specifies $u \times v$.

2. Deduce from (10.1.48) that for $u, v, w \in \mathbb{R}^3$,

$$u \cdot (v \times w) = (u \times v) \cdot w.$$

This establishes (10.1.11).

3. Recall that $T \in SO(3)$ provided that $T \in M(3, \mathbb{R})$ satisfies $T^t T = I$ and $\det T = 1$. Show that

$$T \in SO(3) \Longrightarrow Tu \times Tv = T(u \times v). \tag{10.1.49}$$

Hint. Multiply the 3×3 matrix in (10.1.48) on the left by T.

4. Generalize the identity

$$\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin \theta|$$

to the following, for $u, v, w, x \in \mathbb{R}^3$:

$$\begin{aligned} (u \times v) \cdot (w \times x) &= (u \cdot w)(v \cdot x) - (u \cdot x)(v \cdot w) \\ &= \det \begin{pmatrix} u \cdot w & u \cdot x \\ v \cdot w & v \cdot x \end{pmatrix}. \end{aligned} \tag{10.1.50}$$

Hint. Check this for $u = i,\ v = ai + bj$, in which $u \times v = bk$, and use Exercise 3 to show this suffices. Note that the left side of (10.1.50) is then

$$bk \cdot (w \times x) = \det \begin{pmatrix} 0 & w \cdot i & x \cdot i \\ 0 & w \cdot j & x \cdot j \\ b & w \cdot k & x \cdot k \end{pmatrix}.$$

Show that this equals the right side of (10.1.50).

5. Show that $\kappa : \mathbb{R}^3 \to \mathcal{L}(\mathbb{R}^3)$, given by

$$\kappa(y)x = y \times x, \tag{10.1.51}$$

has the matrix representation

$$\kappa(y) = \begin{pmatrix} 0 & -y_3 & y_2 \\ y_3 & 0 & -y_1 \\ -y_2 & y_1 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \tag{10.1.52}$$

so that actually $\kappa : \mathbb{R}^3 \to \mathrm{Skew}(3)$, the space of antisymmetric 3×3 matrices. Show that, with $[A, B] = AB - BA$,

$$\begin{aligned} \kappa(x \times y) &= [\kappa(x), \kappa(y)], \\ \mathrm{Tr}(\kappa(x)\kappa(y)^t) &= 2x \cdot y. \end{aligned} \tag{10.1.53}$$

6. Show that if $u, v, w \in \mathbb{R}^3$, then the first part of (10.1.53) implies

$$(u \times v) \times w = u \times (v \times w) - v \times (u \times w).$$

Relate this to the Jacobi identity

$$[[A, B], C] = [A, [B, C]] - [B, [A, C]],$$

for $A, B, C \in M(n, \mathbb{R})$ (with $n = 3$). This establishes (10.1.13).

7. Show that if $u, v, w \in \mathbb{R}^3$,

$$v \times (u \times w) = (v \cdot w)u - (v \cdot u)w.$$

This establishes (10.1.14).
Hint. Start with the observation that $v \times (u \times w)$ is in $\mathrm{Span}\{u, w\}$ and is orthogonal to v.
Alternative. Use Exercise 3 to reduce the calculation to the case $u = i,\ w = ai + bj$.

8. As noted before, the identity (10.1.12) follows from the pair of identities (10.1.13)–(10.1.14). Establish the converse:

$$(10.1.12) \Longrightarrow (10.1.13) \quad \text{and} \quad (10.1.14).$$

Hint. To start, given (10.1.12), permute letters to supplement this with

$$\begin{aligned} w \times (v \times u) - (w \times v) \times u &= (v \cdot u)w - (w \cdot v)u, \\ u \times (w \times v) - (u \times w) \times v &= (w \cdot v)u - (u \cdot w)v. \end{aligned} \tag{10.1.54}$$

Then add (10.1.54) to (10.1.12) to obtain (10.1.14).

Exercises 9–12 relate to the representation π of $Sp(1)$ on $\mathbb{R}^3$ defined by (10.1.32) and its derived representation.

9. Note that

$$T_1 Sp(1) = \operatorname{Im}\mathbb{H} = \mathbb{R}^3. \tag{10.1.55}$$

Show that if $u \in \mathbb{R}^3$,

$$\pi(e^{tu})x = e^{tu} x e^{-tu}, \tag{10.1.56}$$

and hence

$$d\pi(u)x = ux - xu. \tag{10.1.57}$$

10. Using (10.1.6) and Exercise 5, show that

$$d\pi(u) = 2\kappa(u). \tag{10.1.58}$$

11. Show that

$$\kappa(k) = \begin{pmatrix} J & \\ & 0 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{10.1.59}$$

and that if $T \in SO(3)$, Exercise 3 yields

$$\begin{aligned} u = Tk &\Rightarrow \kappa(u) = T\kappa(k)T^{-1} \\ &\Rightarrow e^{s\kappa(u)} = T \begin{pmatrix} e^{sJ} & \\ & 1 \end{pmatrix} T^{-1}. \end{aligned} \tag{10.1.60}$$

12. Deduce that

$$\pi(e^{tu}) = T \begin{pmatrix} e^{2tJ} & \\ & 1 \end{pmatrix} T^{-1}, \tag{10.1.61}$$

and relate the factor of 2 in (10.1.58) and in (10.1.61), hence in (10.1.6), to the fact that (10.1.35) is a 2-to-1 map.

13. Show that there is no "product" on $\mathbb{R}^3$, i.e., bilinear map $P : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$, with the property that

$$x, y \in \mathbb{R}^3 \setminus 0 \Longrightarrow P(x, y) \neq 0.$$

Hint. Define $L(x) \in M(3, \mathbb{R})$ by

$$L(x)y = P(x, y).$$

Note that for each $x \in \mathbb{R}^3 \setminus 0$, $L(-x) = -L(x)$, and deduce that (since 3 is odd),

$$\det L(-x) = -\det L(x).$$

Thus any curve connecting x to $-x$ in $\mathbb{R}^3 \setminus 0$ must contain a point z such that $\det L(z) = 0$.

10.2. Quaternionic linear algebra

A quaternionic vector space V is a real vector space, equipped with an $\mathbb{R}$-bilinear map $s : V \times \mathbb{H} \to V$, given by $s(v, a) = va$ (so quaternionic scalars act on the *right*), satisfying

$$v1 = v, \quad (va)b = v(ab), \quad \forall\, v \in V,\ a, b \in \mathbb{H}. \tag{10.2.1}$$

Note that $\mathbb{R}$-bilinearity then implies $va = av$ for all $a \in \mathbb{R}$. The prime example of such a space is $\mathbb{H}^n$, the space of n-tuples $(a_1, \dots, a_n)^t$, $a_j \in \mathbb{H}$, introduced in §1.2.

If V and W are quaternionic vector spaces, we say an $\mathbb{R}$-linear transformation $T : V \to W$ is $\mathbb{H}$-linear provided it commutes with scalar multiplication, i.e., given $T \in \mathcal{L}_{\mathbb{R}}(V, W)$,

$$T \in \mathcal{L}_{\mathbb{H}}(V, W) \Leftrightarrow T(va) = (Tv)a,\ \forall\, v \in V,\ a \in \mathbb{H}. \tag{10.2.2}$$

For such T, we see that $T(v_1a_1 + v_2a_2) = (Tv_1)a_1 + (Tv_2)a_2$, for $v_j \in V,\ a_j \in \mathbb{H}$. If $V = W$, we denote this space by $\mathcal{L}_{\mathbb{H}}(V)$. As noted in §1.2, if $V = \mathbb{H}^n$ and $A = (a_{jk}) \in M(n, \mathbb{H})$, then the standard matrix product

$$(Av)_j = \sum_k a_{jk} v_k, \quad v = (v_1, \dots, v_n)^t \in \mathbb{H}^n, \tag{10.2.3}$$

yields $A \in \mathcal{L}_{\mathbb{H}}(\mathbb{H}^n)$.

An $\mathbb{R}$-linear subspace W of V is said to be an $\mathbb{H}$-linear subspace provided W is stable under quaternionic scalar multiplication, i.e., given $W \subset V$, $\mathbb{R}$-linear,

$$W \text{ is } \mathbb{H}\text{-linear} \Longleftrightarrow wa \in W, \ \ \forall\, w \in W,\ a \in \mathbb{H}. \tag{10.2.4}$$

Clearly such W also has the natural structure of a quaternionic vector space. Examples of $\mathbb{H}$-linear subspaces include

$$\begin{gathered} \mathcal{N}(S) = \{w \in V : Sw = 0\}, \quad \mathcal{R}(T) = \{Tu : u \in X\}, \\ S \in \mathcal{L}_{\mathbb{H}}(V, X), \quad T \in \mathcal{L}_{\mathbb{H}}(X, V). \end{gathered} \tag{10.2.5}$$

Given a set $S = \{v_1, \dots, v_n\} \subset V$, we define

$$\mathcal{J}_S : \mathbb{H}^n \longrightarrow V, \quad \mathcal{J}_S \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = v_1a_1 + \cdots + v_na_n, \tag{10.2.6}$$

so $\mathcal{J}_S \in \mathcal{L}_{\mathbb{H}}(\mathbb{H}^n, V)$. We say

$$\operatorname{Span}_{\mathbb{H}}(S) = \mathcal{R}(\mathcal{J}_S), \tag{10.2.7}$$

and

$$\begin{aligned} &S \text{ is } \mathbb{H}\text{-linearly independent} \Leftrightarrow \mathcal{N}(\mathcal{J}_S) = 0, \\ &S\ \mathbb{H}\text{-spans } V \Leftrightarrow \mathcal{R}(\mathcal{J}_S) = V, \\ &S \text{ is an } \mathbb{H}\text{-basis of } V \Leftrightarrow \mathcal{J}_S \text{ is bijective.} \end{aligned} \tag{10.2.8}$$

Clearly if $\dim_{\mathbb{R}} V < \infty$, V has a finite $\mathbb{H}$-spanning set (namely, any basis over $\mathbb{R}$), any minimal $\mathbb{H}$-spanning subset of which is an $\mathbb{H}$-basis of V. If S as above is an $\mathbb{H}$-basis of V, we say $\dim_{\mathbb{H}} V = n$. Given V isomorphic to $\mathbb{H}^n$ over $\mathbb{H}$, it is a fortiori isomorphic to $\mathbb{H}^n$ over $\mathbb{R}$, so

$$\dim_{\mathbb{H}} V = n \Longleftrightarrow \dim_{\mathbb{R}} V = 4n. \tag{10.2.9}$$

In particular, standard linear algebra over $\mathbb{R}$ implies $\dim_{\mathbb{H}} V$ is well defined. Given this observation, the following is a simple variant of the $\mathbb{R}$-linear analogue.

Proposition 10.2.1. *Given quaternionic vector spaces V, W such that $\dim V < \infty$, and $T \in \mathcal{L}_{\mathbb{H}}(V, W)$,*

$$\dim_{\mathbb{H}} \mathcal{N}(T) + \dim_{\mathbb{H}} \mathcal{R}(T) = \dim_{\mathbb{H}} V. \tag{10.2.10}$$

Proof. The standard result is $\dim_{\mathbb{R}} \mathcal{N}(T) + \dim_{\mathbb{R}} \mathcal{R}(T) = \dim_{\mathbb{R}} V$ (cf. [**42**], §1.3), and (10.2.10) follows from this in view of (10.2.9). ☐

Quaternionic inner product spaces

A quaternionic inner product space is a quaternionic vector space V, equipped with an $\mathbb{R}$-bilinear map $Q : V \times V \to \mathbb{H}$, denoted $\langle u, v\rangle$, satisfying, for all $u, v \in V,\ a \in \mathbb{H}$,

$$\begin{aligned} \langle ua, v\rangle &= \langle u, v\rangle a, \\ \langle v, u\rangle &= \overline{\langle u, v\rangle}, \\ \langle u, u\rangle &> 0, \quad \text{if} \quad u \neq 0. \end{aligned} \tag{10.2.11}$$

A direct consequence of the first two identities is

$$\langle u, va\rangle = \overline{a}\langle u, v\rangle. \tag{10.2.12}$$

The standard example is the inner product on $\mathbb{H}^n$, introduced in §1.2,

$$\langle u, v\rangle = \sum_{k=1}^{n} \overline{v}_k u_k, \tag{10.2.13}$$

for $u = (u_1, \dots, u_n)^t,\ v = (v_1, \dots, v_n)^t \in \mathbb{H}^n$.

A set $S = \{w_1, \dots, w_k\} \subset V$ is said to be an orthonormal set provided

$$\langle w_j, w_\ell\rangle = \delta_{j\ell}, \quad \forall\, j, \ell \in \{1, \dots, k\}. \tag{10.2.14}$$

It readily follows from (10.2.14) that

$$w = \sum w_j a_j \Longrightarrow \langle w, w\rangle = \sum |a_j|^2. \tag{10.2.15}$$

Hence S is linearly independent. Suppose S spans $W \subset V$ (over $\mathbb{H}$). We associate a transformation $P_W \in \mathcal{L}_{\mathbb{H}}(V)$, defined as follows:

$$P_W v = \sum_{j=1}^{k} w_j \langle v, w_j\rangle, \quad v \in V. \tag{10.2.16}$$

Note that, for $1 \le \ell \le k$,

$$\langle P_W v, w_\ell\rangle = \sum_j \langle w_j, w_\ell\rangle\langle v, w_j\rangle = \langle v, w_\ell\rangle, \tag{10.2.17}$$

hence

$$\langle P_W v, w_\ell a_\ell\rangle = \overline{a}_\ell \langle v, w_\ell\rangle = \langle v, w_\ell a_\ell\rangle, \tag{10.2.18}$$

so

$$w \in W \Longrightarrow \langle P_W v, w\rangle = \langle v, w\rangle. \tag{10.2.19}$$

Hence

$$v \in V, \ w \in W \Longrightarrow \langle (I - P_W)v, w\rangle = 0. \tag{10.2.20}$$

Generally, if W is an $\mathbb{H}$-linear subspace of V, we set

$$W^\perp = \{v \in V : \langle v, w\rangle = 0, \quad \forall\, w \in W\}. \tag{10.2.21}$$

It follows from (10.2.11) that $W^\perp$ is an $\mathbb{H}$-linear subspace of V. The content of (10.2.16) and (10.2.20) is that

$$P_W : V \longrightarrow W, \quad I - P_W : V \longrightarrow W^\perp, \tag{10.2.22}$$

under the hypothesis that W has the orthonormal basis $\{w_1, \dots, w_k\}$. The following is a useful complement to (10.2.22).

Proposition 10.2.2. *Let W be an $\mathbb{H}$-linear subspace of V. Take $v \in V$, and assume*

$$v = w_1 + x_1 = w_2 + x_2, \quad w_j \in W, \ x_j \in W^\perp. \tag{10.2.23}$$

Then

$$w_1 = w_2, \quad x_1 = x_2. \tag{10.2.24}$$

Proof. The hypothesis implies

$$w_1 - w_2 = x_2 - x_1 \in W \cap W^\perp = 0, \tag{10.2.25}$$

the last identity by the third item in (10.2.11). $\square$

We are in a position to establish the following.

Proposition 10.2.3. *Let W be an $\mathbb{H}$-linear subspace of V, and assume $\dim_{\mathbb{H}} V = n < \infty$. Assume $\{w_1, \dots, w_k\}$ is an orthonormal basis of W. Then this extends to an orthonormal basis $\{w_1, \dots, w_n\}$ of V.*

Proof. If $W \neq V$, take $v \in V,\ v \notin W$. Then

$$\tilde{w}_{k+1} = (I - P_W)v \in W^\perp, \text{ nonzero.} \tag{10.2.26}$$

Setting

$$w_{k+1} = \frac{1}{a}\tilde{w}_{k+1}, \quad a = \langle \tilde{w}_{k+1}, \tilde{w}_{k+1} \rangle^{1/2} \tag{10.2.27}$$

gives a unit vector $w_{k+1} \in W^\perp$, so we have

$$\{w_1, \dots, w_k, w_{k+1}\}, \tag{10.2.28}$$

an orthonormal set, spanning $W_1 \subset V$, of $\mathbb{H}$-dimension $k+1$. An inductive argument finishes the proof. □

Corollary 10.2.4. *If V is a finite dimensional quaternionic inner product space, of dimension $n \geq 1$, then V has an orthonormal basis.*

Proof. Pick a nonzero $v \in V$, set $w_1 = v/\langle v, v \rangle^{1/2}$, spanning W, and apply Proposition 10.2.3. □

Here is one useful application of an orthonormal basis.

Proposition 10.2.5. *If V is a quaternionic inner product space with the orthonormal basis $S = \{v_1, \dots, v_n\}$, then* (10.2.6) *provides an $\mathbb{H}$-linear isomorphism*

$$\mathcal{J}_S : \mathbb{H}^n \xrightarrow{\approx} V, \tag{10.2.29}$$

preserving the inner products, i.e.,

$$\langle \xi, \eta \rangle = \langle \mathcal{J}_S \xi, \mathcal{J}_S \eta \rangle, \quad \forall\, \xi, \eta \in \mathbb{H}^n, \tag{10.2.30}$$

where the left side uses the inner product (10.2.13) *on $\mathbb{H}^n$, and the right side uses the given inner product on V.*

Proof. We have

$$\begin{aligned} \langle \mathcal{J}_S \xi, \mathcal{J}_S \eta \rangle &= \sum_{j,k=1}^{n} \langle v_j \xi_j, v_k \eta_k \rangle = \sum_{j,k=1}^{n} \overline{\eta}_k \langle v_j, v_k \rangle \xi_j \\ &= \sum_{k=1}^{n} \overline{\eta}_k \xi_k = \langle \xi, \eta \rangle. \end{aligned} \tag{10.2.31}$$

□

Eigenvectors

If $T \in \mathcal{L}_{\mathbb{H}}(V)$, we say a nonzero element $v \in V$ is an eigenvector of T, with eigenvalue $\lambda \in \mathbb{H}$, if

$$Tv = v\lambda. \tag{10.2.32}$$

Note that if this holds, then

$$\xi \in \mathbb{H},\ |\xi| = 1 \Rightarrow T(v\xi) = (Tv)\xi = v\lambda\xi = (v\xi)\overline{\xi}\lambda\xi, \tag{10.2.33}$$

so $v\xi$ is an eigenvector of T, with eigenvalue $\overline{\xi}\lambda\xi$. If $\lambda \in \mathbb{R}$, $\overline{\xi}\lambda\xi \equiv \lambda$. More generally,

$$\lambda = a + \eta, \quad a \in \mathbb{R},\ \eta \in \operatorname{Im}\mathbb{H} \Rightarrow \overline{\xi}\lambda\xi = a + \overline{\xi}\eta\xi, \tag{10.2.34}$$

and the action of $Sp(1)$ on $\operatorname{Im}\mathbb{H}$ covers the $SO(3)$ action, as seen in §10.1. In particular, there exists $\xi \in Sp(1)$ such that $v\xi$ is an eigenvector with eigenvalue in $\mathbb{C} \subset \mathbb{H}$. This observation sets us up to establish the following.

Proposition 10.2.6. *If V is a quaternionic vector space, of dimension $n \in \mathbb{N}$, then each $T \in \mathcal{L}_{\mathbb{H}}(V)$ has an eigenvector.*

Proof. Restrict attention to scalars in $\mathbb{C} \subset \mathbb{H}$. Then V is a complex vector space (of complex dimension $2n$), and $T \in \mathcal{L}_{\mathbb{C}}(V)$. Standard complex linear algebra (cf. [**42**], §2.1) implies there is an eigenvector v such that $Tv = v\lambda$, with $\lambda \in \mathbb{C}$. $\square$

REMARK. The reader should check this proposition out when $n = 1$, $V = \mathbb{H}$, $\xi \in \mathbb{H}$ is given, and $Tv = \xi v$.

Note that

$$Tv = v\lambda \Longrightarrow T^k v = v\lambda^k, \quad \forall\, k \in \mathbb{N}. \tag{10.2.35}$$

In particular, if $T \in \mathcal{L}_{\mathbb{H}}(V)$ is nilpotent, its only eigenvalue is 0, so if such $T \neq 0$, V does not have a basis of eigenvectors of T.

On the other hand, suppose V does have a basis $S = \{v_1, \dots, v_n\}$ of eigenvectors of $T \in \mathcal{L}_{\mathbb{H}}(V)$,

$$Tv_j = v_j\lambda_j. \tag{10.2.36}$$

Define $\mathcal{J}_S : \mathbb{H}^n \overset{\approx}{\to} V$ as in (10.2.6), so $\mathcal{J}_S e_j = v_j$, where $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{H}^n$. Then

$$(\mathcal{J}_S^{-1} T \mathcal{J}_S)e_j = \mathcal{J}_S^{-1} v_j\lambda_j = e_j\lambda_j = \lambda_j e_j, \tag{10.2.37}$$

the last identity holding since $\mathbb{H}$ acts on both the left and the right of $\mathbb{H}^n$, and the only nonzero entry of e_j is 1, which is real. This gives the following result.

Proposition 10.2.7. *Assume $T \in \mathcal{L}_{\mathbb{H}}(V)$ and that V has a basis $S = \{v_j\}$ of eigenvectors of T, satisfying* (10.2.36). *Then $\mathcal{J}_S^{-1} T \mathcal{J}_S \in \mathcal{L}_{\mathbb{H}}(\mathbb{H}^n)$ is given by the diagonal matrix*

$$\mathcal{J}_S^{-1} T \mathcal{J}_S = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}. \tag{10.2.38}$$

Remark. Keep in mind that each eigenvector v_j can be scaled to arrange that $\lambda_j \in \mathbb{C}$.

Adjoints

If V is a quaternionic inner product space and $T \in \mathcal{L}_{\mathbb{H}}(V)$, its adjoint $T^* \in \mathcal{L}_{\mathbb{H}}(V)$ is characterized by

$$\langle u, T^* v\rangle = \langle Tu, v\rangle, \quad u, v \in V. \tag{10.2.39}$$

In case $V = \mathbb{H}^n$, with the inner product (10.2.13), and $A \in \mathcal{L}_{\mathbb{H}}(\mathbb{H}^n)$ is given by (10.2.3), then, as seen in §1.2,

$$A^* = (\overline{a}_{kj}). \tag{10.2.40}$$

Also, if $\mathcal{J}_S : \mathbb{H}^n \to V$ is an isomorphism preserving inner products, as in (10.2.30), then

$$T = \mathcal{J}_S A \mathcal{J}_S^{-1} \Longrightarrow T^* = \mathcal{J}_S A^* \mathcal{J}_S^{-1}. \tag{10.2.41}$$

The following is an important property of adjoints.

Proposition 10.2.8. *Let $T \in \mathcal{L}_{\mathbb{H}}(V)$, and assume $W \subset V$ is an $\mathbb{H}$-linear subspace. Then*

$$T : W \to W \Longrightarrow T^* : W^\perp \to W^\perp. \tag{10.2.42}$$

Proof. Let $w \in W,\ x \in W^\perp$, and assume $T : W \to W$. Then

$$\langle w, T^* x\rangle = \langle Tw, x\rangle = 0. \tag{10.2.43}$$

□

Corollary 10.2.9. *Let $T \in \mathcal{L}_{\mathbb{H}}(V)$ and assume T is either self adjoint ($T^* = T$) or skew adjoint ($T^* = -T$). Let $W \subset V$ be an $\mathbb{H}$-linear subspace. Then*

$$T : W \to W \Longrightarrow T : W^\perp \to W^\perp. \tag{10.2.44}$$

Special classes of T yielding orthonormal bases of eigenvectors

It is an easy step to proceed from Corollary 10.2.9 to the following.

Proposition 10.2.10. *Let V be a finite dimensional quaternionic inner product space. Let $T \in \mathcal{L}_{\mathbb{H}}(V)$ and assume T is either self adjoint or skew adjoint. Then V has an orthonormal basis of eigenvectors of T.*

Proof. By Proposition 10.2.6, T has an eigenvector, say v_1, so $Tv_1 = v_1\lambda_1$. We can scale v_1 to be a unit vector. If v_1 spans V, we are done. Otherwise, if we apply Corollary 10.2.9 to $W = \mathrm{Span}_{\mathbb{H}}(v_1)$, we get $T : W^{\perp} \to W^{\perp}$, and $T|_{W^{\perp}}$ has an eigenvector v_2, so $Tv_2 = v_2\lambda_2$, and $v_2 \perp v_1$. If $\mathrm{Span}_{\mathbb{H}}(v_1, v_2) = V$ we are done, otherwise apply Corollary 10.2.9 again. An inductive argument finishes the proof. ☐

Note that if $T = T^*,\ Tv = v\lambda,\ v \neq 0$, then

$$\begin{aligned} \langle v, v\rangle\lambda = \langle v\lambda, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle \\ = \langle v, v\lambda\rangle = \overline{\lambda}\langle v, v\rangle, \end{aligned} \tag{10.2.45}$$

hence

$$\lambda = \overline{\lambda}, \quad \text{so} \quad \lambda \in \mathbb{R}. \tag{10.2.46}$$

On the other hand, if $T = -T^*,\ Tv = v\lambda,\ v \neq 0$, then

$$\begin{aligned} \langle v, v\rangle\lambda = \langle v\lambda, v\rangle = \langle Tv, v\rangle = -\langle v, Tv\rangle \\ = -\langle v, v\lambda\rangle = -\overline{\lambda}\langle v, v\rangle, \end{aligned} \tag{10.2.47}$$

hence

$$\lambda = -\overline{\lambda}, \quad \text{so} \quad \lambda \in \mathrm{Im}\,\mathbb{H}. \tag{10.2.48}$$

By the remarks around (10.2.33)–(10.2.34), we can scale v to arrange that $\lambda \in \mathbb{C}$, hence

$$\lambda = i\mu, \quad \mu \in \mathbb{R}, \tag{10.2.49}$$

in this case.

The group Sp(V)

We now consider the quaternionic "unitary" case. Let V be a finite dimensional quaternionic inner product space. Given $T \in \mathcal{L}_{\mathbb{H}}(V)$, we say

$$T \in Sp(V) \Longleftrightarrow T^*T = I. \tag{10.2.50}$$

This condition implies that $T : V \to V$ is injective, hence, by (10.2.10), surjective, hence bijective, so

$$T \in Sp(V) \Longrightarrow T^* = T^{-1} \in \mathcal{L}_{\mathbb{H}}(V). \tag{10.2.51}$$

In case $V = \mathbb{H}^n$, with its standard inner product, we have $Sp(\mathbb{H}^n) = Sp(n)$, the group introduced in §1.2, and the major focus of this chapter.

We have the following variant of Proposition 10.2.8.

Proposition 10.2.11. *Let $T \in Sp(V)$. Let $W \subset V$ be an $\mathbb{H}$-linear subspace. Then*

$$T : W \to W \Longrightarrow T : W^\perp \to W^\perp. \tag{10.2.52}$$

Proof. We see that $T|_W \in \mathcal{L}_{\mathbb{H}}(W)$ is injective, hence surjective, so $T^{-1} \in \mathcal{L}_{\mathbb{H}}(V)$ satisfies $T^* = T^{-1} : W \to W$. The fact that $T : W^\perp \to W^\perp$ then follows from Proposition 10.2.8, with T replaced by T^*. □

Having this, we proceed to the following, by the same arguments as used in Proposition 10.2.10.

Proposition 10.2.12. *Let $T \in Sp(V)$. Then V has an orthonormal basis of eigenvectors of T.*

Note that if $T^* = T^{-1},\ Tv = v\lambda,\ v \neq 0$, then

$$\begin{aligned} \langle v, v\rangle\lambda = \langle v\lambda, v\rangle = \langle Tv, v\rangle &= \langle v, T^{-1}v\rangle \\ = \langle v, v\lambda^{-1}\rangle &= \overline{\lambda}^{-1}\langle v, v\rangle, \end{aligned} \tag{10.2.53}$$

hence

$$\lambda = \overline{\lambda}^{-1}, \quad \text{so} \quad |\lambda|^2 = 1. \tag{10.2.54}$$

Again we can scale v to arrange that $\lambda \in \mathbb{C}$, hence

$$\lambda = e^{i\theta}, \quad \theta \in \mathbb{R}. \tag{10.2.55}$$

In light of Proposition 10.2.7, each $T \in Sp(n)$ is conjugate via an element of $Sp(n)$ to a diagonal matrix with entries of the form (10.2.55). Thus, in the terminology introduced in §6.A, we have the following.

Corollary 10.2.13. *The set of $\mathbb{T}$ diagonal matrices in $Sp(n)$ with diagonal elements of the form* (10.2.55) *is a conjugating torus (hence a maximal torus) for $Sp(n)$.*

Thus each $T \in Sp(n)$ is contained in a torus $J\mathbb{T}J^{-1}$, so T is connected to the identity $I \in Sp(n)$, via a continuous curve. This gives the following topological information.

Corollary 10.2.14. *For each $n \geq 1$, $Sp(n)$ is connected.*

Exercises.

Let V be an n-dimensional quaternionic vector space, with scalar action $v \mapsto v\lambda$, $v \in V$, $\lambda \in \mathbb{H}$, the action also denoted by $R(\lambda)$. Note that V is a $2n$-dimensional complex vector space if we restrict the scalars to $\mathbb{C}$.

1. Let $T \in \mathcal{L}_{\mathbb{H}}(V)$. Assume $v \in V$, $\lambda \in \mathbb{H}$, $v \neq 0$, and $Tv = v\lambda$. Say $\lambda = a + \eta$, as in (10.2.33)–(10.2.34), $a \in \mathbb{R}$, $\eta \in \operatorname{Im}\mathbb{H}$, and $\eta \neq 0$. Let

$$E(T, a \pm i|\eta|) = \{v \in V : Tv = v(a \pm i|\eta|)\}.$$

Each of these spaces is a $\mathbb{C}$-linear subspace of V. Show that we have $\mathbb{R}$-linear isomorphisms

$$R(j), R(k) : E(T, a \pm i|\eta|) \xrightarrow{\approx} E(T, a \mp i|\eta|).$$

Show that

$$\operatorname{Span}_{\mathbb{H}} E(T, a \pm i|\eta|) = E(T, a + i|\eta|) \oplus E(T, a - i|\eta|).$$

Denote this span

$$\mathcal{E}(T, \lambda).$$

If $\lambda = a \in \mathbb{R}$, set $\mathcal{E}(T, \lambda) = E(T, a)$.

2. Show that if $T \in \mathcal{L}_{\mathbb{H}}(V)$, $\lambda \in \mathbb{H}$,

$$Tv = v\lambda \Longrightarrow v \in \mathcal{E}(T, \lambda).$$

3. Under the hypotheses of problem 2, show that

$$\xi \in \mathbb{H},\ |\xi| = 1 \Longrightarrow \mathcal{E}(T, \lambda) = \mathcal{E}(T, \bar{\xi}\lambda\xi).$$

4. Show that if $T \in \mathcal{L}_{\mathbb{H}}(V)$, $\lambda \in \mathbb{H}$,

$$T : \mathcal{E}(T, \lambda) \longrightarrow \mathcal{E}(T, \lambda).$$

5. If $T \in \mathcal{L}_{\mathbb{C}}(V)$, let $\det_{\mathbb{C}} T$ denote the standard complex determinant (an element of $\mathbb{C}$). Show that

$$T \in \mathcal{L}_{\mathbb{H}}(V) \Longrightarrow \det_{\mathbb{C}} T \geq 0.$$

6. Show that if $T \in \mathcal{L}_{\mathbb{H}}(V)$, $\lambda \in \mathbb{H}$, $\lambda = a + \eta$, as in exercise 1,

$$\mathcal{E}(T, \lambda) = \operatorname{Ker}(T - a)^2 + |\eta|^2.$$

Hint. Simplify $(T - a - |\eta|R(i))(T - a + |\eta|R(i))$.
Note. $R(i)$ is not $\mathbb{H}$-linear on V, but it is $\mathbb{C}$-linear, and it commutes with T.

7. In this exercise, set

$$V = \mathbb{H}^2, \quad T = \begin{pmatrix} i & \\ & i \end{pmatrix}.$$

(a) Show that

$$E(T, i) = \mathrm{Span}_{\mathbb{C}}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}, \quad E(T, -i) = \mathrm{Span}_{\mathbb{C}}\left\{\begin{pmatrix} j \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ j \end{pmatrix}\right\}.$$

From Exercise 1 we obtain $\mathcal{E}(T, i) = E(T, i) \oplus E(T, -i)$, hence

$$\mathcal{E}(T, i) = V.$$

(b) Show that

$$v_0 = \begin{pmatrix} 1 \\ j \end{pmatrix} \text{ is not an eigenvector of } T.$$

8. Return to the setting of Exercise 1. Suppose $\dim_{\mathbb{H}} \mathcal{E}(T, \lambda) = k$.

(a) Show that a basis $\{v_1, \dots, v_k\}$ (over $\mathbb{C}$) of $E(t, a + i|\eta|)$ provides a basis (over $\mathbb{H}$) of $\mathcal{E}(T, \lambda)$, consisting of eigenvectors of T.

(b) Modify the basis of part (a) to obtain a basis $\{w_1, \dots, w_k\}$ of $\mathcal{E}(T, \lambda)$ satisfying

$$T w_j = w_j \lambda.$$

10.3. Roots, weights, and representations of $\mathbf{Sp}(1)$ and $\mathbf{Sp}(2)$

As seen in §10.1, the Lie algebra of $Sp(1)$ is

$$\mathfrak{sp}(1) = T_1 Sp(1) = \operatorname{Im}\mathbb{H} = \mathbb{R}^3, \tag{10.3.1}$$

with Lie bracket

$$[u, v] = uv - vu = 2u \times v. \tag{10.3.2}$$

Parallel to (4.1.2) in Chapter 4, we set

$$X_1 = \frac{1}{2}i, \quad X_2 = \frac{1}{2}j, \quad X_3 = \frac{1}{2}k, \tag{10.3.3}$$

with commutation relations

$$[X_1, X_2] = X_3, \quad [X_2, X_3] = X_1, \quad [X_3, X_1] = X_2, \tag{10.3.4}$$

yielding an isomorphism $\mathfrak{sp}(1) \approx \mathfrak{su}(2)$. In particular, if we take

$$\mathfrak{h} = \operatorname{Span} X_1, \tag{10.3.5}$$

we have root vectors:

$$[X_1, X_2 \pm \sqrt{-1}X_3] = \mp\sqrt{-1}(X_2 \pm \sqrt{-1}X_3). \tag{10.3.6}$$

Since here i denotes the first standard basis vector of $\mathbb{R}^3$, we will use $\sqrt{-1}$ to denote the complex number formerly known as i.

Now the treatment of roots, weights, and irreducible skew-adjoint representations of $\mathfrak{sp}(1)$ goes through exactly as that for $\mathfrak{su}(2)$ in Chapter 4. Furthermore, the isomorphism $\sigma : SU(2) \to Sp(1)$ of Proposition 10.1.3 yields the representation

$$\mathcal{D}_{k/2} \text{ of } Sp(1) \text{ on } \mathcal{P}_k \approx \mathbb{C}^{k+1}, \tag{10.3.7}$$

where $\mathcal{P}_k$ is the space of polynomials on $\mathbb{C}^2$, homogeneous of degree k, given, for $k \in \mathbb{Z}^+$, by $\mathcal{D}_{k/2}(\xi) = D_{k/2}(g)$, with $D_{k/2}$ given by Proposition 4.1.2 and $\sigma(g) = \xi$. That is,

$$\mathcal{D}_{1/2}(\xi)f(z) = f(g^{-1}z), \quad \xi \in Sp(1),\ \sigma(g) = \xi,\ f \in \mathcal{P}_k. \tag{10.3.8}$$

The group $\mathbf{Sp}(2)$ and its roots

We turn to $Sp(2)$, with Lie algebra

$$\mathfrak{sp}(2) = T_I Sp(2) = \left\{ \begin{pmatrix} u_1 & \xi \\ -\bar{\xi} & u_2 \end{pmatrix} : \xi \in \mathbb{H},\ u_\ell \in \operatorname{Im}\mathbb{H} = \mathbb{R}^3 \right\}. \tag{10.3.9}$$

$Sp(2)$ has a maximal torus $\mathbb{T}$ with Lie algebra

$$\mathfrak{h} = \left\{ \vartheta_{ab} = \begin{pmatrix} ai & \\ & bi \end{pmatrix} : a, b \in \mathbb{R} \right\}. \tag{10.3.10}$$

We see that $Sp(2)$ has dimension 10 and rank 2, hence 8 roots. We seek a root space decomposition

$$\mathfrak{sp}(2)_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus \bigoplus_{\alpha} \mathfrak{g}_{\alpha}, \tag{10.3.11}$$

summing over the eight roots α. A calculation yields

$$\left[\vartheta_{ab}, \begin{pmatrix} u_1 & \xi \\ -\overline{\xi} & u_2 \end{pmatrix}\right] = \begin{pmatrix} a[i,u_1] & ai\xi - b\xi i \\ a\overline{\xi}i - bi\overline{\xi} & b[i,u_2] \end{pmatrix}. \tag{10.3.12}$$

We hence seek root vectors of the form

$$\begin{pmatrix} u_1 & \\ & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & \\ & u_2 \end{pmatrix}, \quad \begin{pmatrix} & \xi \\ -\overline{\xi} & \end{pmatrix}, \quad u_\ell \in \mathbb{C}^3,\ \xi \in \mathbb{H}_{\mathbb{C}}. \tag{10.3.13}$$

Parallel to (10.3.6), we have

$$[i, u_\ell] = \mp 2\sqrt{-1}u_\ell, \quad u_\ell = j \pm \sqrt{-1}k, \tag{10.3.14}$$

yielding 4 root vectors, with roots

$$\mu_m^+(a,b) = 2(-1)^m a, \quad \mu_m^-(a,b) = 2(-1)^m b, \quad m \in \{1,2\}. \tag{10.3.15}$$

Root vectors of the third kind in (10.3.13) arise for $\xi \in \mathbb{H}_{\mathbb{C}}$ that are simultaneously eigenvectors for left multiplication and for right multiplication by i, i.e.,

$$L\xi = i\xi, \quad R\xi = \xi i. \tag{10.3.16}$$

We have

$$\begin{aligned} L1 &= i, \quad Li = -1, \quad Lj = k, \quad Lk = -j, \\ R1 &= i, \quad Ri = -1, \quad Rj = -k, \quad Rk = j, \end{aligned} \tag{10.3.17}$$

and hence

$$\begin{aligned} L(1+\sqrt{-1}i) &= -\sqrt{-1}(1+\sqrt{-1}i), \\ L(1-\sqrt{-1}i) &= \sqrt{-1}(1-\sqrt{-1}i), \\ L(j+\sqrt{-1}k) &= -\sqrt{-1}(j+\sqrt{-1}k), \\ L(j-\sqrt{-1}k) &= \sqrt{-1}(j-\sqrt{-1}k), \end{aligned} \tag{10.3.18}$$

and similarly for R. This yields values of ξ for which $\begin{pmatrix} & \xi \\ -\overline{\xi} & \end{pmatrix}$ is a root vector, with roots given, respectively, by

$$-a+b, \quad a-b, \quad -a-b, \quad a+b. \tag{10.3.19}$$

Here the $\mathbb{R}$-linear map $\xi \mapsto \overline{\xi}$ on $\mathbb{H}$ is extended to be $\mathbb{C}$-linear on $\mathbb{H}_{\mathbb{C}}$.

See Figure 10.3.1 for a picture of these roots, expanded with respect to the dual basis to the basis $\{\vartheta_{10}, \vartheta_{01}\}$ of $\mathfrak{h}$.

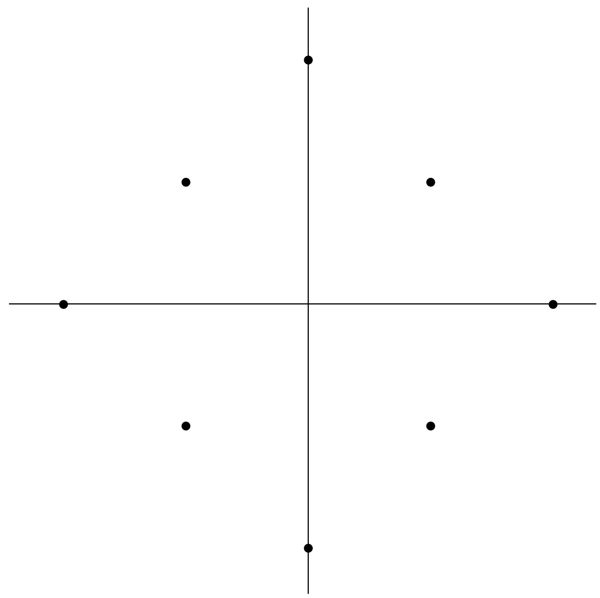

Figure 10.3.1. Roots of Sp(2)

This looks like the root diagram of $SO(5)$, depicted in Figure 7.1.1, except it is rotated by 90° and dilated by a factor of $\sqrt{2}$, i.e., obtained from Figure 7.1.1 by applying the matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{10.3.20}$$

Later in this section we show that $Sp(2)$ and $SO(5)$ are locally isomorphic.

Representation of $\mathbf{Sp}(2)$ on $\mathbb{H}^2$

We now look at the "standard" representation of Sp(2) on $\mathbb{H}^2$,

$$\pi(T)v = Tv, \quad T \in Sp(2),\ v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{H}^2. \tag{10.3.21}$$

The maps $\pi(T)$ are $\mathbb{H}$-linear, but here we consider π as a 4-dimensional complex representation, on $\mathbb{H}^2 \approx \mathbb{C}^4$. The derived representation of $\mathfrak{h}$ is

given by

$$d\pi(\vartheta_{ab})v = \begin{pmatrix} aiv_1 \\ biv_2 \end{pmatrix}. \tag{10.3.22}$$

Now $M_i : \mathbb{H} \to \mathbb{H}$, given by $M_i v_j = iv_j$, is $\mathbb{C}$-linear on $\mathbb{H}$, satisfying

$$M_i 1 = 1 \cdot i, \quad M_i i = i \cdot i; \quad M_i j = -ji, \quad M_i k = -ki. \tag{10.3.23}$$

Hence we have weights

$$a, \quad -a, \quad b, \quad -b, \tag{10.3.24}$$

with weight vectors given, respectively, by

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} j \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ j \end{pmatrix}. \tag{10.3.25}$$

Action of $\mathbf{Sp}(2)$ on $\mathbb{P}^1(\mathbb{H})$

The standard action of $Sp(2)$ on $\mathbb{H}^2$ leads to an action on the quaternionic projective space

$$\mathbb{P}^1(\mathbb{H}) = \mathbb{H}^2 \setminus 0/\sim, \tag{10.3.26}$$

with equivalence relation

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} \sim \begin{pmatrix} \xi\zeta \\ \eta\zeta \end{pmatrix}, \quad \forall\, \zeta \in \mathbb{H} \setminus 0. \tag{10.3.27}$$

Since $Sp(2)$ acts on the left, this action preserves the equivalence relation. We have

$$\tau(T) : \mathbb{P}^1(\mathbb{H}) \longrightarrow \mathbb{P}^1(\mathbb{H}), \quad T \in Sp(2), \tag{10.3.28}$$

each $\tau(T)$ being a diffeomorphism of $\mathbb{P}^1(\mathbb{H})$ onto itself, i.e.,

$$\tau : Sp(2) \longrightarrow \mathrm{Diff}(\mathbb{P}^1(\mathbb{H})). \tag{10.3.29}$$

Let us single out special points on $\mathbb{P}^1(\mathbb{H})$,

$$N = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad S = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{10.3.30}$$

Here, with slight abuse of notation, we let $\binom{1}{0}$ denote the equivalence class, etc. We have "stereographic projections"

$$\begin{aligned} \psi_N : \mathbb{P}^1(\mathbb{H}) \setminus N \to \mathbb{H}, \quad \psi_N \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \xi\eta^{-1}, \\ \psi_S : \mathbb{P}^1(\mathbb{H}) \setminus S \to \mathbb{H}, \quad \psi_S \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \eta\xi^{-1}. \end{aligned} \tag{10.3.31}$$

Then we have

$$\psi_N \circ \psi_S^{-1} : \mathbb{H} \setminus 0 \longrightarrow \mathbb{H} \setminus 0, \tag{10.3.32}$$

given by

$$\psi_N \circ \psi_S^{-1}(\xi) = \psi_N \begin{pmatrix} \xi^{-1} \\ 1 \end{pmatrix} = \xi^{-1} = |\xi|^{-2}\overline{\xi}. \tag{10.3.33}$$

Comparison with standard stereographic projections for the unit sphere $S^4 \subset \mathbb{R}^5$ to $\mathbb{R}^4$ yields a diffeomorhphism

$$\varphi : \mathbb{P}^1(\mathbb{H}) \longrightarrow S^4, \tag{10.3.34}$$

in turn yielding a group of diffeomorphisms of S^4 to itself,

$$\gamma : Sp(2) \longrightarrow \mathrm{Diff}(S^4), \tag{10.3.35}$$

given by

$$\gamma(T) : S^4 \to S^4, \quad \gamma(T) = \varphi \circ \tau(T) \circ \varphi^{-1}, \quad T \in Sp(2). \tag{10.3.36}$$

In order to obtain more information on τ and γ, it is useful to consider

$$K_N = \{A \in Sp(2) : A \cdot N = N\}. \tag{10.3.37}$$

Now

$$A = \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix}, \; N = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Longrightarrow A \cdot N = \begin{pmatrix} \xi_{11} \\ \xi_{21} \end{pmatrix}, \tag{10.3.38}$$

so

$$A \cdot N = N \Longrightarrow \xi_{21} = 0. \tag{10.3.39}$$

Hence

$$\begin{aligned} A = \begin{pmatrix} \xi_{11} & \xi_{12} \\ 0 & \xi_{22} \end{pmatrix} &\Rightarrow A^* = \begin{pmatrix} \overline{\xi}_{11} & 0 \\ \overline{\xi}_{12} & \overline{\xi}_{22} \end{pmatrix} \\ &\Rightarrow A^*A = \begin{pmatrix} |\xi_{11}|^2 & \overline{\xi}_{11}\xi_{12} \\ \overline{\xi}_{12}\xi_{11} & |\xi_{12}|^2 + |\xi_{22}|^2 \end{pmatrix}, \end{aligned} \tag{10.3.40}$$

so

$$A^*A = I \Longrightarrow |\xi_{11}| = 1, \; \xi_{12} = 0, \; |\xi_{22}| = 1. \tag{10.3.41}$$

We have

$$K_N = \left\{ \begin{pmatrix} \xi_1 & \\ & \xi_2 \end{pmatrix} : \xi_j \in Sp(1) \right\}. \tag{10.3.42}$$

For further insight into how K_N acts on $\mathbb{P}^1(\mathbb{H})$, we have, for $\eta \in \mathbb{H}, \; \xi_j \in Sp(1)$,

$$\psi_S \begin{pmatrix} \xi_1 & \\ & \xi_2 \end{pmatrix} \psi_S^{-1}(\eta) = \psi_S \begin{pmatrix} \xi_1 \\ \xi_2\eta \end{pmatrix} = \xi_2\eta\overline{\xi}_1, \tag{10.3.43}$$

the right side giving the action of $Sp(1) \times Sp(1)$ on $\mathbb{H}$ covering the SO(4) action, as indicated in Exercise 7 of §1.2. It follows that $\gamma : K_N \to \mathrm{Diff}(S^4)$

acts as a covering of the standard group of rotations of S^4 about the axis through the origin and N. Hence

$$\gamma : K_N \longrightarrow SO(4) \subset SO(5). \tag{10.3.44}$$

We pause to record a corollary to the computation (10.3.43).

Proposition 10.3.1. *In* (10.3.25), $\operatorname{Ker}\tau = \{\pm I\}$. *Hence* $\operatorname{Ker}\gamma = \{\pm I\}$.

Proof. Clearly $\tau(I) = \tau(-I) = I$. For the converse, note that

$$\begin{aligned} A \in \operatorname{Ker}\tau &\Rightarrow A \in K_N \text{ and } \psi_S A \psi_S^{-1}(\eta) = \eta,\ \forall\, \eta \in \mathbb{H} \setminus 0 \\ &\Rightarrow A \in \operatorname{diag}(\xi_1, \xi_2) \text{ and } \xi_2 \eta \overline{\xi}_1 = \eta,\ \forall\, \eta \in \mathbb{H} \\ &\Rightarrow \xi_2 = \xi_1 \text{ and } \xi_2 \eta = \eta \xi_2,\ \forall\, \eta \in \mathbb{H} \\ &\Rightarrow \xi_1 = \xi_2 = \pm 1. \end{aligned} \tag{10.3.45}$$

□

To proceed, we examine e^{tX}, for $X \in \mathfrak{sp}(2)$ in a space complementary to the Lie algebra of K_N, namely

$$X = \begin{pmatrix} & -\overline{\xi} \\ \xi & \end{pmatrix}, \quad \xi \in \mathbb{H}. \tag{10.3.46}$$

Note that

$$\begin{aligned} \xi \in \mathbb{H},\ |\xi| = 1 \Rightarrow X^2 = -I \Rightarrow e^{tX} &= (\cos t) I + (\sin t) X \\ &= \begin{pmatrix} \cos t & -(\sin t)\overline{\xi} \\ (\sin t)\xi & \cos t \end{pmatrix}. \end{aligned} \tag{10.3.47}$$

Complementing (10.3.39), we have

$$e^{tX} N = \begin{pmatrix} \cos t \\ (\sin t)\xi \end{pmatrix}, \quad \text{hence} \quad \psi_S(e^{tN} N) = (\tan t)\xi. \tag{10.3.48}$$

More generally, complementing (10.3.43), we have, for $\eta \in \mathbb{H}$,

$$\begin{aligned} \psi_S e^{tX} \psi_S^{-1}(\eta) &= \psi_S \begin{pmatrix} c(t) & -s(t)\overline{\xi} \\ s(t)\xi & c(t) \end{pmatrix} \begin{pmatrix} 1 \\ \eta \end{pmatrix} \\ &= \psi_S \begin{pmatrix} c(t) - s(t)\overline{\xi}\eta \\ s(t)\xi + c(t)\eta \end{pmatrix} \\ &= (s(t)\xi + c(t)\eta)(c(t) - s(t)\overline{\xi}\eta)^{-1}, \end{aligned} \tag{10.3.49}$$

where we have set

$$c(t) = \cos t, \quad s(t) = \sin t. \tag{10.3.50}$$

We can factor out ξ on the left and set $\eta = \xi\zeta$, to write

$$\psi_S e^{tX} \psi_S^{-1}(\xi\zeta) = \xi(s(t) + c(t)\zeta)(c(t) - s(t)\zeta)^{-1}, \quad \zeta \in \mathbb{H}, \tag{10.3.51}$$

for X of the form (10.3.46)–(10.3.47).

Representation of $\mathbf{Sp}(2)$ on $\mathbf{M}(2, \mathbb{H})$

We define

$$\kappa : Sp(2) \longrightarrow \mathcal{L}_{\mathbb{R}}(M(2, \mathbb{H})) \tag{10.3.52}$$

by

$$\kappa(A)X = AXA^*, \quad A \in Sp(2),\ X \in M(2, \mathbb{H}). \tag{10.3.53}$$

We give the $\mathbb{R}$-linear space $M(2, \mathbb{H})$ the Hilbert-Schmidt norm

$$\|X\|_{\mathrm{HS}}^2 = \sum_{j,k} |\xi_{jk}|^2, \quad X = \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix},\ \xi_{jk} \in \mathbb{H}. \tag{10.3.54}$$

We also give $M(2, \mathbb{H})$ the associated real inner product. It is clear that

$$X \mapsto AX, \quad A \in Sp(2) \tag{10.3.55}$$

preserves the HS-norm. Also, $X \mapsto X^*$ preserves this norm, so

$$X \mapsto XB, \quad B \in Sp(2) \tag{10.3.56}$$

also does. Hence, for each $A \in Sp(2)$, $\kappa(A)$ preserves the HS-norm on $M(2, \mathbb{H})$. The space $M(2, \mathbb{H})$ has invariant subspaces,

$$M(2, \mathbb{H}) = \mathfrak{sp}(2) \oplus S^2(\mathbb{H}), \tag{10.3.57}$$

where

$$\mathfrak{sp}(2) = \{X^* = -X\}, \quad S^2(\mathbb{H}) = \{X^* = X\}. \tag{10.3.58}$$

The two spaces on the right side of (10.3.57) are mutually orthogonal, and $\kappa(A)$ preserves each factor:

$$\kappa(A) : \mathfrak{sp}(2) \to \mathfrak{sp}(2), \quad \kappa(A) : S^2(\mathbb{H}) \to S^2(\mathbb{H}). \tag{10.3.59}$$

The first action is the adjoint representation, already considered in this section. There is a further decomposition of $S^2(\mathbb{H})$ into mutually orthogonal pieces:

$$\begin{aligned} S^2(\mathbb{H}) &= \{cI : c \in \mathbb{R}\} \oplus S_0^2(\mathbb{H}), \\ S_0^2(\mathbb{H}) &= \left\{ \begin{pmatrix} a & \overline{\xi} \\ \xi & -a \end{pmatrix} : a \in \mathbb{R},\ \xi \in \mathbb{H} \right\}. \end{aligned} \tag{10.3.60}$$

Restricting to $S_0^2(\mathbb{H})$ gives the representation

$$\kappa_0 : Sp(2) \longrightarrow \mathcal{L}_{\mathbb{R}}(S_0^2(\mathbb{H})). \tag{10.3.61}$$

The isomorphism $\mathbf{Sp}(2) \approx \mathbf{Spin}(5)$

At this point, a dimension count is called for:

$$\dim_{\mathbb{R}} M(2,\mathbb{H}) = 16, \quad \dim_{\mathbb{R}} \mathfrak{sp}(2) = 10, \quad \dim_{\mathbb{R}} S_0^2(\mathbb{H}) = 5. \tag{10.3.62}$$

We have that $\kappa(A)$ preserves the HS-norm, so (10.3.61) basically says

$$\kappa_0 : Sp(2) \longrightarrow SO(5). \tag{10.3.63}$$

It is instructive to compute $\kappa_0(A)$ for

$$A = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}, \quad \alpha, \beta \in Sp(1). \tag{10.3.64}$$

Indeed, a calculation gives

$$\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix} \begin{pmatrix} a & \overline{\xi} \\ \xi & -a \end{pmatrix} \begin{pmatrix} \overline{\alpha} & \\ & \overline{\beta} \end{pmatrix} = \begin{pmatrix} a & \alpha\overline{\xi}\,\overline{\beta} \\ \beta\xi\overline{\alpha} & -a \end{pmatrix}. \tag{10.3.65}$$

We now have the following key result.

Proposition 10.3.2. *In* (10.3.63),

$$\operatorname{Ker} \kappa_0 = \{\pm I\}. \tag{10.3.66}$$

Proof. Since κ_0 is a group homomorphism, $\operatorname{Ker} \kappa_0$ is a normal subgroup of $Sp(2)$. Take $x \in \operatorname{Ker} \kappa_0$. We know that

$$\mathbb{T} = \left\{ \begin{pmatrix} e^{i\theta_1} & \\ & e^{i\theta_2} \end{pmatrix} : \theta_j \in \mathbb{R} \right\} \tag{10.3.67}$$

is a conjugating torus for $Sp(2)$. Hence there exists $g \in Sp(2)$ such that

$$g^{-1} x g = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix} \in \mathbb{T}, \tag{10.3.68}$$

and $g^{-1}xg \in \operatorname{Ker} \kappa_0$. The calculation (10.3.65) implies

$$\begin{aligned} &\beta\xi\overline{\alpha} = \xi, \quad \forall\, \xi \in \mathbb{H}, \text{ hence (taking } \xi = 1) \\ &\alpha = \beta, \quad \text{and } \alpha\xi = \xi\alpha, \quad \forall\, \xi \in \mathbb{H}, \text{ hence} \\ &\alpha \in \mathbb{R}, \quad \text{so } \alpha = \beta = \pm 1. \end{aligned} \tag{10.3.69}$$

This implies

$$g^{-1} x g = \pm I, \quad \text{hence } x = \pm I. \tag{10.3.70}$$

□

It follows from (10.3.66) that $d\kappa_0 : \mathfrak{sp}(2) \to \mathfrak{so}(5)$ is injective, hence an isomorphism, since both Lie algebas have dimension 10. Hence $\kappa_0(Sp(2))$ is a 10-dimensional subgroup of $SO(5)$, and since $SO(5)$ is connected, this implies

$$\kappa_0 \text{ in (10.3.63) is onto.} \tag{10.3.71}$$

Thus

$$SO(5) \approx Sp(2)/\{\pm I\}. \tag{10.3.72}$$

We know $Sp(2)$ is connected, so κ_0 in (10.3.63) presents $Sp(2)$ as a two-fold connected covering group of $SO(5)$. Recall from Chapter 7 that Spin(5) has this property. We have seen in Chapter 7 that Spin(n) is simply connected. Hence the Lie group homomorphism

$$\mathrm{Spin}(5) \longrightarrow SO(5) \tag{10.3.73}$$

lifts to a Lie group homomorphism

$$\sigma : \mathrm{Spin}(5) \longrightarrow \mathrm{Sp}(2). \tag{10.3.74}$$

Proposition 10.3.3. *The homomorphism* (10.3.74) *is an isomorphism of Lie groups.*

Proof. The map $\kappa_0 \circ \sigma : \mathrm{Spin}(5) \to SO(5)$ is the standard covering homomorphism, so $\mathrm{Ker}\, \kappa_0 \circ \sigma$ has two elements. Thus $\ker \sigma$ has one element, so σ in (10.3.74) is injective. Surjectivity is established by an argument similar to that used in (10.3.71). □

Back to $\mathbb{P}^1(\mathbb{H})$

We give another perspectice on quaternionic projective space, using a natural one-to-one correspondence between elements of $\mathbb{P}^1(\mathbb{H})$ and orthogonal projections in $\mathcal{L}_{\mathbb{H}}(\mathbb{H}^2)$ of the form

$$P_v u = v\langle u, v\rangle, \quad v \in \mathbb{H}^2,\ |v| = 1. \tag{10.3.75}$$

Compare (10.2.16). Note that if $\zeta \in \mathbb{H},\ |\zeta| = 1$,

$$P_{v\zeta} u = v\zeta\langle u, v\zeta\rangle = v\zeta\overline{\zeta}\langle u, v\rangle = P_v u, \tag{10.3.76}$$

so $P_{v\zeta} = P_v$, i.e., P_v depends only on $\mathrm{Span}_{\mathbb{H}}(v)$. This yields the bijection

$$p : \mathbb{P}^1(\mathbb{H}) \longrightarrow \{P_v : v \in \mathbb{H}^2, |v| = 1\}. \tag{10.3.77}$$

Note that

$$\begin{aligned} A \in Sp(2) \Rightarrow P_{Av} u &= Av\langle u, Av\rangle \\ &= Av\langle A^* u, v\rangle \\ &= AP_v A^* u. \end{aligned} \tag{10.3.78}$$

Next, we set

$$H_v = P_v - P_v^{\perp} = P_v - (I - P_v) = 2P_v - I, \tag{10.3.79}$$

for unit $v \in \mathbb{H}^2$. From (10.3.76) and (10.3.78) we have

$$H_{v\zeta} = H_v, \quad H_{Av} = AH_v A^*, \quad \text{for} \ \ \zeta \in Sp(1),\ A \in Sp(2). \tag{10.3.80}$$

Also (10.3.77) yields a bijection

$$q : \mathbb{P}^1(\mathbb{H}) \longrightarrow \{H_v : v \in \mathbb{H}^2,\ |v| = 1\}. \tag{10.3.81}$$

Note that

$$H_v \in S_0^2(\mathbb{H}), \tag{10.3.82}$$

and we can rewrite the second half of (10.3.80) as

$$A \in Sp(2) \Longrightarrow H_{Av} = \kappa_0(A) H_v. \tag{10.3.83}$$

We compute the matrix representations of P_v and H_v, for unit $v \in \mathbb{H}^2$, starting with

$$\begin{aligned} v = \begin{pmatrix} \zeta \\ \eta \end{pmatrix} \Rightarrow P_v \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} \zeta \\ \eta \end{pmatrix} \overline{\zeta} = \begin{pmatrix} |\zeta|^2 \\ \eta\overline{\zeta} \end{pmatrix} \\ P_v \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= \begin{pmatrix} \zeta \\ \eta \end{pmatrix} \overline{\eta} = \begin{pmatrix} \zeta\overline{\eta} \\ |\eta|^2 \end{pmatrix} \\ \Rightarrow P_v &= \begin{pmatrix} |\zeta|^2 & \zeta\overline{\eta} \\ \eta\overline{\zeta} & |\eta|^2 \end{pmatrix}, \quad |\zeta|^2 + |\eta|^2 = |v|^2 = 1. \end{aligned} \tag{10.3.84}$$

Hence

$$H_v = \begin{pmatrix} 2|\zeta|^2 - 1 & 2\zeta\overline{\eta} \\ 2\eta\overline{\zeta} & 2|\eta|^2 - 1 \end{pmatrix} \in S_0^2(\mathbb{H}). \tag{10.3.85}$$

Note that

$$\begin{aligned} \|P_v\|_{\mathrm{HS}}^2 &= |\zeta|^4 + |\eta|^4 + 2|\eta|^2|\zeta|^2 \\ &= (|\zeta|^2 + |\eta|^2)^2 = 1. \end{aligned} \tag{10.3.86}$$

Also

$$\begin{aligned} &2P_v = H_v + I, \quad H_v \perp I \text{ in } S^2(\mathbb{H}) \\ &\Rightarrow 4\|P_v\|_{\mathrm{HS}}^2 = \|H_v\|_{\mathrm{HS}}^2 + \|I\|_{\mathrm{HS}}^2 \\ &\Rightarrow \|H_v\|_{\mathrm{HS}}^2 = 2. \end{aligned} \tag{10.3.87}$$

This leads to the following

Proposition 10.3.4. *We have*

$$\{H_v : v \in \mathbb{H}^2,\ |v| = 1\} = \{X \in S_0^2(\mathbb{H}) : \|X\|_{\mathrm{HS}}^2 = 2\}. \tag{10.3.88}$$

Proof. By (10.3.87), the left side of (10.3.88) is contained in the right side. By (10.3.63) and (10.3.71), $Sp(2)$ acts transitively on the right side of (10.3.88), and by (10.3.83) $Sp(2)$ leaves the left side invariant. This implies identity in (10.3.88). $\square$

Corollary 10.3.5. *The results* (10.3.81), (10.3.83), *and* (10.3.88) *yield a bijection*

$$q : \mathbb{P}^1(\mathbb{H}) \longrightarrow \{X \in S_0^2(\mathbb{H}) : \|X\|_{\mathrm{HS}}^2 = 2\}, \tag{10.3.89}$$

intertwining the $Sp(2)$ actions.

The representations of $\mathbf{Sp}(2)$

The isomorphism $Sp(2) \approx \mathrm{Spin}(5)$ established in Proposition 10.3.3 allows us to read off the irreducible unitary representations of $Sp(2)$ from the results on $\mathrm{Spin}(n)$ (with $n = 5$) established in Chapter 7. It is worth noting that the spinor representation, filling in the gap mentioned at the end of §7.1, is given by the representation π of $Sp(2)$ on $\mathbb{H}^2$ described in (10.3.21). That is, in the notation of §7.5,

$$\pi \approx D^+_{1/2}. \tag{10.3.90}$$

We note that $\mathbb{H}^2$ and $S_+(6)$ are both complex vector spaces of $\mathbb{C}$-dimension 4. In particular,

$$S_+(6) = \bigoplus_{j \text{ even}} \Lambda^j_{\mathbb{C}}\mathbb{C}^3 = \Lambda^0_{\mathbb{C}}\mathbb{C}^3 \oplus \Lambda^2_{\mathbb{C}}\mathbb{C}^3 \approx \mathbb{C}^4. \tag{10.3.91}$$

Exercises

1. Show that (10.3.88) implies

$$\{H_v : v \in \mathbb{H}^2, |v| = 1\} = \left\{ \begin{pmatrix} a & \overline{\xi} \\ \xi & -a \end{pmatrix} : a \in \mathbb{R},\ \xi \in \mathbb{H},\ a^2 + |\xi|^2 = 1 \right\}.$$

Given H_v of this form, show that

$$P_v = \frac{1}{2} \begin{pmatrix} 1+a & \overline{\xi} \\ \xi & 1-a \end{pmatrix},$$

and that

$$\begin{aligned} \operatorname{Span}_{\mathbb{H}} v = \ &\operatorname{Span}_{\mathbb{H}} \begin{pmatrix} 1+a \\ \xi \end{pmatrix}, \quad \text{if } a \neq -1, \\ &\operatorname{Span}_{\mathbb{H}} \begin{pmatrix} \overline{\xi} \\ 1-a \end{pmatrix}, \quad \text{if } a \neq 1. \end{aligned}$$

2. Set

$$\begin{aligned} \mathfrak{g} &= \mathfrak{sp}(2) = \{X \in M(2, \mathbb{H}) : X^* = -X\}, \\ \mathfrak{k} &= \mathfrak{sp}(1) \oplus \mathfrak{sp}(1) = \left\{ \begin{pmatrix} u_1 & \\ & u_2 \end{pmatrix} : u_j \in \operatorname{Im} \mathbb{H} \right\}, \\ \mathfrak{p} &= \left\{ \begin{pmatrix} & -\overline{\xi} \\ \xi & \end{pmatrix} : \xi \in \mathbb{H} \right\}. \end{aligned}$$

Show that

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}, \quad [\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}.$$

3. Given

$$Y = \begin{pmatrix} u_1 & \\ & u_2 \end{pmatrix} \in \mathfrak{k}, \quad Z = \begin{pmatrix} & -\overline{\xi} \\ \xi & \end{pmatrix} \in \mathfrak{p}, \quad |\xi| = 1,$$

compute

$$e^{tY} = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix},\ \alpha, \beta \in Sp(1), \quad e^{tZ} = (\cos t)I + (\sin t)Z.$$

Then compute

$$\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix} \begin{pmatrix} c(t) & -s(t)\overline{\xi} \\ s(t)\xi & c(t) \end{pmatrix} \in Sp(2).$$

4. Denote the groups generated by $\mathfrak{g}$ and $\mathfrak{k}$ by $G = Sp(2)$ and $K \approx Sp(1) \times Sp(1)$. Given $\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix} = A \in K$ and $Z \in \mathfrak{p}$ (as in Exercise 3), compute

$$AZA^{-1}, \quad \text{representing } K \text{ on } \mathfrak{p}.$$

10.4. Second introduction to Sp(n)

For general $n \in \mathbb{N}$, the group $Sp(n)$, introduced in Chapter 1, §1.2, is defined as

$$Sp(n) = \{A \in M(n, \mathbb{H}) : A^* A = I\}, \tag{10.4.1}$$

where

$$A = (a_{\ell m}),\ a_{\ell m} \in \mathbb{H} \Longrightarrow A^* = (\overline{a}_{m\ell}). \tag{10.4.2}$$

As seen in Chapter 1, $Sp(n) \hookrightarrow O(4n)$, as a consequence of the fact that, given $A \in M(n, \mathbb{H})$, A belongs to $Sp(n)$ if and only if applying A preserves the quaternionic inner product

$$\langle \xi, \eta \rangle = \sum_{\ell} \overline{\eta}_\ell \xi_\ell. \tag{10.4.3}$$

An exercise there gives the more precise result that $Sp(n) \hookrightarrow U(2n)$. Here we use (10.1.41)–(10.1.47) to prove a more precise result. In fact, parallel to (10.1.42), we have the $\mathbb{R}$-linear isomorphism

$$\gamma : \mathbb{H}^n \longrightarrow \mathbb{C}^{2n},$$

and then from (10.1.45)–(10.1.46) we have that, for $\xi, \eta \in \mathbb{H}^n$,

$$\begin{aligned} \mathrm{Co}\langle \xi, \eta \rangle &= ((\gamma(\xi), \gamma(\eta))), \\ \mathrm{Sp}\langle \xi, \eta \rangle &= -\sigma(\gamma(\xi), \gamma(\eta)), \end{aligned} \tag{10.4.4}$$

where $((\ ,\ ))$ denotes the standard Hermitian inner product on $\mathbb{C}^{2n}$, and

$$\sigma((\alpha_1, \alpha_2), (\beta_1, \beta_2)) = \alpha_1 \cdot \beta_2 - \alpha_2 \cdot \beta_1, \tag{10.4.5}$$

is an antisymmetric, $\mathbb{C}$-bilinear form on $\mathbb{C}^{2n}$, called the symplectic form. (Here, $\alpha_\nu, \beta_\nu \in \mathbb{C}^n$, and $\alpha_\mu \cdot \beta_\nu$ is the usual $\mathbb{C}$-bilinear dot product on $\mathbb{C}^n$.) Thus we have the natural injection

$$Sp(n) \hookrightarrow U(2n) \cap Sp(2n, \mathbb{C}), \tag{10.4.6}$$

where $Sp(2n, \mathbb{C})$ denotes the group of elements of $Gl(2n, \mathbb{C})$ that preserve the form σ.

Having reintroduced $Sp(n)$, we recall that its Lie algebra is given by

$$\mathfrak{sp}(n) = \{X \in M(n, \mathbb{H}) : X^* = -X\}. \tag{10.4.7}$$

10.5. Roots of Sp(n), $n \ge 3$

We now look at the roots of $Sp(n)$ for $n \ge 3$. Recall that

$$\mathfrak{sp}(n) = \{X \in M(n,\mathbb{H}) : X^* = -X\}. \tag{10.5.1}$$

Parallel to (10.3.10), we write

$$\mathfrak{h} = \{\vartheta_a : a \in \mathbb{R}^n\}, \quad \vartheta_a = \begin{pmatrix} a_1 i & & \\ & \ddots & \\ & & a_n i \end{pmatrix}. \tag{10.5.2}$$

In case $n = 3$, the parallel to (10.3.9) is

$$X = \begin{pmatrix} u_1 & \xi_{12} & \xi_{13} \\ -\overline{\xi}_{21} & u_2 & \xi_{23} \\ -\overline{\xi}_{13} & -\overline{\xi}_{23} & u_3 \end{pmatrix}, \quad u_\ell \in \operatorname{Im}\mathbb{H},\ \xi_{\ell m} \in \mathbb{H}. \tag{10.5.3}$$

We use similar notation for $X \in \mathfrak{sp}(n)_{\mathbb{C}}$. As in §10.3, we extend the $\mathbb{R}$-linear involution $\xi \mapsto \overline{\xi}$ on $\mathbb{H}$ by $\mathbb{C}$-linearity to an involution on $\mathbb{H}_{\mathbb{C}}$. More generally, $X \in \mathfrak{sp}(n)_{\mathbb{C}}$ has the form $X = (x_{\ell m})$, with

$$\begin{aligned} &x_{\ell\ell} = u_\ell \in (\operatorname{Im}\mathbb{H})_{\mathbb{C}} \approx \mathbb{C}^3, \\ &\ell < m \Rightarrow x_{\ell m} = \xi_{\ell m} \in \mathbb{H}_{\mathbb{C}},\ x_{m\ell} = -\overline{\xi}_{\ell m}. \end{aligned} \tag{10.5.4}$$

We extend the commutator calculation (10.3.12). In case $n = 3$, this becomes

$$[\vartheta_a, X] = \begin{pmatrix} a_1[i,u_1] & a_1 i\xi_{12} - a_2\xi_{12} i & a_1 i\xi_{13} - a_3\xi_{13} i \\ * & a_2[i,u_2] & a_2 i\xi_{23} - a_3\xi_{23} i \\ * & * & a_3[i,u_3] \end{pmatrix}, \tag{10.5.5}$$

where the "$*$"s indicate that $Y = [\vartheta_a, X]$ satisfies $Y^* = -Y$. More generally, we have

$$\begin{aligned} &[\vartheta_a, X] = (y_{\ell m}),\ y_{\ell\ell} = a_\ell[i,u_\ell], \\ &\ell < m \Rightarrow y_{\ell m} = a_\ell i\xi_{\ell m} - a_m\xi_{\ell m} i,\ y_{m\ell} = -\overline{y}_{\ell m}. \end{aligned} \tag{10.5.6}$$

Parallel to (10.3.13), we have two types of candidates for root vectors, either E_ℓ, a diagonal matrix with ℓth diagonal entry $u_\ell \in (\operatorname{Im}\mathbb{H})_{\mathbb{C}} \approx \mathbb{C}^3$ and all other entries zero, or $E_{\ell m}$, $\ell < m$, with (ℓ, m)-entry $\xi_{\ell m} \in \mathbb{H}_{\mathbb{C}}$, (m,ℓ)-entry $-\overline{\xi}_{\ell m}$, all other entries zero. Calculations parallel to (10.3.14)–(10.3.19) yield the following.

Proposition 10.5.1. *With respect to the basis of $\mathfrak{h}$ defined by $\mathfrak{h} \approx \mathbb{R}^n$, $\vartheta_a \leftrightarrow a$, roots with root vectors of the form E_ℓ are given by*

$$\rho_\ell^\pm(a) = \pm 2a_\ell, \tag{10.5.7}$$

and roots with root vectors of the form $E_{\ell m}$ are given by

$$\rho_{\ell m}^{\alpha\beta}(a) = (-1)^\alpha a_\ell + (-1)^\beta a_m, \quad \ell < m,\ \alpha, \beta \in \{0, 1\}. \tag{10.5.8}$$

We see that $\mathfrak{sp}(n)_{\mathbb{C}}$ has $2n$ roots of the first sort and $4(n^2 - n)/2 = 2(n^2 - n)$ roots of the second sort, hence a total of $2n^2$ roots. As a check, note that

$$\dim \mathfrak{sp}(n) = 2n^2 + n, \quad \dim \mathfrak{h} = n. \tag{10.5.9}$$

If we use the standard lexicographical ordering, we see that the n^2 positive roots of $Sp(n)$ are

$$\rho_\ell^+(a) = 2a_\ell, \quad \rho_{\ell m}^{0\beta}(a) = a_\ell + (-1)^\beta a_m, \tag{10.5.10}$$

with, respectively, $1 \le \ell \le n$ and $1 \le \ell < m \le n$. Recall from §6.6 that the set Σ of simple roots consists of those positive roots that cannot be written as a sum of two positive roots. Now

$$\rho_\ell^+ = \rho_{\ell m}^{00} + \rho_{\ell m}^{01} \ \text{ if } \ \ell < m, \tag{10.5.11}$$

so among $\{\rho_\ell^+\}$, only ρ_n^+ is simple. Also

$$\rho_{\ell m}^{00} = \rho_{\ell m}^{01} + \rho_m^+ \ \text{ if } \ \ell < m, \tag{10.5.12}$$

so none of the roots $\rho_{\ell m}^{00}$ are simple. Furthermore,

$$a_\ell - a_{\ell+j+1} = a_\ell - a_{\ell+1} + a_{\ell+1} - a_{\ell+j+1}, \tag{10.5.13}$$

so $\rho_{\ell m}^{01}$ is simple only for $m = \ell + 1,\ 1 \le \ell \le n - 1$. We summarize:

Proposition 10.5.2. *The set Σ of simple roots of $Sp(n)$ consists of*

$$\rho_{\ell,\ell+1}^{01},\ 1 \le \ell \le n-1, \quad \textit{and } \rho_n^+. \tag{10.5.14}$$

Let us denote these simple roots by

$$\begin{aligned} &\hat\rho_\ell \ (1 \le \ell \le n-1), \text{ and } \hat\rho_n, \text{ so} \\ &\hat\rho_\ell(a) = a_\ell - a_{\ell+1} \ (1 \le \ell \le n-1) \ \ \hat\rho_n(a) = 2a_n. \end{aligned} \tag{10.5.15}$$

We have

$$\begin{aligned} \langle \hat\rho_\ell, \hat\rho_\ell \rangle = 2 \quad &\text{if } \ 1 \le \ell \le n-1, \\ 4 \quad &\text{if } \ \ell = n. \end{aligned} \tag{10.5.16}$$

and

$$\begin{aligned} \langle \hat\rho_\ell, \hat\rho_{\ell+1} \rangle = -1 \quad &\text{if } \ 1 \le \ell \le n-2, \\ -2 \quad &\text{if } \ \ell = n-1, \end{aligned} \tag{10.5.17}$$

other inner products being 0 (subject to symmetry). Thus, following the general construction described in §6.6, we have the following Dynkin diagram.

Figure 10.5.1. Dynkin diagram of $\mathrm{Sp}(n)$

Note the similarity with the Dynkin diagram of $SO(2n+1)$ presented in §6.6, except the vertices decorated by dots are reversed. Only for $n = 2$ $(2n + 1 = 5)$ are the two diagrams equivalent.

10.6. Weights and representations of $\mathbf{Sp}(n)$, $n \geq 3$

It is a special case of Proposition 6.2.1 that, if $\lambda \in \mathfrak{h}'$, then for λ to be the weight of a unitary representation of $Sp(n)$, it is necessary that

$$2\,\frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}, \tag{10.6.1}$$

for each root α. It suffices to check the positive roots. If we identify $\mathfrak{h}$ with $\mathbb{R}^n$ via $\vartheta_a \leftrightarrow a$, and write

$$\lambda(a) = \lambda_1 a_1 + \cdots + \lambda_n a_n, \tag{10.6.2}$$

we see that

$$2\,\frac{\langle \lambda, \rho_\ell^+ \rangle}{\langle \rho_\ell^+, \rho_\ell^+ \rangle} = \lambda_\ell, \quad 2\,\frac{\langle \lambda, \rho_{\ell m}^{0\beta} \rangle}{\langle \rho_{\ell m}^{0\beta}, \rho_{\ell m}^{0\beta} \rangle} = \lambda_\ell + (-1)^\beta \lambda_m, \tag{10.6.3}$$

so (10.6.1) holds for each root α if and only if

$$\lambda_\ell \in \mathbb{Z}, \quad \forall\, \ell \in \{1, \dots, n\}. \tag{10.6.4}$$

In other words, the integer lattice $\mathbb{Z}^n \subset \mathbb{R}^n \approx \mathfrak{h}'$ forms the lattice of potential weights of representations of $Sp(n)$.

Furthermore, as defined in §6.2, an element $\lambda \in \mathfrak{h}'$ is a dominant integral weight if and only if the quantity in (10.6.1) is a non-negative integer for each positive root α. The calculations (10.6.3) then give the following.

Proposition 10.6.1. *An element $\lambda \in \mathfrak{h}'$ is a dominant integral weight if and only if* (10.6.2) *holds with*

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0, \quad \text{all} \;\; \lambda_\ell \in \mathbb{Z}^+. \tag{10.6.5}$$

Proof. The first part of (10.6.3) gives $\lambda_\ell \in \mathbb{Z}^+$ for each ℓ, and the second part gives $\lambda_\ell - \lambda_m \in \mathbb{Z}^+$ for $\ell < m$. □

The relevance to the issue at hand is stated as follows.

Theorem 10.6.2. *An element $\lambda \in \mathfrak{h}'$ is the highest weight of an irreducible unitary representation of $Sp(n) \Leftrightarrow \lambda$ is a dominant integral weight.*

We note that the implication "⇒" is established in great generality in Proposition 6.2.1. Given that $Sp(n)$ is simply connected (which we have not yet established) Theorem 10.6.2 is a special case of the Theorem of the Highest Weight, stated in §6.2. In light of general results of Chapter 6, this result has the following corollary.

Corollary 10.6.3. *There is a natural one-to-one correspondence between the set of elements $\lambda \in \mathfrak{h}'$ satisfying* (10.6.5) *and the set of equivalence classes of irreducible unitary representations of $Sp(n)$.*

Accordingly, if we have an irreducible unitary representation of $Sp(n)$ with highest weight λ, satisfying (10.6.2), we denote the representation by

$$D_{(\lambda_1,\lambda_2,\dots,\lambda_n)}. \tag{10.6.6}$$

Of course, $D_{(0,0,\dots,0)}$ is the trivial representation of $Sp(n)$ on $\mathbb{C}$. Note that this notation differs from that used (for $n = 1$) in (10.3.7).

Representation of $\mathbf{Sp}(n)$ on $\mathbb{H}^n$

We look at the standard representation of $Sp(n)$ on $\mathbb{H}^n$,

$$\pi(T)v = Tv, \quad T \in Sp(n),\ v = (v_1, \dots, v_n)^t \in \mathbb{H}^n. \tag{10.6.7}$$

The maps $\pi(T)$ are $\mathbb{H}$-linear, but we consider π as a $2n$-dimensional complex representation on $\mathbb{H}^n \approx \mathbb{C}^{2n}$. The derived representation of $\mathfrak{h}$ is given by

$$d\pi(\vartheta_a)v = \begin{pmatrix} a_1 i v_1 \\ \vdots \\ a_n i v_n \end{pmatrix}. \tag{10.6.8}$$

Calculations similar to (10.3.23)–(10.3.25) show that the weights of this representation are

$$\lambda_\ell^\pm(a) = \pm a_\ell, \quad 1 \leq \ell \leq n. \tag{10.6.9}$$

In particular, the highest weight is $\lambda_1^+(a) = a_1$, and we have

$$\pi = D_{(1,0,\dots,0)}. \tag{10.6.10}$$

Actually, to conclude (10.6.10), we need to know that π is irreducible. Here is a "lazy" proof. If not, π would split into irreducible pieces, each of which would have a non-raisable weight vector, and its weight must be a dominant integral weight. But among the weights (10.6.9), only λ_1^+ is dominant integral, and all the weight spaces are one-dimensional.

We compute the character of this representation, at points on the maximal torus $\mathbb{T} = \operatorname{Exp}(\mathfrak{h})$. We have

$$\operatorname{Exp}(\vartheta_a) = \begin{pmatrix} e^{ia_1} & & \\ & \ddots & \\ & & e^{ia_n} \end{pmatrix}, \tag{10.6.11}$$

and we want to compute the trace of the $(2n) \times (2n)$ matrix representing its action on $\mathbb{C}^n$, arising from the $\mathbb{C}$-linear isomorphism $\mathbb{H} \to \mathbb{C}^2$, given by

$$x + iy + jz + kw = (x + iy) + j(z - iw) \mapsto \begin{pmatrix} x + iy \\ z - iw \end{pmatrix}. \tag{10.6.12}$$

One has left multiplication by $\xi \in \mathbb{H}$ on $\mathbb{H}$ represented by the 2×2 complex matrix $M(\xi)$, given by

$$\begin{aligned} M(1) &= \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \ M(i) = \begin{pmatrix} i & \\ & -i \end{pmatrix}, \\ M(j) &= \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}, \ M(k) = \begin{pmatrix} & -i \\ -i & \end{pmatrix}. \end{aligned} \tag{10.6.13}$$

It follows that

$$\begin{aligned} \chi_\pi(\operatorname{Exp}(\vartheta_a)) &= 2 \operatorname{Re} \sum_j e^{ia_j} \\ &= 2 \sum_j \cos a_j. \end{aligned} \tag{10.6.14}$$

Representation of Sp(n) **on** $M(n, \mathbb{H})$

We define

$$\kappa : Sp(n) \longrightarrow \mathcal{L}_{\mathbb{R}}(M(n, \mathbb{H})), \quad \kappa(A)X = AXA^*, \tag{10.6.15}$$

for $A \in Sp(n)$, $X \in M(n, \mathbb{H})$. We give the $\mathbb{R}$-linear space $M(n, \mathbb{H})$ the Hilbert-Schmidt norm, parallel to (10.3.54), and the associated real inner product. Parallel to (10.3.55)–(10.3.56), we have that for each $A \in Sp(n)$, $\kappa(A)$ preserves the HS-norm on $M(n, \mathbb{H})$. Also, parallel to (10.3.57)–(10.3.60), we have the orthogonal decomposition

$$M(n, \mathbb{H}) = \mathfrak{sp}(n) \oplus \{cI : c \in \mathbb{R}\} \oplus S_0^2(\mathbb{H}^n), \tag{10.6.16}$$

where

$$\begin{aligned} S_0^2(\mathbb{H}^n) = \{X = (\xi_{\ell m}) \in M(n, \mathbb{H}) : \xi_{\ell\ell} \in \mathbb{R}, \ \xi_{\ell m} = \overline{\xi}_{m\ell}, \\ \xi_{11} + \cdots + \xi_{nn} = 0\}. \end{aligned} \tag{10.6.17}$$

We have the representation

$$\kappa_0 : Sp(n) \longrightarrow \mathcal{L}_{\mathbb{R}}(S_0^2(\mathbb{H}^n)). \tag{10.6.18}$$

Of course, $\kappa|_{\mathfrak{sp}(n)} = \operatorname{Ad}$, and κ acts trivially on $\{cI\}$. To study the weights, we extend these representations to the complexifications, e.g., to $\mathfrak{sp}(n)_{\mathbb{C}}$ (already considered earlier in this section) and to $S_0^2(\mathbb{H}^n)_{\mathbb{C}}$. The weights of Ad consist of the roots discussed above, plus the weight 0, whose associated weight space is $\mathfrak{h}_{\mathbb{C}}$, of dimension n.

We turn our attention to the representation κ_0. As in (10.3.64)–(10.3.65), it is instructive to compute $\kappa_0(A)$ for

$$A = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix}, \quad \alpha_j \in Sp(1). \tag{10.6.19}$$

If $X = (\xi_{\ell m})$ is as in (10.6.17), we have

$$\kappa_0(A)X = Y = (\eta_{\ell m}), \quad \eta_{\ell\ell} = \xi_{\ell\ell},\ \eta_{\ell m} = \alpha_\ell \xi_{\ell m} \overline{\alpha}_m,\ \eta_{m\ell} = \overline{\eta}_{\ell m}. \tag{10.6.20}$$

This leads to the following extension of Proposition 10.3.2.

Proposition 10.6.4. *In* (10.6.18),

$$\operatorname{Ker} \kappa_0 = \{\pm I\}. \tag{10.6.21}$$

Proof. $\operatorname{Ker} \kappa_0$ is a normal subgroup of $Sp(n)$. Take $x \in \operatorname{Ker} \kappa_0$. We know that

$$\mathbb{T} = \{\operatorname{diag}(e^{i\theta_1}, \dots, e^{i\theta_n}) : \theta_j \in \mathbb{R}\} \tag{10.6.22}$$

is a conjugating torus for $Sp(n)$. Hence there exists $g \in Sp(n)$ such that

$$g^{-1}xg = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \in \mathbb{T}, \tag{10.6.23}$$

and $g^{-1}xg \in \operatorname{Ker} \kappa_0$. The calculation (10.6.20) implies

$$\begin{aligned} &\alpha_\ell \xi_{\ell m} \overline{\alpha}_m = \xi_{\ell m},\ \forall\, \xi_{\ell m} \in \mathbb{H},\ \text{hence (taking } \xi_{\ell m} = 1) \\ &\alpha_\ell = \alpha_m,\ \text{and } \alpha_\ell \xi = \xi \alpha_\ell,\ \forall\, \xi \in \mathbb{H},\ \text{hence} \\ &\alpha_\ell \in \mathbb{R},\ \text{so } \alpha_1 = \cdots \alpha_n = \pm 1. \end{aligned} \tag{10.6.24}$$

This implies

$$g^{-1}xg = \pm I, \quad \text{hence} \quad x = \pm I. \tag{10.6.25}$$

□

Regarding the weights of κ_0, an argument parallel to that leading from (10.5.6) to Proposition 10.5.1, with $[i, u_\ell]$ replaced by $[i, \xi_{\ell\ell}] = 0$, yields the following result.

Proposition 10.6.5. *With respect to the isomorphism $\mathfrak{h} \approx \mathbb{R}^n$ defined by $\vartheta_a \leftrightarrow a$, the weights of the representation κ_0 of $Sp(n)$ on $S_0^2(\mathbb{H}^n)_{\mathbb{C}}$ are given by* (10.5.8), *i.e.,*

$$\rho_{\ell m}^{\alpha\beta}(a) = (-1)^\alpha a_\ell + (-1)^\beta a_m, \quad \ell < m,\ \alpha, \beta \in \{0, 1\}. \tag{10.6.26}$$

Note that

$$\text{the highest weight of } \kappa_0 \text{ is } \rho_{12}^{00}(a) = a_1 + a_2, \tag{10.6.27}$$

and this is the only dominant integral weight that occurs in (10.6.26).

Corollary 10.6.6. *The representation κ_0 is irreducible, and*

$$\kappa_0 = D_{(1,1,0,\dots,0)}. \tag{10.6.28}$$

Returning to the adjoint representation Ad of $Sp(n)$ on $\mathfrak{sp}(n)_{\mathbb{C}}$, we recall that its weights are the roots (10.5.7)–(10.5.8). Thus

$$\text{the highest weight of Ad is } \rho_1^+(a) = 2a_1. \tag{10.6.29}$$

In this case, there are two dominant integral weights that occur in the list of weights for Ad, namely ρ_1^+ and ρ_{12}^{00}. This leads to the following.

Proposition 10.6.7. *The adjoint representation* Ad *is an irreducible representation of* Sp(n) *on* $\mathfrak{sp}(n)_{\mathbb{C}}$*, and*

$$\mathrm{Ad} = D_{(2,0,\dots,0)}. \tag{10.6.30}$$

The reprsentations $\Lambda^k\pi$ on $\Lambda^k_{\mathbb{C}}\mathbb{H}^n$

The space $\mathbb{H}^n$, with $\mathbb{C}$ acting on the right, is a complex vector space isomorphic to $\mathbb{C}^{2n}$, and for each $k \in \{1, \dots, 2n\}$, $\Lambda^k\pi$ is a unitary representation of $Sp(n)$ on $\Lambda^k_{\mathbb{C}}\mathbb{H}^n$. We aim to prove the following.

Proposition 10.6.8. *For $1 \le k \le n$, the highest weight of $\Lambda^k\pi$ is $\lambda_1^+ + \cdots + \lambda_k^+$. Hence $\Lambda^k_{\mathbb{C}}\mathbb{H}^n$ contains a $\mathbb{C}$-linear subspace V_{nk} on which $\Lambda^k\pi$ acts irreducibly, with highest weight $\lambda_1^+ + \cdots + \lambda_k^+$. i.e., $Sp(n)$ acts on V_{nk} as*

$$D_{(1,\dots,1,\dots0)} \quad (k \text{ ones}). \tag{10.6.31}$$

Proof. The analysis of π on $\mathbb{H}^n$ yields an orthonormal basis $\{u_j^\pm : 1 \le j \le n\}$ of the $(2n)$-dimensional complex vector space $\mathbb{H}^n$, satisfying

$$\pi(\Gamma(a))u_j^\pm = e^{i\lambda_j^\pm(a)}u_j^\pm, \tag{10.6.32}$$

with

$$\Gamma(a) = \mathrm{Exp}(\vartheta_a). \tag{10.6.33}$$

Now, for $1 \le k \le 2n$, $\Lambda^k_{\mathbb{C}}\mathbb{H}^n$ is spanned by monomials

$$u_{j_1}^{s_1} \wedge \cdots \wedge u_{j_k}^{s_k}, \quad 1 \le j_1 \le \cdots \le j_k \le n,\ s_\nu \in \{\pm\}, \tag{10.6.34}$$

where the pairs (j_ν, s_ν) are all distinct, and we have

$$\begin{aligned}\Lambda^k\pi(\Gamma(a))&u_{j_1}^{s_1} \wedge \cdots \wedge u_{j_k}^{s_k} \\ &= e^{i[\lambda_{j_1}^{s_1}(a)+\cdots+\lambda_{j_k}^{s_k}(a)]}u_{j_1}^{s_1} \wedge \cdots \wedge u_{j_k}^{s_k}.\end{aligned} \tag{10.6.35}$$

Thus, for $1 \leq k \leq 2n$, all the weights of $\Lambda^k \pi$ are

$$\lambda_{j_1}^{s_1} + \cdots + \lambda_{j_k}^{s_k}, \tag{10.6.36}$$

with (j_ν, s_ν) as described above. This yields the assertion on the highest weight made in Proposition 10.6.8. If $\Lambda^k \pi$ is decomposed into irreducible factors, one of these must have $\lambda_1^+ + \cdots + \lambda_k^+$ as its highest weight, as long as $k \leq n$, and we conclude that (10.6.31) holds. □

Remark. If $k = n + j$, $1 \leq j \leq n$, we obtain that the highest weight of $\Lambda^k \pi$ is $\lambda_1^+ + \cdots + \lambda_{n-j}^+$. In such a case, $\Lambda^k \pi$ decomposes into irreducible factors, and one of them has this as its highest weight.

Let us focus on the case $k = 2$. Comparing (10.6.28) and (10.6.31), we see that $\Lambda^2 \pi$ contains κ_0. We have the following dimension count:

$$\dim_{\mathbb{C}} \Lambda^2_{\mathbb{C}} \mathbb{H}^n = \binom{2n}{2} = 2n^2 - n, \quad \dim_{\mathbb{C}} S^2_0(\mathbb{H}^n)_{\mathbb{C}} = 2n^2 - n - 1. \tag{10.6.37}$$

We conclude that

$$\Lambda^2 \pi \approx \kappa_0 \oplus 1, \tag{10.6.38}$$

where 1 denotes the trivial representation of $Sp(n)$ on $\mathbb{C}$.

Returning to the main thrust of Proposition 10.6.8, we see that, as a consequence, the theorem of the highest weight holds for $Sp(n)$. This has the following topological consequence.

Proposition 10.6.9. *For each $n \in \mathbb{N}$, the group $Sp(n)$ is simply connected.*

Proof. Results of Appendix E.3 imply that the universal covering group $\widetilde{Sp(n)}$ is compact. If this is a k-to-1 cover of $Sp(n)$ and $k > 1$, the Peter-Weyl theorem implies that it must have irreducible unitary representations with highest weight different from those on the list of irreducible representations of $Sp(n)$, hence with highest weights that are not dominant integral weights, an impossibility. □

Chapter 11

The Octonions and the group G_2

There is a continuation of the construction leading from the real numbers to the complex numbers that proceeds for two more steps, yielding first the quaternions (of dimension 4 over $\mathbb{R}$) and then the octonions (of dimension 8 over $\mathbb{R}$). As seen already in §1.2, a quaternion $\xi \in \mathbb{H}$ is given by

$$\xi = a + bi + cj + dk, \quad a, b, c, d \in \mathbb{R}. \tag{11.0.1}$$

Addition is performed componentwise, and multiplication is an $\mathbb{R}$-bilinear map $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$ in which 1 is a unit, products of distinct factors i, j, k behave like the cross product on $\mathbb{R}^3$, and $i^2 = j^2 = k^2 = -1$. This product is not commutative, but it is associative. An octonion $x \in \mathbb{O}$ is given by

$$x = (\xi, \eta), \quad \xi, \eta \in \mathbb{H}. \tag{11.0.2}$$

Addition is again defined componentwise, and multiplication is an $\mathbb{R}$-bilinear map $\mathbb{O} \times \mathbb{O} \to \mathbb{O}$, whose definition (given in §11.1) is a somewhat subtle modification of the definition of multiplication on $\mathbb{H}$. One major difference is that multiplication on $\mathbb{O}$ is no longer associative. Nevertheless, this non-associative algebra $\mathbb{O}$ has a very beautiful algebraic structure, whose study is the principal object of this chapter.

In §11.1 we define the product xy of two elements of $\mathbb{O}$ and establish basic properties. We introduce a norm and cross product, and study 4-dimensional subalgebras of $\mathbb{O}$ that are isomorphic to $\mathbb{H}$. These include $\mathcal{A} = \mathrm{Span}\{1, u_1, u_2, u_1u_2\}$, when u_1 and u_2 are orthonormal elements of $\mathrm{Im}(\mathbb{O})$. Analysis of the relationship of $\mathcal{A}$ and $\mathcal{A}^\perp$ gives rise to orthogonal linear maps on $\mathbb{O}$ that are shown to preserve the product, i.e., to automorphisms of $\mathbb{O}$, which form a group, denoted $\text{Aut}(\mathbb{O})$.

Section 11.2 goes further into $\text{Aut}(\mathbb{O})$, noting that it is a compact, connected Lie group of dimension 14, and analyzing some of its subgroups, including groups isomorphic to $SO(4)$ and groups isomorphic to $SU(3)$. Both of these types of subgroups contain two-dimensional tori. It is shown that $\text{Aut}(\mathbb{O})$ contains no tori of larger dimension. These facts are used in §11.3 to show that $\text{Aut}(\mathbb{O})$ is simple and to analyze its root system. The analysis reveals that $\text{Aut}(\mathbb{O})$ falls into the classification of compact simple Lie groups as the group denoted G_2. In §11.4 we make further comments about the Lie algebra of $\text{Aut}(\mathbb{O})$, including the fact that it has a $\mathbb{Z}/(3)$ grading.

The Lie group G_2 is the first in a list of 5 exceptional compact Lie groups, denoted G_2, F_4, E_6, E_7, and E_8, with complexified Lie algebras denoted $\mathfrak{G}_2, \mathfrak{F}_4, \mathfrak{E}_6, \mathfrak{E}_7$, and $\mathfrak{E}_8$.

Further material on the algebra of octonions, its automorphism group, and other concepts arising here can be found in a number of sources, including the survey article [**4**] and the books [**35**], [**30**], and [**21**].

Acknowledgment. Thanks to Robert Bryant for valuable conversations on the material of this chapter, particularly for sharing his insights on the Moufang identities.

11.1. Octonions

The set of octonions (also known as Cayley numbers) is a special but intriguing example of a nonassociative algebra. This space is

$$\mathbb{O} = \mathbb{H} \oplus \mathbb{H}, \tag{11.1.1}$$

with product given by

$$(\alpha, \beta) \cdot (\gamma, \delta) = (\alpha\gamma - \overline{\delta}\beta, \delta\alpha + \beta\overline{\gamma}), \quad \alpha, \beta, \gamma, \delta \in \mathbb{H}, \tag{11.1.2}$$

with conjugation $\delta \mapsto \overline{\delta}$ on $\mathbb{H}$ defined as in §10.1. We mention that, with $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}$, the product in $\mathbb{H}$ is also given by (11.1.2), with $\alpha, \beta, \gamma, \delta \in \mathbb{C}$. Furthermore, with $\mathbb{C} = \mathbb{R} \oplus \mathbb{R}$, the product in $\mathbb{C}$ is given by (11.1.2), with $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. In the setting of $\mathbb{O} = \mathbb{H} \oplus \mathbb{H}$, the product in (11.1.2) is clearly $\mathbb{R}$-bilinear, but it is neither commutative nor associative. However, it does retain a vestige of associativity, namely

$$x(yz) = (xy)z \ \text{ whenever any two of } x, y, z \in \mathbb{O} \text{ coincide.} \tag{11.1.3}$$

We define a conjugation on $\mathbb{O}$:

$$x = (\alpha, \beta) \Longrightarrow \overline{x} = (\overline{\alpha}, -\beta). \tag{11.1.4}$$

We set $\operatorname{Re} x = (x + \overline{x})/2 = (\operatorname{Re} \alpha, 0)$. Note that $a = \operatorname{Re} x$ lies in the center of $\mathbb{O}$ (i.e., commutes with each element of $\mathbb{O}$), and $\overline{x} = 2a - x$. It is straightforward to check that

$$x, y \in \mathbb{O} \Longrightarrow \operatorname{Re} xy = \operatorname{Re} yx. \tag{11.1.5}$$

We have a decomposition

$$x = a + u, \quad a = \operatorname{Re} x, \ u = x - \operatorname{Re} x = \operatorname{Im} x, \tag{11.1.6}$$

parallel to (10.1.4). Again we call u the vector part of x, and we say that $u \in \operatorname{Im}(\mathbb{O})$. If also $y = b + v$, then

$$xy = ab + av + bu + uv, \tag{11.1.7}$$

with a similar formula for yx, yielding

$$xy - yx = uv - vu. \tag{11.1.8}$$

We now define the inner product

$$\langle x, y \rangle = \operatorname{Re}(x\overline{y}), \quad x, y \in \mathbb{O}. \tag{11.1.9}$$

To check symmetry, note that if $x = a + u, \ y = b + v$,

$$\langle x, y \rangle = ab - \operatorname{Re}(uv), \tag{11.1.10}$$

and (11.1.5) then implies

$$\langle x, y \rangle = \langle y, x \rangle. \tag{11.1.11}$$

In fact, (11.1.9) yields the standard Euclidean inner product on $\mathbb{O} \approx \mathbb{R}^8$, with square norm $|x|^2 = \sqrt{\langle x, x\rangle}$. We have

$$x = (\alpha, \beta) \Longrightarrow x\overline{x} = (\alpha\overline{\alpha} + \overline{\beta}\beta, 0) = (|x|^2, 0). \tag{11.1.12}$$

As a consequence, we see that

$$x \in \mathbb{O},\ x \neq 0,\ y = |x|^{-2}\overline{x} \Longrightarrow xy = yx = 1, \tag{11.1.13}$$

where $1 = (1, 0)$ is the multiplicative unit in $\mathbb{O}$.

Returning to conjugation on $\mathbb{O}$, we have, parallel to (10.1.20),

$$x, y \in \mathbb{O} \Longrightarrow \overline{xy} = \overline{y}\,\overline{x}, \tag{11.1.14}$$

via a calculation using the definition (11.1.2) of the product. Using the decomposition $x = a + u,\ y = b + v$, this is equivalent to $\overline{uv} = vu$, and since $\overline{uv} = 2\,\mathrm{Re}(uv) - uv = -2\langle u, v\rangle - uv$, this is equivalent to

$$u, v \in \mathrm{Im}(\mathbb{O}) \Longrightarrow uv + vu = -2\langle u, v\rangle. \tag{11.1.15}$$

In turn, (11.1.15) follows by expanding $(u + v)^2$ and using $w^2 = -|w|^2$ for $w \in \Im(\mathbb{O})$, with $w = u, v$, and $u + v$. We next establish the following parallel to (10.1.22).

Proposition 11.1.1. *Given $x, y \in \mathbb{O}$,*

$$|xy| = |x|\,|y|. \tag{11.1.16}$$

Proof. To begin, we bring in the following variant of (11.1.3),

$$x, y \in \mathbb{O} \Longrightarrow (xy)(yx) = ((xy)y)x, \tag{11.1.17}$$

which can be verified from the definition (11.1.2) of the product. Taking into account $\overline{x} = 2a - x,\ \overline{y} = 2b - y$, and (11.1.14), we have

$$\begin{aligned}(xy)(\overline{xy}) &= (xy)(\overline{y}\,\overline{x}) = ((xy)\overline{y})\overline{x} \\ &= (x|y|^2)\overline{x} = |x|^2|y|^2,\end{aligned} \tag{11.1.18}$$

which gives (11.1.16), since $|xy|^2 = (xy)(\overline{xy})$. ∎

Continuing to pursue parallels with §10.1, we define a cross product on $\mathrm{Im}(\mathbb{O})$ as follows. Given $u, v \in \mathrm{Im}(\mathbb{O})$, set

$$u \times v = \frac{1}{2}(uv - vu). \tag{11.1.19}$$

By (11.1.5), this is an element of $\mathrm{Im}(\mathbb{O})$. Also, if $x = a + u, y = b + v$,

$$xy - yx = 2u \times v. \tag{11.1.20}$$

Compare (10.1.6). Putting together (11.1.15) and (11.1.19), we have

$$uv = -\langle u, v\rangle + u \times v, \quad u, v \in \mathrm{Im}(\mathbb{O}). \tag{11.1.21}$$

Hence

$$|uv|^2 = |\langle u, v\rangle|^2 + |u \times v|^2. \tag{11.1.22}$$

Now (11.1.16) implies $|uv|^2 = |u|^2|v|^2$, and of course $\langle u, v\rangle = |u|\,|v| \cos\theta$, where θ is the angle between u and v. Hence, parallel to (10.1.26),

$$|u \times v|^2 = |u|^2|v|^2|\sin\theta|^2, \quad \forall\, u, v \in \mathrm{Im}(\mathbb{O}). \tag{11.1.23}$$

We have the following complement.

Proposition 11.1.2. *If $u, v \in \mathrm{Im}(\mathbb{O})$, then*

$$w = u \times v \Longrightarrow \langle w, u\rangle = \langle w, v\rangle = 0. \tag{11.1.24}$$

Proof. We know that $w \in \mathrm{Im}(\mathbb{O})$. Hence, by (11.1.21),

$$\begin{aligned} \langle w, v\rangle &= \langle uv, v\rangle = \mathrm{Re}((uv)\overline{v}) \\ &= \mathrm{Re}(u(v\overline{v})) = |v|^2 \,\mathrm{Re}\, u = 0, \end{aligned} \tag{11.1.25}$$

the third identity by (11.1.3) (applicable since $\overline{v} = -v$). The proof that $\langle w, u\rangle = 0$ is similar. ☐

Returning to basic observations about the product (11.1.2), we note that it is uniquely determined as the $\mathbb{R}$-bilinear map $\mathbb{O} \times \mathbb{O} \to \mathbb{O}$ satisfying

$$\begin{aligned} (\alpha, 0) \cdot (\gamma, 0) &= (\alpha\gamma, 0), \quad (0, \beta) \cdot (\gamma, 0) = (0, \beta\overline{\gamma}), \\ (\alpha, 0) \cdot (0, \delta) &= (0, \delta\alpha), \quad (0, \beta) \cdot (0, \delta) = (-\overline{\delta}\beta, 0), \end{aligned} \tag{11.1.26}$$

for $\alpha, \beta, \gamma, \delta \in \mathbb{H}$. In particular, $\mathbb{H} \oplus 0$ is a subalgebra of $\mathbb{O}$, isomorphic to $\mathbb{H}$. As we will see, $\mathbb{O}$ has lots of subalgebras isomorphic to $\mathbb{H}$. First, let us label the "standard" basis of $\mathbb{O}$ as

$$\begin{aligned} 1 = (1, 0), \quad e_1 = (i, 0), \quad e_2 = (j, 0), \quad e_3 = (k, 0), \\ f_0 = (0, 1), \quad f_1 = (0, i), \quad f_2 = (0, j), \quad f_3 = (0, k), \end{aligned} \tag{11.1.27}$$

and describe the associated multiplication table. The mutiplication table for $1, e_1, e_2, e_3$ is the same as (10.1.2)–(10.1.3), of course. We also have $f_\ell^2 = -1$ and all the distinct e_ℓ and f_m anticommute. These results are special cases of the fact that

$$u, v \in \mathrm{Im}(\mathbb{O}),\ |u| = 1,\ \langle u, v\rangle = 0 \Longrightarrow u^2 = -1 \ \text{ and } \ uv = -vu, \tag{11.1.28}$$

which is a consequence of (11.1.15).

To proceed with the multiplication table for $\mathbb{O}$, note that (11.1.26) gives

$$(\alpha, 0) f_0 = (0, \alpha), \tag{11.1.29}$$

so

$$e_\ell f_0 = f_\ell, \quad 1 \le \ell \le 3. \tag{11.1.30}$$

By (11.1.28), $f_0 e_\ell = -f_\ell$. Using the notation $\varepsilon_1 = i, \varepsilon_2 = j, \varepsilon_3 = k \in \mathbb{H}$, we have

$$e_\ell f_m = (\varepsilon_\ell, 0) \cdot (0, \varepsilon_m) = (0, \varepsilon_m \varepsilon_\ell), \quad 1 \le \ell, m \le 3, \tag{11.1.31}$$

and the multiplication table (10.1.2)–(10.1.3) gives the result as $-f_0$ if $\ell = m$, and $\pm f_\mu$ if $\ell \neq m$, where $\{\ell, m, \mu\} = \{1, 2, 3\}$. Again by (11.1.28), $f_m e_\ell = -e_\ell f_m$. To complete the multiplication table, we have

$$f_0 f_m = (0, 1) \cdot (0, \varepsilon_m) = (\varepsilon_m, 0) = e_m, \quad 1 \le m \le 3, \tag{11.1.32}$$

and

$$f_\ell f_m = (0, \varepsilon_\ell) \cdot (0, \varepsilon_m) = (\varepsilon_m \varepsilon_\ell, 0) = e_m e_\ell, \quad 1 \le \ell, m \le 3. \tag{11.1.33}$$

The following is a succinct summary of the results described above on the multiplication table for $\mathbb{O}$. In each row listed in (11.1.34), consisting of three elements (say u_j), $\mathrm{Span}\{1, u_1, u_2, u_3\}$ is an algebra, isomorphic to $\mathbb{H}$ under $i \mapsto u_1, j \mapsto u_2, k \mapsto u_3$.

$$\begin{array}{ccc} i & j & k \\ e_1 & e_2 & e_3 \\ e_\ell & f_0 & f_\ell \\ f_1 & e_3 & f_2 \\ f_2 & e_1 & f_3 \\ f_3 & e_2 & f_1 \end{array} \tag{11.1.34}$$

See Figure 11.1.1 for a diagram of this multiplication table. In each case, one has a triple recorded in (11.1.34), lying along a line (or a circle), equipped with an arrow indicating the appropriate order.

We turn to the general task of constructing subalgebras of $\mathbb{O}$, having just seen several examples. To start, pick

$$u_1 \in \mathrm{Im}(\mathbb{O}), \quad \text{such that } \ |u_1| = 1. \tag{11.1.35}$$

By (11.1.28), $u_1^2 = -1$, and we have the subalgebra of $\mathbb{O}$,

$$\mathrm{Span}\{1, u_1\} \approx \mathbb{C}. \tag{11.1.36}$$

To proceed, pick

$$u_2 \in \mathrm{Im}(\mathbb{O}), \quad \text{such that } \ |u_2| = 1 \text{ and } \langle u_1, u_2 \rangle = 0, \tag{11.1.37}$$

and set

$$u_3 = u_1 u_2. \tag{11.1.38}$$

By (11.1.28),

$$u_2^2 = -1, \ \text{and } u_2 u_1 = -u_1 u_2 = -u_3. \tag{11.1.39}$$

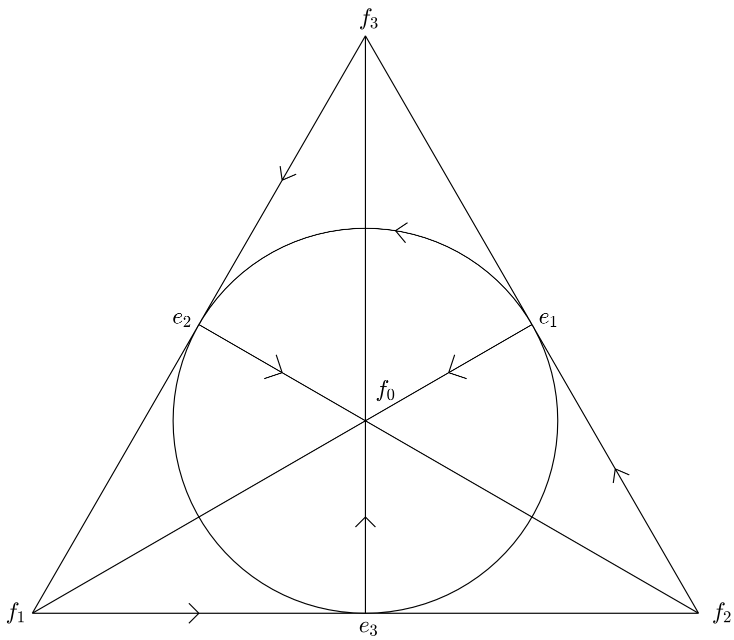

Figure 11.1.1. Multiplication table schematic for $\mathbb{O}$

Note that

$$\operatorname{Re} u_3 = \operatorname{Re}(u_1 u_2) = -\langle u_1, u_2 \rangle = 0. \tag{11.1.40}$$

Also, by (11.1.16), $|u_3| = 1$, so

$$1 = u_3 \overline{u}_3 = -u_3^2. \tag{11.1.41}$$

Furthermore, by (11.1.3),

$$\begin{aligned} u_1 u_3 &= u_1(u_1 u_2) = (u_1 u_1) u_2 = -u_2, \quad \text{and} \\ u_3 u_2 &= (u_1 u_2) u_2 = u_1(u_2 u_2) = -u_1. \end{aligned} \tag{11.1.42}$$

Let us also note that

$$u_3 = u_1 \times u_2. \tag{11.1.43}$$

Hence, by Proposition 11.1.2,

$$\langle u_3, u_1 \rangle = \langle u_3, u_2 \rangle = 0, \tag{11.1.44}$$

and, again by (11.1.28), $u_3 u_1 = -u_1 u_3$ and $u_2 u_3 = -u_3 u_2$. Thus we have for each such choice of u_1 and u_2 a subalgebra of $\mathbb{O}$,

$$\operatorname{Span}\{1, u_1, u_2, u_3\} \approx \mathbb{H}. \tag{11.1.45}$$

At this point we can make the following observation.

Proposition 11.1.3. *Given any two elements $x_1, x_2 \in \mathbb{O}$, the algebra $\mathcal{A}$ generated by $1, x_1$, and x_2 is isomorphic to either $\mathbb{R}, \mathbb{C}$, or $\mathbb{H}$. In particular, it is associative.*

Proof. Consider $V = \operatorname{Span}\{1, x_1, x_2\}$. If $\dim V = 1$, then $\mathcal{A} \approx \mathbb{R}$. If $\dim V = 2$, the argument yielding (11.1.36) gives $\mathcal{A} \approx \mathbb{C}$. If $\dim V = 3$, then $\Im x_1$ and $\Im x_2$ are linearly independent. We can pick orthonormal elements u_1 and u_2 in their span. Then $\mathcal{A}$ is the algebra generated by $1, u_1$, and u_2, and the analysis (11.1.35)–(11.1.45) gives $\mathcal{A} \approx \mathbb{H}$. $\square$

The last assertion of Proposition 11.1.3 contains (11.1.3) and (11.1.17) as special cases. The failure of $\mathbb{O}$ to be associative is clearly illustrated by (11.1.31), which implies

$$e_\ell(e_m f_0) = (e_m e_\ell) f_0, \quad \text{for } 1 \le \ell, m \le 3, \tag{11.1.46}$$

so

$$e_\ell(e_m f_0) = -(e_\ell e_m) f_0, \quad \text{if } \ell \neq m.$$

Bringing in also (11.1.33) yields

$$f_\ell(e_m f_0) = e_m e_\ell, \quad \text{while } (f_\ell e_m) f_0 = e_\ell e_m. \tag{11.1.47}$$

We next explore how the subalgebra

$$\mathcal{A} = \operatorname{Span}\{1, u_1, u_2, u_3\}, \tag{11.1.48}$$

from (11.1.45), interacts with its orthogonal complement $\mathcal{A}^\perp$. Pick

$$v_0 \in \mathcal{A}^\perp, \quad |v_0| = 1. \tag{11.1.49}$$

Note that $v_0 \in \operatorname{Im}(\mathbb{O})$. Taking a cue from (11.1.30), we set

$$v_\ell = u_\ell v_0, \quad 1 \le \ell \le 3. \tag{11.1.50}$$

Note that $\operatorname{Re} v_\ell = -\langle u_\ell, v_0 \rangle = 0$, so $v_\ell \in \operatorname{Im}(\mathbb{O})$. We claim that

$$\{v_0, v_1, v_2, v_3\} \text{ is an orthonormal set in } \mathbb{O}. \tag{11.1.51}$$

To show this, we bring in the following operators. Given $x \in \mathbb{O}$, define the $\mathbb{R}$-linear maps

$$L_x, R_x : \mathbb{O} \longrightarrow \mathbb{O}, \quad L_x y = xy, \ R_x y = yx. \tag{11.1.52}$$

By (11.1.16), for $y \in \mathbb{O}$,

$$|x| = 1 \Longrightarrow |L_x y| = |R_x y| = |y|. \tag{11.1.53}$$

Hence L_x and R_x are orthogonal transformations. Since the unit sphere in $\mathbb{O}$ is connected, $\det L_x$ and $\det R_x$ are $\equiv 1$ for such x, so

$$|x| = 1 \Longrightarrow L_x, R_x \in SO(\mathbb{O}). \tag{11.1.54}$$

Hence $R_{v_0} \in SO(\mathbb{O})$. Since

$$v_0 = R_{v_0}1, \quad v_\ell = R_{v_0}u_\ell \ \text{ for } \ 1 \le \ell \le 3, \tag{11.1.55}$$

we have (11.1.51). We next claim that

$$v_\ell \perp u_m, \quad \forall\, \ell, m \in \{1,2,3\}. \tag{11.1.56}$$

In fact, since $L_{u_\ell} \in SO(\mathbb{O})$,

$$\begin{aligned} \langle v_\ell, u_m\rangle &= \langle u_\ell v_0, u_m\rangle = \langle u_\ell(u_\ell v_0), u_\ell u_m\rangle \\ &= \langle (u_\ell u_\ell)v_0, u_\ell u_m\rangle = -\langle v_0, u_\ell u_m\rangle = 0, \end{aligned} \tag{11.1.57}$$

the third identity by (11.1.3).

It follows that

$$\mathcal{A}^\perp = \operatorname{Span}\{v_0, v_1, v_2, v_3\}. \tag{11.1.58}$$

Consequently

$$\{1, u_1, u_2, u_3, v_0, v_1, v_2, v_3\} \ \text{ is an orthonormal basis of } \ \mathbb{O}. \tag{11.1.59}$$

Results above imply that

$$R_{v_0} : \mathcal{A} \xrightarrow{\approx} \mathcal{A}^\perp. \tag{11.1.60}$$

Such an argument applies to any unit length $v \perp \mathcal{A}$. Consequently

$$x \in \mathcal{A},\ y \in \mathcal{A}^\perp \Longrightarrow xy \in \mathcal{A}^\perp. \tag{11.1.61}$$

Noting that if also $x \in \operatorname{Im}(\mathbb{O})$ then $xy = -yx$, we readily deduce that

$$x \in \mathcal{A},\ y \in \mathcal{A}^\perp \Longrightarrow yx \in \mathcal{A}^\perp. \tag{11.1.62}$$

Furthermore, since $|x| = 1 \Rightarrow L_x, R_x \in SO(\mathbb{O})$, we have

$$x \in \mathcal{A}^\perp \Longrightarrow L_x, R_x : \mathcal{A}^\perp \longrightarrow \mathcal{A}, \tag{11.1.63}$$

hence

$$x, y \in \mathcal{A}^\perp \Longrightarrow xy \in \mathcal{A}. \tag{11.1.64}$$

Note that for the special case

$$\mathcal{H} = \mathbb{H} \oplus 0, \quad \mathcal{H}^\perp = 0 \oplus \mathbb{H}, \tag{11.1.65}$$

the results (11.1.61)–(11.1.64) follow immediately from (11.1.26).

We have the following important result about the correspondence between the bases (11.1.27) and (11.1.59) of $\mathbb{O}$.

Proposition 11.1.4. *Let $u_\ell, v_\ell \in \operatorname{Im}(\mathbb{O})$ be given as in* (11.1.48)–(11.1.50). *Then the orthogonal transformation $K : \mathbb{O} \to \mathbb{O}$, defined by*

$$K1 = 1, \quad Ke_\ell = u_\ell, \quad Kf_\ell = v_\ell, \tag{11.1.66}$$

preserves the product on $\mathbb{O}$:

$$K(xy) = K(x)K(y), \quad \forall\, x, y \in \mathbb{O}. \tag{11.1.67}$$

That is to say, K is an automorphism of $\mathbb{O}$.

Proof. What we need to show is that $\{u_1, u_2, u_3, v_0, v_1, v_2, v_3\}$ has the same multiplication table as $\{e_1, e_2, e_3, f_0, f_1, f_2, f_3\}$. That products involving only $\{u_\ell\}$ have such behavior follows from the arguments leading to (11.1.45). That $e_\ell f_0 = f_\ell$ is paralleled by $u_\ell v_0 = v_\ell$, for $1 \le \ell \le 3$, is the definition (3.49). It remains to show that the products $u_\ell v_m$ and $v_\ell v_m$ mirror the products $e_\ell f_m$ and $f_\ell f_m$, as given in (11.1.31)–(11.1.33).

First, we have, for $1 \le m \le 3$,

$$v_0 v_m = -v_m v_0 = -(u_m v_0)v_0 = -u_m(v_0 v_0) = u_m, \tag{11.1.68}$$

mirroring (11.1.32). Mirroring the case $\ell = m$ of (11.1.31), we have

$$u_\ell v_\ell = u_\ell(u_\ell v_0) = (u_\ell u_\ell)v_0 = -v_0. \tag{11.1.69}$$

The analogue of (11.1.31) for $\ell = m$ is simple, thanks to (11.1.15):

$$v_\ell v_\ell = -1. \tag{11.1.70}$$

It remains to establish the following:

$$u_\ell v_m = (u_m u_\ell)v_0, \quad v_\ell v_m = u_m u_\ell, \quad \text{for} \ \ 1 \le \ell, m \le 3, \ \ell \ne m. \tag{11.1.71}$$

Expanded out, the required identities are

$$u_\ell(u_m v_0) = (u_m u_\ell)v_0, \quad 1 \le \ell, m \le 3, \ \ell \ne m, \tag{11.1.72}$$

and

$$(u_\ell v_0)(u_m v_0) = u_m u_\ell, \quad 1 \le \ell, m \le 3, \ \ell \ne m. \tag{11.1.73}$$

$\square$

Such identities as (11.1.72)–(11.1.73) are closely related to an important class of identities known as "Moufang identities," which we now introduce.

Proposition 11.1.5. *Given $x, y, z \in \mathbb{O}$,*

$$(xyx)z = x(y(xz)), \quad z(xyx) = ((zx)y)x, \tag{11.1.74}$$

and

$$(xy)(zx) = x(yz)x. \tag{11.1.75}$$

Regarding the paucity of parentheses here, we use the notation xwx to mean

$$xwx = (xw)x = x(wx), \tag{11.1.76}$$

the last identity by (11.1.3). Note also that the two identities in (11.1.74) are equivalent, respectively, to

$$L_{xyx} = L_x L_y L_x, \quad \text{and} \quad R_{xyx} = R_x R_y R_x. \tag{11.1.77}$$

A proof of Proposition 11.1.5 will be given in §11.2. We now show how (11.1.74)–(11.1.75) can be used to establish (11.1.72)–(11.1.73).

We start with (11.1.73), which is equivalent to

$$(v_0 u_\ell)(u_m v_0) = u_\ell u_m. \tag{11.1.78}$$

In this case, (11.1.75) yields

$$\begin{aligned} (v_0 u_\ell)(u_m v_0) &= v_0(u_\ell u_m)v_0 \\ &= -(u_\ell u_m)v_0 v_0 \quad (\text{if } \ell \neq m) \\ &= u_\ell u_m, \end{aligned} \tag{11.1.79}$$

via a couple of applications of (11.1.15). This gives (11.1.73).

Moving on, applying L_{v_0}, we see that (11.1.72) is equivalent to

$$v_0(u_\ell(u_m v_0)) = v_0(u_m u_\ell)v_0, \tag{11.1.80}$$

hence to

$$v_0(u_\ell(v_0 u_m)) = v_0(u_\ell u_m)v_0. \tag{11.1.81}$$

Now the first identity in (11.1.74) implies that the left side of (11.1.81) is equal to

$$(v_0 u_\ell v_0)u_m = u_\ell u_m, \tag{11.1.82}$$

the latter identity because $v_0 u_\ell v_0 = -u_\ell v_0 v_0 = u_\ell$. On the other hand, if $\ell \neq m$, then

$$v_0(u_\ell u_m)v_0 = -(u_\ell u_m)v_0 v_0 = u_\ell u_m, \tag{11.1.83}$$

agreeing with the right side of (11.1.82). Thus we have (11.1.81), hence (11.1.72).

Rather than concluding that Proposition 11.1.4 is now proved, we must reveal that the proof of Proposition 11.1.5 given in the next section actually uses Proposition 11.1.4. Therefore, it is necessary to produce an alternative endgame to the proof of Proposition 11.1.4.

We begin by noting that the approach to the proof of Proposition 11.1.4 described above uses the identities (11.1.74)–(11.1.75) with

$$x = v_0, \quad y = u_\ell, \quad z = u_m, \quad \ell \neq m, \tag{11.1.84}$$

hence $xy = -v_\ell, zx = v_m, yz = \pm u_h$, $\{h, \ell, m\} = \{1, 2, 3\}$. Thus the application of the first identity of (11.1.74) in (11.1.82) is justified by the following special case of (11.1.77):

Proposition 11.1.6. *If $\{u, v\} \in \mathrm{Im}(\mathbb{O})$ is an orthonormal set, then*

$$L_{uvu} = L_v = L_u L_v L_u. \tag{11.1.85}$$

Proof. Under these hypotheses, $u^2 = -1$ and $uv = -vu$. Bringing in (3.3), we have

$$uvu = -u^2 v = v, \tag{11.1.86}$$

which gives the first identity in (11.1.85). We also have

$$a \in \mathrm{Im}(\mathbb{O}) \Longrightarrow L_a^2 = L_{a^2} = -|a|^2 I, \tag{11.1.87}$$

the first identity by (11.1.3). Thus

$$\begin{aligned} -2I = L_{(u+v)}^2 &= (L_u + L_v)(L_u + L_v) \\ &= L_u^2 + L_v^2 + L_u L_v + L_v L_u, \end{aligned} \tag{11.1.88}$$

so

$$L_u L_v = -L_v L_u, \tag{11.1.89}$$

and hence

$$L_u L_v L_u = -L_v L_u^2 = L_v, \tag{11.1.90}$$

giving the second identity in (11.1.85). ☐

As for the application of (11.1.75) to (11.1.79), we need the special case

$$(uv)(wu) = u(vw)u, \tag{11.1.91}$$

for $u = v_0, v = u_\ell, w = u_m, \ell \neq m, 1 \leq \ell, m \leq 3$ (so $uv = -v_\ell$), in which cases

$$\{u, v, w, uv\},\ \{u, vw\} \subset \mathrm{Im}(\mathbb{O}),\ \text{are orthonormal sets.} \tag{11.1.92}$$

In such a case, $u(vw)u = -(vw)u^2 = vw$, so it suffices to show that

$$(uv)(wu) = vw, \tag{11.1.93}$$

for

$$\{u, v, w, uv\} \subset \mathrm{Im}(\mathbb{O}), \quad \text{orthonormal.} \tag{11.1.94}$$

When (11.1.94) holds, we say $\{u, v, w\}$ is a *Cayley triangle*. The following takes care of our needs.

Proposition 11.1.7. *Assume* $\{u, v, w\}$ *is a Cayley triangle. Then*

$$v(uw) = -(vu)w, \tag{11.1.95}$$

$$\langle uv, uw \rangle = 0, \quad \textit{so } \{u, v, uw\} \textit{ is a Cayley triangle,} \tag{11.1.96}$$

and (11.1.93) *holds.*

Proof. To start, the hypotheses imply

$$vu = -uv, \quad vw = -wv, \quad uw = -wu, \quad (vu)w = -w(vu), \tag{11.1.97}$$

so

$$\begin{aligned} v(uw) + (vu)w &= -v(wu) - w(vu) \\ &= (v^2 + w^2)u - (v + w)(vu + wu) \\ &= (v + w)^2 u - (v + w)((v + w)u) \\ &= 0, \end{aligned} \tag{11.1.98}$$

and we have (11.1.95). Next,

$$\langle uv, uw \rangle = \langle L_u v, L_u w \rangle = \langle u, w \rangle = 0, \tag{11.1.99}$$

since $L_u \in SO(\mathbb{O})$. Thus $\{u, v, uw\}$ is a Cayley triangle. Applying (11.1.95) to this Cayley triangle (and bringing in (11.1.3)) then gives

$$\begin{aligned} (vu)(uw) &= -v(u(uw)) \\ &= -v(u^2 w) \\ &= vw, \end{aligned} \tag{11.1.100}$$

yielding (11.1.93). $\square$

At this point, we have a complete proof of Proposition 11.1.4. The proof of Proposition 11.1.5 will be given in the following section.

11.2. The automorphism group of $\mathbb{O}$

The set of automorphisms of $\mathbb{O}$ is denoted $\mathrm{Aut}(\mathbb{O})$. Note that $\mathrm{Aut}(\mathbb{O})$ is a group, i.e.,

$$K_j \in \mathrm{Aut}(\mathbb{O}) \Longrightarrow K_1K_2,\ K_j^{-1} \in \mathrm{Aut}(\mathbb{O}). \tag{11.2.1}$$

Clearly $K \in \mathrm{Aut}(\mathbb{O}) \Rightarrow K1 = 1$. The following result will allow us to establish a converse to Proposition 11.1.4.

Proposition 11.2.1. *Assume $K \in \mathrm{Aut}(\mathbb{O})$. Then*

$$K : \mathrm{Im}(\mathbb{O}) \longrightarrow \mathrm{Im}(\mathbb{O}). \tag{11.2.2}$$

Consequently

$$K\overline{x} = \overline{Kx}, \quad \forall\, x \in \mathbb{O}, \tag{11.2.3}$$

and

$$|Kx| = |x|, \quad \forall\, x \in \mathbb{O}, \tag{11.2.4}$$

so $K : \mathbb{O} \to \mathbb{O}$ is an orthogonal transformation.

Proof. To start, we note that, given $x \in \mathbb{O}$, x^2 is real if and only if either x is real or $x \in \mathrm{Im}(\mathbb{O})$. Now, given $u \in \mathrm{Im}(\mathbb{O})$,

$$(Ku)^2 = K(u^2) = -|u|^2K1 = -|u|^2 \quad \text{(real)}, \tag{11.2.5}$$

so either $Ku \in \mathrm{Im}(\mathbb{O})$ or $Ku = a$ is real. In the latter case, we have $K(a^{-1}u) = 1$, so $a^{-1}u = 1$, so $u = a$, contradicting the hypothesis that $u \in \mathrm{Im}(\mathbb{O})$. This gives (11.2.2). The result (11.2.3) is an immediate consequence. Thus, for $x \in \mathbb{O}$,

$$|Kx|^2 = (Kx)(\overline{Kx}) = (Kx)(K\overline{x}) = K(x\overline{x}) = |x|^2, \tag{11.2.6}$$

giving (4.4). □

Now, given $K \in \mathrm{Aut}(\mathbb{O})$, define u_1, u_2, and v_0 by

$$u_1 = Ke_1, \quad u_2 = Ke_2, \quad v_0 = Kf_0. \tag{11.2.7}$$

By Proposition 11.2.1, these are orthonormal elements of $\mathrm{Im}(\mathbb{O})$. Also, $\mathcal{A} = K(\mathcal{H})$, spanned by $1, u_1, u_2$, and $u_1u_2 = u_1 \times u_2$, is a subalgebra of $\mathbb{O}$, and $v_0 \in \mathcal{A}^\perp$. These observations, together with Proposition 11.1.4, yield the following.

Proposition 11.2.2. *The formulas* (11.2.7) *provide a one-to-one correspondence between the set of automorphisms of $\mathbb{O}$ and*

(11.2.8)
the set of ordered orthonormal triples (u_1, u_2, v_0) in $\mathrm{Im}(\mathbb{O})$
such that v_0 is also orthogonal to $u_1 \times u_2$, that is,
the set of Cayley triangles in $\mathrm{Im}(\mathbb{O})$.

It can be deduced from (11.2.8) that $\mathrm{Aut}(\mathbb{O})$ is a compact, connected Lie group of dimension 14.

We return to the Moufang identities and use the results on $\mathrm{Aut}(\mathbb{O})$ established above to prove them.

Proof of Proposition 11.1.5. Consider the first identity in (11.1.74), i.e.,

$$(xyx)z = x(y(xz)), \quad \forall\, x, y, z \in \mathbb{O}. \tag{11.2.9}$$

We begin with a few simple observations. First, (11.2.9) is clearly true if any one of x, y, z is scalar, or if any two of them coincide (thanks to Proposition 11.1.3). Also, both sides of (11.2.9) are linear in y and in z. Thus, it suffices to treat (11.2.9) for $y, z \in \mathrm{Im}(\mathbb{O})$. Meanwhile, multiplying by a real number and applying an element of $\mathrm{Aut}(\mathbb{O})$, we can assume $x = a + e_1$, for some $a \in \mathbb{R}$.

To proceed, (11.2.9) is clear for $y \in \mathrm{Span}(1, x)$, so, using the linearity in y, and applying Proposition 11.2.2 again, we can arrange that $y = e_2$. Given this, (11.2.9) is clear for $z \in \mathcal{H} = \mathrm{Span}(1, e_2, e_2, e_3 = e_1e_2)$. Thus, using linearity of (11.2.9) in z, it suffices to treat $z \in \mathcal{H}^\perp$, and again applying an element of $\mathrm{Aut}(\mathbb{O})$, we can assume $z = f_1$.

At this point, we have reduced the task of proving (11.2.9) to checking it for

$$x = a + e_1, \quad y = e_2, \quad z = f_1, \quad a \in \mathbb{R}, \tag{11.2.10}$$

and this is straightforward. Similar arguments applied to the second identity in (11.1.74), and to (11.1.75), reduce their proofs to a check in the case (11.2.10). $\square$

We next look at some interesting subgroups of $\mathrm{Aut}(\mathbb{O})$. Taking $Sp(1)$ to be the group of unit quaternions, as in (10.1.27), we have group homomorphisms

$$\alpha, \beta : Sp(1) \longrightarrow \mathrm{Aut}(\mathbb{O}), \tag{11.2.11}$$

given by

$$\begin{aligned} \alpha(\xi)(\zeta, \eta) &= (\xi\zeta\overline{\xi}, \xi\eta\overline{\xi}), \\ \beta(\xi)(\zeta, \eta) &= (\zeta, \xi\eta), \end{aligned} \tag{11.2.12}$$

where $\zeta, \eta \in \mathbb{H}$ define $(\zeta, \eta) \in \mathbb{O}$. As in (10.1.33)–(10.1.38), for $\xi \in Sp(1)$, $\pi(\xi)\zeta = \xi\zeta\overline{\xi}$ gives an automorphism of $\mathbb{H}$, and it commutes with conjugation in $\mathbb{H}$, so the fact that $\alpha(\xi)$ is an automorphism of $\mathbb{O}$ follows from the definition (11.1.2) of the product in $\mathbb{O}$. The fact that $\beta(\xi)$ is an automorphism of $\mathbb{O}$ also follows directly from (11.1.2). Parallel to (10.1.36),

$$\mathrm{Ker}\, \alpha = \{\pm 1\} \subset Sp(1), \tag{11.2.13}$$

so the image of $Sp(1)$ under α is a subgroup of $\mathrm{Aut}(\mathbb{O})$ isomorphic to $SO(3)$. Clearly β is one-to-one, so it yields a subgroup of $\mathrm{Aut}(\mathbb{O})$ isomorphic to $Sp(1)$.

These two subgroups of $\mathrm{Aut}(\mathbb{O})$ do not commute with each other. In fact, we have, for $\xi_j \in Sp(1)$, $(\zeta, \eta) \in \mathbb{O}$,

$$\begin{aligned} \alpha(\xi_1)\beta(\xi_2)(\zeta,\eta) &= (\xi_1\zeta\overline{\xi}_1, \xi_1\xi_2\eta\overline{\xi}_1), \\ \beta(\xi_2)\alpha(\xi_1)(\zeta,\eta) &= (\xi_1\zeta\overline{\xi}_1, \xi_2\xi_1\eta\overline{\xi}_1). \end{aligned} \tag{11.2.14}$$

Note that, since $\xi_2\xi_1 = \xi_1(\overline{\xi}_1\xi_2\xi_1)$,

$$\beta(\xi_2)\alpha(\xi_1) = \alpha(\xi_1)\beta(\overline{\xi}_1\xi_2\xi_1). \tag{11.2.15}$$

It follows that

$$G_{\mathcal{H}} = \{\alpha(\xi_1)\beta(\xi_2) : \xi_j \in Sp(1)\} \tag{11.2.16}$$

is a subgroup of $\mathrm{Aut}(\mathbb{O})$. It is clear from (11.2.12) that each automorphism $\alpha(\xi_1), \beta(\xi_2)$, and hence each element of $G_{\mathcal{H}}$, preserves $\mathcal{H}$ (and also $\mathcal{H}^\perp$). The converse also holds:

Proposition 11.2.3. *The group $G_{\mathcal{H}}$ is the group of all automorphisms of $\mathbb{O}$ that preserve $\mathcal{H}$.*

Proof. Indeed, suppose $K \in \mathrm{Aut}(\mathbb{O})$ preserves $\mathcal{H}$. Then $K|_{\mathcal{H}}$ is an automorphism of $\mathcal{H} \approx \mathbb{H}$. Arguments in the paragraph containing (10.1.37)–(10.1.40) imply that there exists $\xi_1 \in Sp(1)$ such that $K|_{\mathcal{H}} = \alpha(\xi_1)|_{\mathcal{H}}$, so $K_0 = \alpha(\xi_1)^{-1}K \in \mathrm{Aut}(\mathbb{O})$ is the identity on $\mathcal{H}$. Now $K_0 f_1 = (0, \xi_2)$ for some $\xi_2 \in Sp(1)$, and it then follows from Proposition 11.2.2 that $K_0 = \beta(\xi_2)$. Hence $K = \alpha(\xi_1)\beta(\xi_2)$, as desired. ☐

For another perspective on $G_{\mathcal{H}}$, we bring in

$$\tilde{\alpha} : Sp(1) \longrightarrow \mathrm{Aut}(\mathbb{O}), \quad \tilde{\alpha}(\xi) = \beta(\overline{\xi})\alpha(\xi). \tag{11.2.17}$$

Note that

$$\tilde{\alpha}(\xi)(\zeta,\eta) = (\xi\zeta\overline{\xi}, \eta\overline{\xi}), \tag{11.2.18}$$

so $\tilde{\alpha}$ is a group homomorphism. Another easy consequence of (11.2.18) is that $\tilde{\alpha}(\xi_1)$ and $\beta(\xi_2)$ commute, for each $\xi_j \in Sp(1)$. We have a surjective group homomorphism

$$\tilde{\alpha} \times \beta : Sp(1) \times Sp(1) \longrightarrow G_{\mathcal{H}}. \tag{11.2.19}$$

Note that $\mathrm{Ker}(\tilde{\alpha} \times \beta) = \{(1,1), (-1,-1)\}$, with 1 denoting the unit in $\mathbb{H}$. It follows that

$$G_{\mathcal{H}} \approx SO(4). \tag{11.2.20}$$

We now take a look at one-parameter families of automorphisms of $\mathbb{O}$, of the form

$$K(t) = e^{tA}, \quad A \in \mathcal{L}(\mathbb{O}), \tag{11.2.21}$$

where e^{tA} is the matrix exponential, studied in §1.3. To see when such linear transformations on $\mathbb{O}$ are automorphisms, we differentiate the identity

$$K(t)(xy) = (K(t)x)(K(t)y), \quad x, y \in \mathbb{O}, \tag{11.2.22}$$

obtaining

$$A(xy) = (Ax)y + x(Ay), \quad x, y \in \mathbb{O}. \tag{11.2.23}$$

When (11.2.23) holds, we say

$$A \in \mathrm{Der}(\mathbb{O}). \tag{11.2.24}$$

Proposition 11.2.4. *Given $A \in \mathcal{L}(\mathbb{O})$, $e^{tA} \in \mathrm{Aut}(\mathbb{O})$ for all $t \in \mathbb{R}$ if and only if $A \in \mathrm{Der}(\mathbb{O})$.*

Proof. The implication $\Rightarrow$ was established above. For the converse, suppose A satisfies (11.2.23). Take $x, y \in \mathbb{O}$, and set

$$X(t) = (e^{tA}x)(e^{tA}y). \tag{11.2.25}$$

Applying d/dt gives

$$\begin{aligned} \frac{dX}{dt} &= (Ae^{tA}x)(e^{tA}y) + (e^{tA}x)(Ae^{tA}y) \\ &= A\big((e^{tA}x)(e^{tA}y)\big) \\ &= AX(t), \end{aligned} \tag{11.2.26}$$

the second identity by (11.2.23). Since $X(0) = xy$, it follows from the standard uniqueness argument of ODE, cf. Chapter 3 of [**44**], that

$$X(t) = e^{tA}(xy), \tag{11.2.27}$$

so indeed $e^{tA} \in \mathrm{Aut}(\mathbb{O})$. $\square$

The set $\mathrm{Der}(\mathbb{O})$ has the following structure.

Proposition 11.2.5. *The set $\mathrm{Der}(\mathbb{O})$ is a linear subspace of $\mathcal{L}(\mathbb{O})$ satisfying*

$$A, B \in \mathrm{Der}(\mathbb{O}) \Longrightarrow [A, B] \in \mathrm{Der}(\mathbb{O}), \tag{11.2.28}$$

where $[A, B] = AB - BA$. That is, $\mathrm{Der}(\mathbb{O})$ is a Lie algebra.

Proof. That $\mathrm{Der}(\mathbb{O})$ is a linear space is clear from the defining property (4.23). Furthermore, if $A, B \in \mathrm{Der}(\mathbb{O})$, then, for all $x, y \in \mathbb{O}$,

$$\begin{aligned} AB(xy) &= A((Bx)y) + A(x(By)) \\ &= (ABx)y + (Bx)(Ay) + (Ax)(By) + x(ABy), \end{aligned} \tag{11.2.29}$$

and similarly

$$BA(xy) = (BAx)y + (Ax)(By) + (Bx)(Ay) + x(BAy), \tag{11.2.30}$$

so

$$[A, B](xy) = ([A, B]x)y + x([A, B]y), \tag{11.2.31}$$

and we have (11.2.28). □

By Proposition 11.2.1, if $A \in \mathrm{Der}(\mathbb{O})$, then e^{tA} is an orthogonal transformation for each $t \in \mathbb{R}$. We have

$$(e^{tA})^* = e^{tA^*}, \tag{11.2.32}$$

so

$$A \in \mathrm{Der}(\mathbb{O}) \Longrightarrow A^* = -A, \tag{11.2.33}$$

i.e., A is skew-adjoint. It is clear that

$$A \in \mathrm{Der}(\mathbb{O}) \Longrightarrow A : \mathrm{Im}(\mathbb{O}) \to \mathrm{Im}(\mathbb{O}), \tag{11.2.34}$$

and since $\mathrm{Im}(\mathbb{O})$ is odd dimensional, the structural analysis in Chapter 2, §11 of [**44**] implies

$$A \in \mathrm{Der}(\mathbb{O}) \Longrightarrow \mathcal{N}(A) \cap \mathrm{Im}(\mathbb{O}) \neq 0. \tag{11.2.35}$$

As long as $A \neq 0$, we can also deduce from Proposition 11.4 in Chapter 2 of [**44**] that $\mathrm{Im}(\mathbb{O})$ contains a two-dimensional subspace with orthonormal basis $\{u_1, u_2\}$, invariant under A, and with respect to which A is represented by a 2×2 block

$$\begin{pmatrix} 0 & -\lambda \\ \lambda & 0 \end{pmatrix}. \tag{11.2.36}$$

Then, by (11.2.23),

$$\begin{aligned} A(u_1u_2) &= (Au_1)u_2 + u_1(Au_2) \\ &= \lambda u_2^2 - \lambda u_1^2 \\ &= 0, \end{aligned} \tag{11.2.37}$$

so $u_1u_2 = u_1 \times u_2 \in \mathcal{N}(A) \cap \mathrm{Im}(\mathbb{O})$. As in (11.1.37)–(11.1.45), $\mathrm{Span}\{1, u_1, u_2, u_3 = u_1u_2\} = \mathcal{A}$ is a subalgebra of $\mathbb{O}$ isomorphic to $\mathbb{H}$. We see that A preserves $\mathcal{A}$, so the associated one-parameter group of automorphisms e^{tA} preserves $\mathcal{A}$.

Using Proposition 11.2.2, we can pick $K \in \mathrm{Aut}(\mathbb{O})$ taking $\mathcal{A}$ to $\mathcal{H}$, and deduce the following.

Proposition 11.2.6. *Given* $A \in \mathrm{Der}(\mathbb{O})$*, there exists* $K \in \mathrm{Aut}(\mathbb{O})$ *such that*

$$Ke^{tA}K^{-1} \in G_{\mathcal{H}}, \quad \forall\, t \in \mathbb{R}. \tag{11.2.38}$$

Note that then

$$Ke^{tA}K^{-1} = e^{t\widetilde{A}}, \quad \widetilde{A} = KAK^{-1} \in \mathrm{Der}(\mathbb{O}), \tag{11.2.39}$$

and (11.2.38) is equivalent to

$$\widetilde{A} : \mathcal{H} \longrightarrow \mathcal{H}, \quad \widetilde{A} \in \mathrm{Der}(\mathbb{O}), \tag{11.2.40}$$

which also entails $\widetilde{A} : \mathcal{H}^{\perp} \to \mathcal{H}^{\perp}$, since $\widetilde{A}$ is skew-adjoint. When (11.2.40) holds, we say

$$\widetilde{A} \in D_{\mathcal{H}}. \tag{11.2.41}$$

Going further, suppose we have d commuting elements of $\mathrm{Der}(\mathbb{O})$:

$$A_j \in \mathrm{Der}(\mathbb{O}), \quad A_jA_k = A_kA_j, \quad j,k \in \{1,\dots,d\}. \tag{11.2.42}$$

A modification of the arguments leading to Proposition 11.4 of [**44**] yields a two-dimensional subspace of $\mathrm{Im}(\mathbb{O})$, with orthonormal basis $\{u_1, u_2\}$, invariant under each A_j, with respect to which each A_j is represented by a 2×2 block as in (11.2.36), with λ replaced by λ_j (possibly 0). As in (11.2.37),

$$A_j(u_1u_2) = 0, \quad 1 \le j \le d, \tag{11.2.43}$$

so each A_j preserves $\mathcal{A} = \mathrm{Span}\{1, u_1, u_2, u_3 = u_1u_2\}$, and so does each one-parameter group of automorphisms e^{tA_j}. Bringing in $K \in \mathrm{Aut}(\mathbb{O})$, taking $\mathcal{A}$ to $\mathcal{H}$, we have the following variant of Proposition 11.2.6.

Proposition 11.2.7. *Given commuting $A_j \in \mathrm{Der}(\mathbb{O})$, $1 \le j \le d$, there exists $K \in \mathrm{Aut}(\mathbb{O})$ such that*

$$Ke^{tA_j}K^{-1} \in G_{\mathcal{H}}, \quad \forall\, t \in \mathbb{R},\ j \in \{1,\dots,d\}. \tag{11.2.44}$$

As a consequence, we have

$$\widetilde{A}_j = KA_jK^{-1} \in D_{\mathcal{H}}, \quad \widetilde{A}_j\widetilde{A}_k = \widetilde{A}_k\widetilde{A}_j, \quad 1 \le j,k \le d. \tag{11.2.45}$$

Consequently, $e^{t\widetilde{A}_j}$ are mutually commuting one-parameter subgroups of $G_{\mathcal{H}}$, i.e.,

$$e^{t_j\widetilde{A}_j} \in G_{\mathcal{H}}, \quad e^{t_j\widetilde{A}_j}e^{t_k\widetilde{A}_k} = e^{t_k\widetilde{A}_k}e^{t_j\widetilde{A}_j}, \quad 1 \le j,k \le d. \tag{11.2.46}$$

One can produce *pairs* of such commuting groups, as follows. Take

$$\tilde{\alpha}(\xi_1(t_1)),\ \beta(\xi_2(t_2)) \in G_{\mathcal{H}}, \tag{11.2.47}$$

with β as in (11.2.11)–(11.2.12), $\tilde{\alpha}$ as in (11.2.17)–(11.2.18), and $\xi_\nu(t)$ one-parameter subgroups of $Sp(1)$, for example

$$\xi_\nu(t) = e^{t\omega_\nu}, \quad \omega_\nu \in \mathrm{Im}(\mathbb{H}) = \mathrm{Span}\{i,j,k\}. \tag{11.2.48}$$

The exponential $e^{t\omega_\nu}$ is amenable to a treatment parallel to that given in §1.3. Mutual commutativity in (11.2.47) follows from the general mutual commutativity of $\tilde{\alpha}$ and β. The following important structural information on $\mathrm{Aut}(\mathbb{O})$ says $d = 2$ is as high as one can go.

Proposition 11.2.8. *If $A_j \in \mathrm{Der}(\mathbb{O})$ are mutually commuting, for $j \in \{1, \dots, d\}$, and if $\{A_j\}$ is linearly independent in $\mathcal{L}(\mathbb{O})$, then $d \le 2$.*

Proof. To start, we obtain from A_j the mutually commuting one-parameter groups $Ke^{tA_j}K^{-1}$, subgroups of $G_{\mathcal{H}}$. Taking inverse images under the two-to-one surjective homomorphism (11.2.19), we get mutually commuting one-parameter subgroups $\gamma_j(t)$ of $Sp(1) \times Sp(1)$, which can be written

$$\gamma_j(t) = \begin{pmatrix} e^{\omega_j t} & \\ & e^{\sigma_j t} \end{pmatrix}, \quad \omega_j, \sigma_j \in \mathrm{Im}(\mathbb{H}), \quad 1 \le j \le d. \tag{11.2.49}$$

Parallel to Proposition 3.7.6 of [**42**], this commutativity requires $\{\omega_j : 1 \le j \le d\}$ to commute in $\mathbb{H}$ and it also requires $\{\sigma_j : 1 \le j \le d\}$ to commute in $\mathbb{H}$. These conditions in turn require each ω_j to be a real multiple of some $\omega^\# \in \mathrm{Im}(\mathbb{H})$ and each σ_j to be a real multiple of some $\sigma^\# \in \mathrm{Im}(\mathbb{H})$.

Now the linear independence of $\{A_j : 1 \le j \le d\}$ in $\mathrm{Der}(\mathbb{O})$ implies the linear independence of $\{(\omega_j, \sigma_j) : 1 \le j \le d\}$ in $\mathrm{Im}(\mathbb{H}) \oplus \mathrm{Im}(\mathbb{H})$, and this implies $d \le 2$. $\square$

We turn to the introduction of another interesting subgroup of $\mathrm{Aut}(\mathbb{O})$. Note that, by Proposition 11.2.2, given any unit $u_1 \in \mathrm{Im}(\mathbb{O})$, there exists $K \in \mathrm{Aut}(\mathbb{O})$ such that $Ke_1 = u_1$. Consequently, $\mathrm{Aut}(\mathbb{O})$, acting on $\Im(\mathbb{O})$ as a group of orthogonal transformations, acts *transitively* on the unit sphere S in $\mathrm{Im}(\mathbb{O}) \approx \mathbb{R}^7$, i.e., on $S \approx S^6$. We are hence interested in the group

$$\{K \in \mathrm{Aut}(\mathbb{O}) : Ke_1 = e_1\} = \mathcal{G}_{e_1}. \tag{11.2.50}$$

We claim that

$$\mathcal{G}_{e_1} \approx SU(3). \tag{11.2.51}$$

As preparation for the demonstration, note that each $K \in \mathcal{G}_{e_1}$ is an orthogonal linear transformation on $\mathbb{O}$ that leaves invariant $\mathrm{Span}\{1, e_1\}$, and hence it also leaves invariant the orthogonal complement

$$V = \mathrm{Span}\{1, e_1\}^\perp = \mathrm{Span}\{e_2, e_3, f_0, f_1, f_2, f_3\}, \tag{11.2.52}$$

a linear space of $\mathbb{R}$-dimension 6. We endow V with a complex structure. Generally, a complex structure on a real vector space V is an $\mathbb{R}$-linear map $J : V \to V$ such that $J^2 = -I_V$. One can check that this requires $\dim_{\mathbb{R}} V$ to be even, say $2k$. Then (V, J) has the structure of a complex vector space, with

$$(a + ib)v = av + bJv, \quad a, b \in \mathbb{R}, \ v \in V. \tag{11.2.53}$$

One has $\dim_{\mathbb{C}}(V, J) = k$. If V is a real inner product space, with inner product $\langle\ ,\ \rangle$, and if J is orthogonal (hence skew-adjoint) on V, then (V, J)

gets a natural Hermitian inner product

$$(u,v) = \langle u,v\rangle + i\langle u,Jv\rangle. \tag{11.2.54}$$

If $T : V \to V$ preserves $\langle\ ,\ \rangle$ and commutes with J, then it also preserves $(\ ,\ )$, so it is a unitary transformation on (V,J).

We can apply this construction to V as in (11.2.52), with

$$Jv = L_{e_1}v = e_1 v, \tag{11.2.55}$$

noting that L_{e_1} is an orthogonal map on $\mathbb{O}$ that preserves $\mathrm{Span}\{1,e_1\}$, and hence also preserves V. To say that an $\mathbb{R}$-linear map $K : V \to V$ is $\mathbb{C}$-linear is to say that $K(e_1v) = e_1K(v)$, for all $v \in V$. Clearly this holds if $K \in \mathrm{Aut}(\mathbb{O})$ and $Ke_1 = e_1$. Thus each element of $\mathcal{G}_{e_1}$ defines a complex linear orthogonal (hence unitary) transformation on V, and we have an injective group homomorphism

$$\mathcal{G}_{e_1} \longrightarrow U(V,J). \tag{11.2.56}$$

Note that the 6 element real orthonormal basis of V in (11.2.52) yields the 3 element orthonormal basis of (V,J),

$$\{e_2, f_0, f_2\}, \tag{11.2.57}$$

since

$$e_3 = e_1e_2, \quad f_1 = e_1f_0, \quad f_3 = -e_1f_2, \tag{11.2.58}$$

the latter two identities by (11.1.30)–(11.1.31). This choice of basis yields the isomorphism

$$U(V,J) \approx U(3). \tag{11.2.59}$$

We aim to identify the image of $\mathcal{G}_{e_1}$ in $U(3)$ that comes from (11.2.56) and (11.2.59).

To accomplish this, we reason as follows. From Proposition 11.2.2 it follows that there is a natural one-to-one correspondence between the elements of $\mathcal{G}_{e_1}$ and

$$\begin{aligned}&\text{the set of ordered orthonormal pairs } \{u_2, v_0\} \text{ in } V\\ &\text{such that also } v_0 \perp e_1u_2,\end{aligned} \tag{11.2.60}$$

or, equivalently,

$$\text{the set of ordered orthonormal pairs } \{u_2, v_0\} \text{ in } (V,J), \tag{11.2.61}$$

where (V,J) carries the Hermitian inner product (11.2.54). In fact, the correspondence associates to $K \in \mathcal{G}_{e_1}$ (i.e., $K \in \mathrm{Aut}(\mathbb{O})$ and $Ke_1 = e_1$) the pair

$$u_2 = Ke_2, \quad v_0 = Kf_0. \tag{11.2.62}$$

Then the image of $\mathcal{G}_{e_1}$ in $U(V, J)$ in (11.2.56) is uniquely determined by the action of K on the third basis element in (11.2.57), as

$$Kf_2 = K(e_2 f_0) = K(e_2)K(f_0) = u_2 v_0 = u_2 \times v_0, \tag{11.2.63}$$

where we recall from (11.1.30) that $f_2 = e_2 f_0$, and the last identity in (11.2.63) follows from (11.1.21).

From (11.2.60)–(11.2.61), it can be deduced that $\mathcal{G}_{e_1}$ is a compact, connected Lie group of dimension 8. Then (11.2.55) and (11.2.58) present $\mathcal{G}_{e_1}$ as isomorphic to a subgroup (call it $\widetilde{\mathcal{G}}$) of $U(3)$ that is a compact, connected Lie group of dimension 8. Meanwhile, $\dim U(3) = 9$, so $\widetilde{\mathcal{G}}$ has codimension 1. We claim that this implies

$$\widetilde{\mathcal{G}} = SU(3). \tag{11.2.64}$$

We sketch a proof of (11.2.64).

To start, one can show that a connected, codimension-one subgroup of a compact, connected Lie group must be *normal*. Hence $\widetilde{\mathcal{G}}$ is a normal subgroup of $U(3)$. This implies $U(3)/\widetilde{\mathcal{G}}$ is a group. This quotient is a compact Lie group of dimension 1, hence isomorphic to $S^1 = \{z \in \mathbb{C} : |z| = 1\}$, and the projection $U(3) \to U(3)/\widetilde{\mathcal{G}}$ produces a continuous, surjective group homomorphism

$$\vartheta : U(3) \longrightarrow S^1, \quad \operatorname{Ker} \vartheta = \widetilde{\mathcal{G}}. \tag{11.2.65}$$

Now a complete list of such homomorphisms is given by

$$\vartheta_j(K) = (\det K)^j, \quad j \in \mathbb{Z} \setminus 0, \tag{11.2.66}$$

and in such a case, $\operatorname{Ker} \vartheta_j$ has $|j|$ connected components. Then connectivity of $\widetilde{\mathcal{G}}$ forces $\vartheta = \vartheta_{\pm 1}$ in (11.2.65), which in turn gives (11.2.64).

It is useful to take account of various subgroups of $\operatorname{Aut}(\mathbb{O})$ that are conjugate to $G_{\mathcal{H}}$ (given by (11.2.16)) or to $\mathcal{G}_{e_1}$ (given by (11.2.50)). In particular, when $\mathcal{A} \subset \mathbb{O}$ is a four-dimensional subalgebra, we set

$$G_{\mathcal{A}} = \{K \in \operatorname{Aut}(\mathbb{O}) : K(\mathcal{A}) \subset \mathcal{A}\}, \tag{11.2.67}$$

and if $u \in \operatorname{Im}(\mathbb{O}), |u| = 1$, we set

$$\mathcal{G}_u = \{K \in \operatorname{Aut}(\mathbb{O}) : Ku = u\}. \tag{11.2.68}$$

We see that each group $G_{\mathcal{A}}$ is conjugate to $G_{\mathcal{H}}$, and isomorphic to $SO(4)$, and each group $\mathcal{G}_u$ is conjugate to $\mathcal{G}_{e_1}$, and isomorphic to $SU(3)$.

It is of interest to look at $\mathcal{G}_u \cap \mathcal{G}_v$, where u and v are unit elements of $\operatorname{Im}(\mathbb{O})$ that are not collinear. Then

$$\mathcal{G}_u \cap \mathcal{G}_v = \{K \in \operatorname{Aut}(\mathbb{O}) : K = I \text{ on } \operatorname{Span}\{u, v\}\}. \tag{11.2.69}$$

Now we can write $\mathrm{Span}\{u, v\} = \mathrm{Span}\{u_1, u_2\}$, with $u_1 = u, u_2 \perp u_1$, and note that $Ku_j = u_j \Rightarrow K(u_1u_2) = u_1u_2$, so (4.69) is equal to

$$\mathcal{G}_{\mathcal{A}} = \{K \in \mathrm{Aut}(\mathbb{O}) : K = I \text{ on } \mathcal{A}\}, \tag{11.2.70}$$

where $\mathcal{A} = \mathrm{Span}\{1, u_1, u_2, u_1u_2\}$ is a four-dimensional subalgebra of $\mathbb{O}$. Clearly

$$\mathcal{G}_{\mathcal{A}} \subset G_{\mathcal{A}}, \quad \text{and } \ \mathcal{G}_{\mathcal{A}} \approx Sp(1) \approx SU(2). \tag{11.2.71}$$

In fact, $\mathcal{G}_{\mathcal{A}}$ is conjugate to $\mathcal{G}_{\mathcal{H}} = \beta(Sp(1))$, with β as in (11.2.11)–(11.2.12).

Extending (11.2.52), we have associated to each unit $u \in \mathrm{Im}(\mathbb{O})$ the space

$$V_u = \mathrm{Span}\{1, u\}^{\perp}, \tag{11.2.72}$$

and $L_u : V_u \to V_u$ gives a complex structure $J_u = L_u|_{V_u}$, so (V_u, J_u) is a three-dimensional complex vector space. Parallel to (11.2.56), we have an injective group homomorphism

$$\mathcal{G}_u \longrightarrow U(V_u, J_u), \tag{11.2.73}$$

whose image is a codimension-one subgroup isomorphic to $SU(3)$. Associated to the family (V_u, J_u) is the following interesting geometrical structure. Consider the unit sphere $S \approx S^6$ in $\mathrm{Im}(\mathbb{O})$. There is a natural identification of V_u with the tangent space T_uS to S at u:

$$T_uS = V_u, \tag{11.2.74}$$

and the collection of complex structures J_u gives S what is called an *almost complex structure*. Now an element $K \in \mathrm{Aut}(\mathbb{O})$ acts on S, thanks to Proposition 11.2.1. Furthermore, for each $u \in S$,

$$K : V_u \longrightarrow V_{Ku} \tag{11.2.75}$$

is an isometry, and it is $\mathbb{C}$-linear, since

$$v \in V_u \Longrightarrow K(uv) = K(u)K(v). \tag{11.2.76}$$

Thus $\mathrm{Aut}(\mathbb{O})$ acts as a group of rotations on S that preserve its almost complex structure. In fact, this property characterizes $\mathrm{Aut}(\mathbb{O})$. To state this precisely, we bring in the following notation. Set

$$\iota : \mathrm{Aut}(\mathbb{O}) \longrightarrow SO(\mathrm{Im}(\mathbb{O})), \quad \iota(K) = K\big|_{\Im(\mathbb{O})}. \tag{11.2.77}$$

This is an injective group homomorphism, whose image we denote

$$A^b(\mathbb{O}) = \iota\, \mathrm{Aut}(\mathbb{O}). \tag{11.2.78}$$

The inverse of the isomorphism $\iota : \mathrm{Aut}(\mathbb{O}) \to A^b(\mathbb{O})$ is given by

$$\begin{gathered} j\Big|_{A^b(\mathbb{O})}, \quad j : SO(\mathrm{Im}(\mathbb{O})) \to SO(\mathbb{O}), \\ J(K_0)(a+u) = a + K_0u. \end{gathered} \tag{11.2.79}$$

Our result can be stated as follows.

Proposition 11.2.9. *The group* Γ *of rotations on* $\mathrm{Im}(\mathbb{O})$ *that preserve the almost complex structure of* S *is equal to* $A^b(\mathbb{O})$.

Proof. We have seen that $A^b(\mathbb{O}) \subset \Gamma$. It remains to prove that $\Gamma \subset A^b(\mathbb{O})$, so take $K_0 \in \Gamma$, and set $K = j(K_0)$, as in (11.2.79). We need to show that $K \in \mathrm{Aut}(\mathbb{O})$. First, one readily checks that, if $K = j(K_0)$, then

$$K \in \mathrm{Aut}(\mathbb{O}) \Longleftrightarrow K(uv) = K(u)K(v), \ \forall\, u, v \in \mathrm{Im}(\mathbb{O}), \tag{11.2.80}$$

and furthermore we can take $|u| = 1$. Now the condition $K_0 \in \Gamma$ implies

$$K_0(uv) = K_0(u)K_0(v), \quad \forall\, u \in \mathrm{Im}(\mathbb{O}),\ v \in V_u. \tag{11.2.81}$$

To finish the argument, we simply note that if $K_0 \in \Gamma$ and $K = j(K_0)$, and if u is a unit element of $\mathrm{Im}(\mathbb{O})$ and $v \in V_u$, then for all $a \in \mathbb{R}$,

$$\begin{aligned} K(u(au+v)) &= K(-a+uv) \\ &= -a + K_0(uv) \\ &= -a + K_0(u)K_0(v), \end{aligned} \tag{11.2.82}$$

while

$$\begin{aligned} (Ku)(K(au+v)) &= (K_0u)(aK_0u + K_0v) \\ &= a(K_0u)^2 + (K_0u)(K_0v) \\ &= -a + K_0(u)K_0(v). \end{aligned} \tag{11.2.83}$$

This finishes the proof. □

Further results on almost complex 6-dimensional submanifolds, including submanifolds of $\mathbb{O}$, can be found in [**6**] and [**7**].

11.3. Simplicity and root structure of Aut($\mathbb{O}$)

Our first goal in this section is to establish the following.

Proposition 11.3.1. *The group* Aut($\mathbb{O}$) *is simple.*

We will deduce this from the facts that Aut($\mathbb{O}$) is a compact, connected Lie group of dimension 14 and that it has rank 2. We recall from basic Lie group theory that if G is a compact Lie group, it has a maximal torus, and any two such are conjugate. The dimension of such a maximal torus is the rank of G. That Aut($\mathbb{O}$) has rank 2 follows from Proposition 11.2.8. The following general result basically does the trick.

Proposition 11.3.2. *Let G be a compact Lie group of rank 2. If its Lie algebra $\mathfrak{g}$ has a non-trivial ideal, $\mathfrak{h}$, then* $\dim G \le 6$.

Proof. Give $\mathfrak{g}$ an ad-invariant inner product. If $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then ad $\mathfrak{g}$ preserves both $\mathfrak{h}$ and $\mathfrak{h}^\perp$, so $\mathfrak{h}^\perp$ is also an ideal, and each $X \in \mathfrak{h}$ commutes with each $Y \in \mathfrak{h}^\perp$. Now if $\mathfrak{h}$ and $\mathfrak{h}^\perp$ are both nonzero,

$$\text{Rank } \mathfrak{g} = 2 \Longrightarrow \text{Rank } \mathfrak{h} = \text{ Rank } \mathfrak{h}^\perp = 1.$$

But, as is well known,

$$\text{Rank } \mathfrak{h} = 1 \Longrightarrow \dim \mathfrak{h} = 1 \quad \text{or} \quad 3,$$

so we have the conclusion that $\dim G \le 6$. □

It follows from Proposition 11.3.2 that the Lie algebra Der($\mathbb{O}$) of Aut($\mathbb{O}$) has no nontrivial ideals. A connected Lie group with this property is typically said to be simple.

Remark. Our analysis of Aut($\mathbb{O}$) as a compact simple Lie group, of rank 2 and dimension 14, implies that it is isomorphic to the group G_2, introduced in §6.6.

Going further, we establish the more precise result that Aut($\mathbb{O}$) contains no nontrivial normal subgroups. Indeed, if H were such a subgroup, so would be its closure, so it suffices to consider the case when H is closed. (The reader can show that a proper *dense* subgroup of a noncommutative, connected Lie group cannot be normal.) Then H is a Lie group, and Proposition 11.3.2 implies that either $H = \text{Aut}(\mathbb{O})$ or H is discrete, hence finite. In such a case, H normal implies H is the center of Aut($\mathbb{O}$), so our task is reduced to showing

$$\text{Aut}(\mathbb{O}) \ \text{ has trivial center.} \tag{11.3.1}$$

Indeed, suppose K_0 belongs to the center of Aut($\mathbb{O}$). Then K_0 belongs to a one-parameter subgroup e^{tA}, and (11.2.35) applies, to yield $u \in S \subset \text{Im}(\mathbb{O})$, fixed under the action of e^{tA}, hence fixed by K_0. Then, for each $K \in \mathrm{Aut}(\mathbb{O})$, $KK_0K^{-1} = K_0$ fixes Ku, and since $\mathrm{Aut}(\mathbb{O})$ acts transitively on the unit sphere $S \subset \mathrm{Im}(\mathbb{O})$, K_0 must fix each point of $\mathrm{Im}(\mathbb{O})$, so $K_0 = I$, and we have (11.3.1).

Our next goal is to analyze the root structure of $\mathrm{Aut}(\mathbb{O})$. We start by recalling the general definition. Let G be a compact, connected Lie group, with maximal torus $\mathbb{T}$, having associated Lie algebras $\mathfrak{t} \subset \mathfrak{g}$. Give $\mathfrak{g}$ an Ad-invariant inner product. The adjoint representation Ad of G on $\mathfrak{g}_{\mathbb{C}}$ has derived Lie algebra representation ad of $\mathfrak{g}$ by skew-adjoint transformations on $\mathfrak{g}_{\mathbb{C}}$, which simultaneously diagonalize when restricted to $\mathfrak{t}$. We have the root space decomposition

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{t}_{\mathbb{C}} \oplus \bigoplus_{\alpha} \mathfrak{g}_{\alpha}, \tag{11.3.2}$$

where, given $\alpha \in \mathfrak{t}'$, $\alpha \neq 0$,

$$\mathfrak{g}_{\alpha} = \{z \in \mathfrak{g}_{\mathbb{C}} : [x, z] = i\alpha(x)z,\ \forall\, x \in \mathfrak{t}\}. \tag{11.3.3}$$

If $\mathfrak{g}_{\alpha} \neq 0$, we call α a root, and nonzero elements of $\mathfrak{g}_{\alpha}$ are called root vectors. As sen in §6.1, we have

$$[\mathfrak{g}_{\alpha}, \mathfrak{g}_{\beta}] \subset \mathfrak{g}_{\alpha+\beta}, \tag{11.3.4}$$

that

$$\alpha \ \text{ root } \Longrightarrow \dim \mathfrak{g}_{\alpha} = 1, \tag{11.3.5}$$

and that if $\mathfrak{z}$ denotes the center of $\mathfrak{g}$,

$$\mathfrak{z} = 0 \Longrightarrow \ \text{the set of roots spans } \ \mathfrak{t}'. \tag{11.3.6}$$

Before tackling the particulars for $G = \mathrm{Aut}(\mathbb{O})$, we recall the most classical case $SU(n)$, from §§4.2–4.5.

The group $SU(n)$ has maximal torus

$$\mathbb{T} = \left\{ \begin{pmatrix} e^{ix_1} & & \\ & \ddots & \\ & & e^{ix_n} \end{pmatrix} : x_j \in \mathbb{R},\ \sum_j x_j = 0 \right\}, \tag{11.3.7}$$

leading to the identification

$$\mathfrak{t} = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_1 + \cdots + x_n = 0\}. \tag{11.3.8}$$

Then the set Δ of roots of $SU(n)$ is given by

$$\Delta = \{\omega_{jk} : j \neq k, 1 \leq j, k \leq n\}, \tag{11.3.9}$$

where $\omega_{jk} \in \mathfrak{t}'$ is given by

$$\omega_{jk}(x) = x_j - x_k, \quad x \in \mathfrak{t}. \tag{11.3.10}$$

See (4.2.14).

Such results, for $n = 3$, actually yield half the roots of Aut($\mathbb{O}$), as we now explain. As seen in §11.2, we have the subgroup

$$\mathcal{G}_{e_1} = \{K \in \mathrm{Aut}(\mathbb{O}) : Ke_1 = e_1\} \approx SU(3). \tag{11.3.11}$$

Thus a maximal torus $\mathbb{T}$ of $\mathcal{G}_{e_1}$ is two-dimensional, and, by Proposition 11.2.8, this must also be a maximal torus of Aut($\mathbb{O}$). The adjoint action of $\mathfrak{t}$ on Der($\mathbb{O}$) leaves invariant the Lie algebra $\mathfrak{g}_{e_1}$ of $\mathcal{G}_{e_1}$, so, with the identification (11.3.8), we see that

$$\{\omega_{jk} : 1 \le j, k \le 3, j \ne k\}, \quad \omega_{jk}(x) = x_j - x_k, \tag{11.3.12}$$

are roots of Aut($\mathbb{O}$). This gives six roots. Since $\dim \mathrm{Aut}(\mathbb{O}) = 14$ and $\mathfrak{t}$ has dimension 2, it follows from (11.3.2)–(11.3.5) that Aut($\mathbb{O}$) has 12 roots. It remains to find the other six.

Let us abstract the setting. Let G be a compact, connected Lie group, $H \subset G$ a compact, connected subgroup, and assume that a maximal torus $\mathbb{T}$ of G is contained in H, i.e., $\mathbb{T} \subset H$. Then the adjoint action of G on $\mathfrak{g}_{\mathbb{C}}$, restricted to $\mathbb{T}$, is also the restriction to $\mathbb{T}$ of the action of H on $\mathfrak{g}_{\mathbb{C}}$, obtained by restricting Ad from G to H. This latter is a unitary representation of H on $\mathfrak{g}_{\mathbb{C}}$, which we will denote by π. Thus the roots of G coincide with the weights of π.

We recall the definition of weights. Let H be as above, with maximal torus $\mathbb{T}$, whose Lie algebra is $\mathfrak{t}$, and let π be a unitary representation of H on a finite-dimensional complex inner-product space V. Then there is an orthogonal decomposition

$$V = \bigoplus_{\lambda} V_\lambda, \tag{11.3.13}$$

where, for $\lambda \in \mathfrak{t}'$,

$$V_\lambda = \{v \in V : d\pi(x)v = i\lambda(x)v, \ \forall\, x \in \mathfrak{t}\}. \tag{11.3.14}$$

If $V_\lambda \ne 0$, we call λ a weight, and any nonzero $v \in V_\lambda$ a weight vector. Generally, if π is a representation of H on V, we define the *contragredient* representation $\overline{\pi}$ of H on V' by $\overline{\pi}(g) = \pi(g^{-1})^t$. It is readily verified that $\lambda \in \mathfrak{t}'$ is a weight of π if and only if $-\lambda$ is a weight of $\overline{\pi}$.

To take an example, let $H = SU(n)$, with maximal torus given by (11.3.7) and $\mathfrak{t}$ as in (11.3.8), and let π_0 be the standard representation of $SU(n)$ on $\mathbb{C}^n$. Then the weights of π_0 are

$$\{\lambda_j : 1 \le j \le n\}, \quad \lambda_j(x) = x_j, \tag{11.3.15}$$

with associated weight spaces $V_{\lambda_j} = \mathrm{Span}\{e_j\}$, where $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{C}^n$. The weights of the contragredient representation $\overline{\pi}_0$

of $SU(n)$ on $\mathbb{C}^n$ are given by

$$\{-\lambda_j : 1 \le j \le n\}. \tag{11.3.16}$$

We return to the situation introduced three paragraphs above, with $\mathbb{T} \subset H \subset G$, and π the restriction to H of the adjoint representation of G on $\mathfrak{g}$ (and on its complexification $\mathfrak{g}_{\mathbb{C}}$). Taking an Ad-invariant inner product on $\mathfrak{g}$, we can write

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp, \tag{11.3.17}$$

and both pieces are invariant under π, say

$$\pi = \pi_{\mathfrak{h}} \oplus \pi_1. \tag{11.3.18}$$

Of course, $\pi_{\mathfrak{h}}$ is simply the adjoint action of H on $\mathfrak{h}$. We need to analyze π_1.

To do this, it is convenient to look at the homogeneous space $M = G/H$, on which G acts transitively. Then H is the subgroup of elements of G that fix the point $p = eH \in M$. This gives rise to an action of H on T_pM, i.e., a real representation ρ of H on T_pM. Furthermore, we have natural equivalences

$$\mathfrak{h}^\perp \approx T_pM, \quad \pi_1 \approx \rho. \tag{11.3.19}$$

We now apply this to

$$G = \mathrm{Aut}(\mathbb{O}), \quad H = \mathcal{G}_{e_1}, \quad M = S \subset \mathrm{Im}(\mathbb{O}), \quad p = e_1. \tag{11.3.20}$$

Then, as seen in §11.2, T_pS carries a complex structure, with respect to which, via the isomorphism $\mathcal{G}_{e_1} \approx SU(3)$ set up in §11.2, ρ becomes the standard representation π_0 of $SU(3)$ on $\mathbb{C}^3$.

However, we need to regard ρ as a real representation on T_pS, and then complexify this vector space. When this is done, the resulting representation on $(\mathfrak{h}^\perp)_{\mathbb{C}}$ is seen to be

$$\pi_0 \oplus \overline{\pi}_0, \tag{11.3.21}$$

with weights

$$\{\lambda_j, -\lambda_j : 1 \le j \le 3\}, \quad \lambda_j(x) = x_j. \tag{11.3.22}$$

We have the following conclusion.

Proposition 11.3.3. *The roots of* Aut($\mathbb{O}$) *are the linear functionals on*

$$\mathfrak{t} = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\} \tag{11.3.23}$$

given by (11.3.12) *and* (11.3.22).

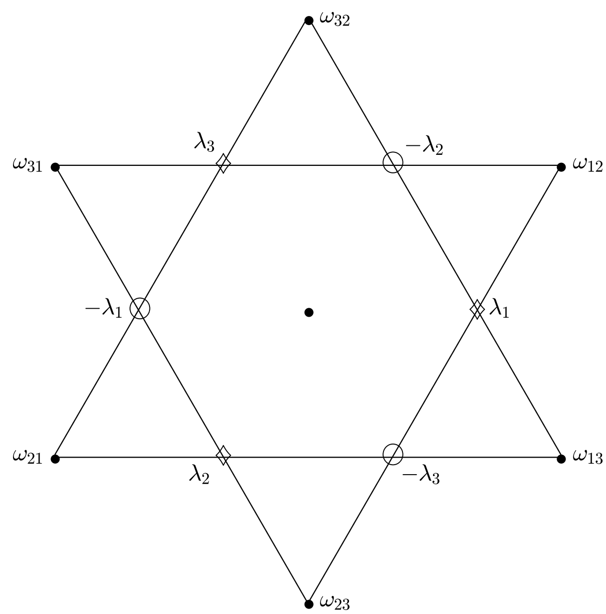

Figure 11.3.1. Root system of Aut($\mathbb{O}$) $\approx G_2$

The root system of Aut($\mathbb{O}$) $\approx G_2$ is depicted in Figure 11.3.1 Results on

$$\mathfrak{g}_0 = \mathfrak{t}_{\mathbb{C}} \oplus \bigoplus_{j \neq k} \mathfrak{g}_{\omega_{jk}}, \quad \mathfrak{g}_1 = \bigoplus_{j} \mathfrak{g}_{\lambda_j}, \quad \mathfrak{g}_{-1} = \bigoplus_{j} \mathfrak{g}_{-\lambda_j} \tag{11.3.24}$$

are discussed in the next section.

We want to investigate the Weyl group of Aut($\mathbb{O}$). Generally, if G is a compact, connected Lie group with maximal torus $\mathbb{T}$, the Weyl group of G is

$$W(G) = N(\mathbb{T})/\mathbb{T}, \quad N(\mathbb{T}) = \{g \in G : g^{-1}\mathbb{T}g = \mathbb{T}\}. \tag{11.3.25}$$

We define the representation $\mathcal{W}$ of $N(\mathbb{T})$ on $\mathfrak{t}$ by

$$\mathcal{W}(g) = \mathrm{Ad}(g)\Big|_{\mathfrak{t}}, \quad \text{for} \quad g \in N(\mathbb{T}), \tag{11.3.26}$$

and denote by $\overline{\mathcal{W}}$ the contragredient representation on $\mathfrak{t}'$ (and its complexification). Of course, these two representations are equivalent via the isomorphism $\mathfrak{t} \approx \mathfrak{t}'$ induced by the Ad-invariant inner product we use on $\mathfrak{g}$.

A key example is

$$W(SU(n)) \approx S_n, \tag{11.3.27}$$

the symmetric group on n symbols, which arises as follows (if n is odd). For $\sigma \in S_n$, the permutation matrix $E_\sigma \in U(n)$, defined on the standard basis $\{u_1, \dots, u_n\}$ of $\mathbb{C}^n$ by $E_\sigma u_k = u_{\sigma(k)}$, has the property that

$$C(E_\sigma) : \mathbb{T} \longrightarrow \mathbb{T}, \quad C(E_\sigma)V = E_\sigma^{-1} V E_\sigma, \tag{11.3.28}$$

with $\mathbb{T}$ as in (11.3.7). Since $\det E_\sigma = \operatorname{sgn} \sigma$, we need to alter (11.3.28) to get an element of $N(\mathbb{T}) \subset SU(n)$. For n odd, we can just replace E_σ in (11.3.28) by

$$\widetilde{E}_\sigma = (\operatorname{sgn} \sigma) E_\sigma. \tag{11.3.29}$$

For n even, see (6.3.35)–(6.3.37). Of immediate interest here is the case $n = 3$. Note that $\sigma \mapsto \widetilde{E}_\sigma$ gives a group homomorphism

$$S_3 \longrightarrow N(\mathbb{T}) \subset SU(3), \tag{11.3.30}$$

whose composition with $\widetilde{E}_\sigma \mapsto C(\widetilde{E}_\sigma) : \mathbb{T} \to \mathbb{T}$ yields an isomorphism of S_3 with the image of $W(SU(3))$ under the map $\overline{\mathcal{W}}$. In connection with these facts, we mention the following general results regarding $W(G)$, for an arbitrary compact, connected, semisimple Lie group G. For details, see §6.3 of this text, and Chapter 8 of [**34**].

Proposition 11.3.4. *Let π be a unitary representation of G on V, with weight space decomposition $V = \oplus V_\lambda$. Then*

$$g \in N(\mathbb{T}) \Longrightarrow \pi(g) : V_\lambda \to V_{\overline{\mathcal{W}}(g)\lambda}. \tag{11.3.31}$$

Proposition 11.3.5. *If $g \in G$ and $g^{-1}ug = u$ for each $u \in \mathbb{T}$, then $g \in \mathbb{T}$. Hence if $g \in N(\mathbb{T})$ and $\mathcal{W}(g) = I$ on $\mathfrak{t}$, then $g \in \mathbb{T}$. Consequently, we can identify $W(G)$ with its image under $\mathcal{W}$ in $G\ell(\mathfrak{t})$, and therefore also with its image under $\overline{\mathcal{W}}$ in $G\ell(\mathfrak{t}')$.*

Proposition 11.3.6. *The image of $W(G)$ under $\overline{\mathcal{W}}$ in $G\ell(\mathfrak{t}')$ is generated by the set of reflections S_α across hyperplanes in $\mathfrak{t}'$ orthogonal to α, as α runs over the set of roots of G.*

It is straightforward to verify these results for $G = SU(3)$. Note that, under $\overline{\mathcal{W}}$, the Weyl group $W(SU(3))$ acts transitively on each of the sets

$$\{\omega_{jk} : j \neq k, 1 \le j, k \le 3\}, \quad \{\lambda_j : 1 \le j \le 3\}, \quad \{-\lambda_j : 1 \le j \le 3\}, \tag{11.3.32}$$

defined as in (11.3.12), (11.3.15), and (11.3.16). The first set is the set of roots for $SU(3)$, and the last two sets are, respectively, the sets of weights for π_0 and $\overline{\pi}_0$.

Composing the map $\sigma \mapsto \widetilde{E}_\sigma$ in (11.3.30) with the inclusion $SU(3) \approx \mathcal{G}_{e_1} \subset \operatorname{Aut}(\mathbb{O})$ yields the injective group homomorphism

$$W(SU(3)) \longrightarrow W(\operatorname{Aut}(\mathbb{O})). \tag{11.3.33}$$

However, $W(\text{Aut}(\mathbb{O}))$ is bigger than $W(SU(3))$. By Proposition 11.3.6, the image under $\overline{\mathcal{W}}$ of $W(SU(3))$ is generated by the reflections in $\mathfrak{t}'$ across lines orthogonal to ω_{12}, ω_{23}, and ω_{31}, respectively. The image under $\overline{\mathcal{W}}$ of $W(\text{Aut}(\mathbb{O})$ is generated by these 3 reflections plus 3 more: reflections in $\mathfrak{t}'$ across lines orthogonal to λ_1, λ_2, and λ_3, respectively. In particular, the image under $\overline{\mathcal{W}}$ of $W(\text{Aut}(\mathbb{O}))$ acts transitively on each of the sets

$$\{\omega_{jk} : j \neq k, 1 \leq j, k \leq 3\}, \quad \{\lambda_j, -\lambda_j : 1 \leq j \leq 3\}, \tag{11.3.34}$$

which together give all the roots of Aut($\mathbb{O}$). We see that $W(\text{Aut}(\mathbb{O}))$ is isomorphic to the group of isometries of a regular hexagon.

11.4. More on the Lie algebra of $\mathrm{Aut}(\mathbb{O})$

As seen in §§11.2–11.3, the Lie algebra $\mathrm{Der}(\mathbb{O})$ of $\mathrm{Aut}(\mathbb{O})$ can be written as a vector space sum

$$\mathrm{Der}(\mathbb{O}) = \mathfrak{su}(3) \oplus V, \tag{11.4.1}$$

where $\mathfrak{su}(3)$ is the Lie algebra of $\mathcal{G}_{e_1} \approx SU(3)$ and V, the orthogonal complement of $\mathfrak{su}(3)$, is isomorphic to $T_{e_1}S$, a vector space of $\mathbb{R}$-dimension 6, with a complex structure J, so $(V, J) \approx \mathbb{C}^3$, and the natural action ρ of $\mathcal{G}_{e_1}$ on V is equivalent to the standard action of $SU(3)$ on $\mathbb{C}^3$. Thus an element of $\mathrm{Der}(\mathbb{O})$ can be represented as a pair (X, v), with $X \in \mathfrak{su}(3),\ v \in V$. If also $(Y, w) \in \mathrm{Der}(\mathbb{O})$, we want to look at the Lie bracket

$$[(X, v), (Y, w)] = [X, Y] + [X, w] + [v, Y] + [v, w]. \tag{11.4.2}$$

Of course, $[X, Y]$ is the standard bracket on $\mathfrak{su}(3)$. Meanwhile, by (5.18)–(5.19),

$$[X, w] = d\rho(X)w \in V, \tag{11.4.3}$$

and similarly $[v, Y] = -[Y, v] = -d\rho(Y)v$.

It remains to examine $[v, w]$, which will typically have a component in $\mathfrak{su}(3)$ and a component in V. The component in $\mathfrak{su}(3)$ is specified by

$$\begin{aligned} \langle X, [v, w]\rangle &= \langle X, \mathrm{ad}(v), w\rangle \\ &= -\langle \mathrm{ad}(v)X, w\rangle \\ &= \langle d\rho(X)v, w\rangle. \end{aligned} \tag{11.4.4}$$

For further analysis of $[v, w]$, it is convenient to bring in the complexification

$$V_{\mathbb{C}} = V_1 \oplus V_{-1}, \tag{11.4.5}$$

where

$$V_\mu = \{v \in V_{\mathbb{C}} : Jv = \mu i v\}, \quad \mu = \pm 1. \tag{11.4.6}$$

Since $d\rho(X)$ commutes with J, we have, for $v, w \in V_{\mathbb{C}}$,

$$\begin{aligned} \langle X, [Jv, w]\rangle &= \langle d\rho(X)Jv, w\rangle \\ &= \langle J d\rho(X)v, w\rangle \\ &= -\langle d\rho(X)v, Jw\rangle \\ &= -\langle X, [v, Jw]\rangle, \end{aligned} \tag{11.4.7}$$

and hence

$$\begin{aligned} \langle X, [Jv, Jw]\rangle &= -\langle X, [v, J^2 w]\rangle \\ &= \langle X, [v, w]\rangle. \end{aligned} \tag{11.4.8}$$

Meanwhile, for each $\mu = \pm 1$,

$$v, w \in V_\mu \Longrightarrow [Jv, Jw] = -[v, w], \tag{11.4.9}$$

so

$$v, w \in V_\mu \Longrightarrow [v, w] \perp X, \quad \forall\, X \in \mathfrak{su}(3). \tag{11.4.10}$$

More precisely, we can show that

$$v, w \in V_\mu \Longrightarrow [v, w] \in V_{-\mu}. \tag{11.4.11}$$

This can be seen from the root space decomposition, established in §11.3. With $\mathfrak{g}_{\mathbb{C}}$ denoting the complexification of Der($\mathbb{O}$), we have

$$\begin{aligned} &\mathfrak{g}_{\mathbb{C}} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_{-1}, \\ &\mathfrak{g}_0 = \mathfrak{su}(3)_{\mathbb{C}}, \quad \mathfrak{g}_1 = V_1, \quad \mathfrak{g}_{-1} = V_{-1}, \end{aligned} \tag{11.4.12}$$

and

$$\mathfrak{g}_0 = \mathfrak{t}_{\mathbb{C}} \oplus \bigoplus_{j \neq k} \mathfrak{g}_{\omega_{jk}}, \quad \mathfrak{g}_1 = \bigoplus_j \mathfrak{g}_{\lambda_j}, \quad \mathfrak{g}_{-1} = \bigoplus_j \mathfrak{g}_{-\lambda_j}, \tag{11.4.13}$$

with $\{\omega_{jk}\}$ and $\{\pm\lambda_j\}$ as in (11.3.10), (11.3.15). It follows from (11.3.4) that

$$[\mathfrak{g}_j, \mathfrak{g}_k] \subset \mathfrak{g}_\ell, \quad \ell = j + k \bmod 3. \tag{11.4.14}$$

In particular,

$$[\mathfrak{g}_1, \mathfrak{g}_{-1}] \subset \mathfrak{g}_0, \tag{11.4.15}$$

so this bracket action is completely determined by (11.4.4). It remains to analyze

$$[\mathfrak{g}_1, \mathfrak{g}_1] \to \mathfrak{g}_{-1}, \quad \text{and} \quad [\mathfrak{g}_{-1}, \mathfrak{g}_{-1}] \to \mathfrak{g}_1, \tag{11.4.16}$$

or equivalently

$$[V_1, V_1] \to V_{-1}, \quad [V_{-1}, V_{-1}] \to V_1, \tag{11.4.17}$$

with $V_{\pm 1}$ as in (11.4.5)–(11.4.6). The following observation is useful.

Proposition 11.4.1. *If (V, J) is a vector space with complex structure J, equipped with a Hermitian inner product $(\ ,\ )$, and $V_{\mathbb{C}} = V_1 \oplus V_{-1}$, as in* (11.4.5)–(11.4.6)*, then there are natural $\mathbb{C}$-linear isomorphisms*

$$V_1' \approx V_{-1} \quad \textit{and} \quad V_{-1}' \approx V_1. \tag{11.4.18}$$

Proof. The inner product $(\ ,\ )$ on V extends to a $\mathbb{C}$-bilinear form on $V_{\mathbb{C}}$. If $u - iJu \in V_1$ and $v + iJv \in V_{-1}$ (with $u, v \in V$), then

$$\begin{aligned} (u - iJu, v + iJv) &= (u, v) - i(Ju, v) + i(u, Jv) + (Ju, Jv) \\ &= 2(u, v), \end{aligned} \tag{11.4.19}$$

so the left side yields a $\mathbb{C}$-linear dual pairing of V_1 and V_{-1}. Note that $(i(u - iJu), v + iJv) = 2(Ju, v) = 2i(u, v)$ and $(u - iJu, i(v + iJv)) = -2(u, Jv) = 2i(u, v)$. $\square$

It follows that the bilinear maps in (11.4.17) yield tri-linear maps

$$\varphi : V_1 \times V_1 \times V_1 \to \mathbb{C}, \quad \psi : V_{-1} \times V_{-1} \times V_{-1} \to \mathbb{C}, \tag{11.4.20}$$

via

$$\varphi(u, v, w) = ([u, v], w), \quad u, v, w \in V_1, \tag{11.4.21}$$

and analogously for ψ. Note that

$$\begin{aligned} ([u, v], w) &= (\operatorname{ad} u(v), w) \\ &= -(v, \operatorname{ad} u(w)) \\ &= -(v, [u, w]) \\ &= -([u, w], v), \end{aligned}$$

so φ is anti-symmetric in its arguments. On the other hand,

$$\dim_{\mathbb{C}} V_1 = 3 \Longrightarrow \Lambda^3_{\mathbb{C}} V_1 \approx \mathbb{C}, \tag{11.4.22}$$

so φ is uniquely determined, up to a scalar multiple, by the anti-symmetry property. Let us note that φ in (11.4.20) is not zero, i.e., the bracket $[V_1, V_1] \hookrightarrow V_{-1}$ is not identically zero. In fact, for example, $[\mathfrak{g}_{\lambda_1}, \mathfrak{g}_{\lambda_3}]$ has nonzero image in $\mathfrak{g}_{-\lambda_2}$ (cf. Proposition 6.1.6).

Appendix A

Background in advanced calculus and ODE

This appendix provides some background in topics of advanced calculus and ODE that come up in the text. More leisurely developments of this material can be found in Chapters 2–3 of the advanced calculus text [**40**], in Chapter 4 of the ODE text [**44**], and in Chapter 1 (Basic theory of ODE and vector fields) of the text [**39**].

Section A.1 proves the submersion mapping theorem, which came up in §1.1 as a tool to show that various matrix groups are in fact smooth submanifolds of the linear spaces $M(n, \mathbb{R})$ and $M(n, \mathbb{C})$. This appendix derives the needed result from the inverse function theorem, whose proof is given in [**40**]. Another application of the inverse function theorem treated in §A.1 involves holomorphic mappings with invertible derivatives. We obtain a holomorphic inverse function theorem, of use in the treatment of the complexification of a Lie group, in Chapter 6.

Section A.2 discusses metric tensors on smooth manifolds, and the volume elements they induce, which allow one to define an integral. Section A.3 gives basic material on differential forms and how to integrate them.

Section A.4 discusses how a vector field X generates a flow. Section A.5 discusses how such a flow acts on another vector field, Y, how the derivative of such an action yields a Lie derivative $\mathcal{L}_X Y$, and how this is equal to the Lie bracket $[X, Y]$. Section A.6 uses this interplay between the Lie derivative and the Lie bracket to establish Frobenius' theorem. This result has the important consequence that if $\mathfrak{g}$ is the Lie algebra of G, then a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is the Lie algebra of a subgroup of G.

Section A.7 derives a formula for the variation of a flow as the vector field generating the flow is varied.

Section A.8 has material on the Laplace-Beltrami operator on a Riemannian manifold. One key result is that this operator is invariant under the action of isometries.

A.1. The inverse function theorem and submersion mapping theorem

Let V and W be finite dimensional real vector spaces, $\Omega \subset V$ open, and $F : \Omega \to W$. We recall that F is said to be differentiable at $x \in \Omega$ if and only if there exists a linear map $L : V \to W$ such that, for small $y \in V$,

$$F(x+y) = F(x) + Ly + r(x,y), \quad |r(x,y)| = o(|y|). \tag{A.1.1}$$

Then we set $DF(x) = L$. If F is differentiable at each $x \in \Omega$, we have $DF : \Omega \to \mathcal{L}(V,W)$. If DF is continuous, we say F is C^1. Then we can consider differentiability of DF, etc., and naturally arrive at the concept of $F \in C^k$, in the standard fashion. The following is the Inverse Function Theorem.

Theorem A.1.1. *Assume* $\dim V = \dim W$. *Let $F : \Omega \to W$ be C^k $(k \geq 1)$. Take $x_0 \in \Omega$ and assume $DF(x_0) : V \to W$ is an isomorphism. Then there exists a neighborhood U of x_0 and a neighborhood $\mathcal{O}$ of $y_0 = F(x_0)$ such that $F : U \to \mathcal{O}$ is bijective, and its inverse $F^{-1} : \mathcal{O} \to U$ is C^k.*

We assume this is part of the reader's background. Proofs can be found in Chapter 2 of [**40**], Chapter 1 of [**39**], and Appendix B of [**41**].

Our first application of this result is to the following Submersion Mapping Theorem, of use in §1.1. In this case, $\dim V \geq \dim W$.

Theorem A.1.2. *Let V and W be finite dimensional vector spaces, and $F : V \to W$ a C^k map, $k \geq 1$. Fix $p \in W$, and consider*

$$S = \{x \in V : F(x) = p\}. \tag{A.1.2}$$

Assume that, for each $x \in S$, $DF(x) : V \to W$ is surjective. Then S is a C^k submanifold of V. Furthermore, for each $x \in S$,

$$T_x S = \ker DF(x). \tag{A.1.3}$$

Proof. Given $q \in S$, set $K_q = \ker DF(q)$ and define

$$G_q : V \longrightarrow W \oplus K_q, \quad G_q(x) = (F(x), P_q(x-q)), \tag{A.1.4}$$

where P_q is a projection of V onto K_q. Note that

$$G_q(q) = (F(q), 0) = (p, 0). \tag{A.1.5}$$

Also

$$DG_q(x) = (DF(x), P_q), \quad x \in V. \tag{A.1.6}$$

We claim that

$$DG_q(q) = (DF(q), P_q) : V \to W \oplus K_q \text{ is an isomorphism.} \tag{A.1.7}$$

This is a special case of the following general observation. ☐

Lemma A.1.3. *If $A : V \to W$ is a surjective linear map and P is a projection of V onto* $\ker A$*, then*

$$(A, P) : V \longrightarrow W \oplus \ker A \quad \textit{is an isomorphism.} \tag{A.1.8}$$

We postpone the proof of this lemma and proceed with the proof of Theorem A.1.2. Having (A.1.7), we can apply the Inverse Function Theorem to obtain a neighborhood U of q in V and a neighborhood $\mathcal{O}$ of $(p, 0)$ in $W \oplus K_q$ such that $G_q : U \to \mathcal{O}$ is bijective, with C^k inverse

$$G_q^{-1} : \mathcal{O} \longrightarrow U, \quad G_q^{-1}(p, 0) = q. \tag{A.1.9}$$

By (A.1.4), given $x \in U$,

$$x \in S \Longleftrightarrow G_q(x) = (p, v), \text{ for some } v \in K_q. \tag{A.1.10}$$

Hence $S \cap U$ is the image under the C^k diffeomorphism G_q^{-1} of $\mathcal{O} \cap \{(p, v) : v \in K_q\}$. Hence S is smooth of class C^k and $\dim T_q S = \dim K_q$. It follows from the chain rule that $T_q S \subset K_q$, so the dimension count yields $T_q S = K_q$. This proves Theorem A.1.2.

It remains to prove Lemma A.1.3. Indeed, given that $A : V \to W$ is surjective, the fundamental theorem of linear algebra implies $\dim V = \dim(W \oplus \ker A)$, and it is clear that (A, P) in (A.1.8) is injective, so the isomorphism property follows.

Remark. In case $V = \mathbb{R}^n$ and $W = \mathbb{R}$, $DF(x)$ is typically denoted $\nabla F(x)$, the hypothesis on $DF(x)$ becomes $\nabla F(x) \neq 0$, and (A.1.3) is equivalent to the assertion that $\dim S = n - 1$ and, for $x \in S$,

$$\nabla F(x) \perp T_x S.$$

This result (at least for $n = 2, 3$) appears in standard multivariable calculus courses.

We next treat an extension of Theorem A.1.1 to the setting of holomorphic maps. Let $\Omega \subset \mathbb{C}^n$ be an open set, $F : \Omega \to \mathbb{C}^n$ a C^1 map. We say F is holomorphic on Ω if it satisfies the Cauchy-Riemann equations

$$\frac{\partial F}{\partial x_j} = \frac{1}{i} \frac{\partial F}{\partial y_j}, \tag{A.1.11}$$

where $(x, y) \in \Omega$, $x, y \in \mathbb{R}^n$. Consequently, if $F = u + iv$, u and v taking values in $\mathbb{R}^n$, we have

$$D_x u = D_y v, \quad D_y u = -D_x v. \tag{A.1.12}$$

Another way to phrase this is to write $F = \binom{u}{v}$, so

$$DF(x,y) = \begin{pmatrix} D_x u & D_y u \\ D_x v & D_y v \end{pmatrix}, \tag{A.1.13}$$

and the Cauchy-Riemann equations say

$$DF(x,y) = \begin{pmatrix} D_y v & -D_x v \\ -D_y u & D_x u \end{pmatrix}. \tag{A.1.14}$$

Now if we set

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \in M(2n, \mathbb{R}), \quad J^2 = \begin{pmatrix} -I & 0 \\ 0 & -I \end{pmatrix}, \tag{A.1.15}$$

we see that the Cauchy-Riemann equations hold (i.e., (A.1.13) is equal to (A.1.14)) if and only if

$$DF(x,y)J = J\,DF(x,y). \tag{A.1.16}$$

Given this, the following holomorphic inverse function theorem is a straightforward consequence of Theorem A.1.1.

Theorem A.1.4. *Let $\Omega \subset \mathbb{C}^n$ be open, $F : \Omega \to \mathbb{C}^n$ be holomorphic. Assume $p \in \Omega$, $q = F(p)$, and $DF(p)$ invertible. Then there are neighborhoods U of p and $\mathcal{O}$ of q such that $F : U \to \mathcal{O}$ is bijective, with inverse $G = F^{-1} : \mathcal{O} \to U$. Furthermore, G is holomorphic.*

Proof. We need only show that G is holomorphic. Say $q_0 \in \mathcal{O}$, $G(q_0) = p_0 \in U$. Then

$$DG(q_0) = DF(p_0)^{-1}, \tag{A.1.17}$$

and it suffices to note that

$$DF(p_0) \text{ commutes with } J \Longrightarrow DG(q_0) \text{ commutes with } J. \tag{A.1.18}$$

$\square$

A.2. Metric tensors and volume elements

Let M be a C^1 manifold of dimension n. A continuous metric tensor on M gives a continuous inner product on tangent vectors to M. In a local coordinate system $(x_1, \dots, x_n)$, identifying an open subset of M with an open set $\mathcal{O} \subset \mathbb{R}^n$, the metric tensor is given by a positive definite $n \times n$ matrix $G(x) = (g_{jk}(x))$, and the inner product of vectors U and V is given by

$$\langle U, V \rangle = U \cdot G(x) V = \sum_{j,k} g_{jk}(x) u_j(x) v_k(x), \tag{A.2.1}$$

where $U = \sum_j u_j(x) e_j$, $V = \sum_k v_k(x) e_k$, and $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{R}^n$. If we change coordinates by a C^1 diffeomorphism $F : \mathcal{O} \to \Omega$, the metric tensor $H(y) = (h_{jk}(y))$ in the coordinate system $y = F(x)$ is related to $G(x)$ by

$$DF(x)U \cdot H(y) DF(x) V = \langle U, V \rangle = U \cdot G(x) V, \tag{A.2.2}$$

at $y = F(x)$, i.e.,

$$G(x) = DF(x)^t H(y) DF(x), \tag{A.2.3}$$

or

$$g_{jk}(x) = \sum_{i,\ell} \frac{\partial F_i}{\partial x_j} \frac{\partial F_\ell}{\partial x_k} h_{i\ell}(y). \tag{A.2.4}$$

Now, for an integrable function u supported on a coordinate patch, the integral is given by

$$\int u \, dV = \int u(x) \sqrt{g} \, dx, \quad g(x) = \det G(x). \tag{A.2.5}$$

To see that (A.2.5) is well defined, note that under the change of coordinates $y = F(x)$ we have by (A.2.3) that $\det G(x) = (\det DF(x))^2 \det H(y)$. Hence

$$\sqrt{h} = |\det DF|^{-1} \sqrt{g}, \quad h = \det H, \tag{A.2.6}$$

so, by the standard change of variable formula for the integral,

$$\begin{aligned} \int u(y) \sqrt{h} \, dy &= \int u(F(x)) |\det DF|^{-1} \sqrt{g} |\det DF| \, dx \\ &= \int u(F(x)) \sqrt{g} \, dx. \end{aligned} \tag{A.2.7}$$

More generally, $\int_M u \, dV$ is defined by writing u as a sum of terms supported on coordinate charts. We see that we have a well defined integral over M, determined by the metric tensor.

In case M is an n-dimensional submanifold of $\mathbb{R}^m$ and a local coordinate chart arises via a C^1 map $\varphi : \mathcal{O} \to \mathbb{R}^m$, the metric tensor induced on $\mathcal{O}$ is given by

$$G(x) = D\varphi(x)^t D\varphi(x), \tag{A.2.8}$$

i.e.,

$$g_{jk}(x) = \sum_{\ell=1}^{m} \frac{\partial \varphi_\ell}{\partial x_j} \frac{\partial \varphi_\ell}{\partial x_k} = \frac{\partial \varphi}{\partial x_j} \cdot \frac{\partial \varphi}{\partial x_k}, \tag{A.2.9}$$

using the dot product on $\mathbb{R}^m$.

A.3. Integration of differential forms

The calculus of differential forms provides a convenient setting for integration on manifolds, as we will explain in this appendix, due to the efficient way it keeps track of changes of variables.

A k-form β on an open set $\mathcal{O} \subset \mathbb{R}^n$ has the form

$$\beta = \sum_j b_j(x)\, dx_{j_1} \wedge \cdots \wedge dx_{j_k}. \tag{A.3.1}$$

Here $j = (j_1, \dots, j_k)$ is a k-multi-index. We write $\beta \in \Lambda^k(\mathcal{O})$. The wedge product used in (A.3.1) has the anti-commutative property

$$dx_\ell \wedge dx_m = -dx_m \wedge dx_\ell, \tag{A.3.2}$$

so that if σ is a permutation of $\{1, \dots, k\}$, we have

$$dx_{j_1} \wedge \cdots \wedge dx_{j_k} = (\operatorname{sgn} \sigma)\, dx_{j_{\sigma(1)}} \wedge \cdots \wedge dx_{j_{\sigma(k)}}. \tag{A.3.3}$$

In particular, an n-form α on $\mathcal{O} \subset \mathbb{R}^n$ can be written

$$\alpha = A(x)\, dx_1 \wedge \cdots \wedge dx_n. \tag{A.3.4}$$

If $A \in L^1(\mathcal{O}, dx)$, we write

$$\int_{\mathcal{O}} \alpha = \int_{\mathcal{O}} A(x)\, dx, \tag{A.3.5}$$

the right side being the usual integral on Euclidean space.

Suppose now $\Omega \subset \mathbb{R}^n$ is open and there is a C^1 diffeomorphism $F : \Omega \to \mathcal{O}$. We define the *pull-back* $F^*\beta$ of the k-form β in (A.3.1) as

$$F^*\beta = \sum_j b_j(F(x))\, (F^* dx_{j_1}) \wedge \cdots \wedge (F^* dx_{j_k}), \tag{A.3.6}$$

where

$$F^* dx_j = \sum_\ell \frac{\partial F_j}{\partial x_\ell}\, dx_\ell, \tag{A.3.7}$$

the algebraic computation in (A.3.6) being performed using the rule (A.3.3).

If $B = (b_{\ell m})$ is an $n \times n$ matrix, then, by (A.3.3) and the standard formula for the determinant,

$$\begin{aligned} &\Big(\sum_m b_{1m}\, dx_m\Big) \wedge \Big(\sum_m b_{2m}\, dx_m\Big) \wedge \cdots \wedge \Big(\sum_m b_{nm}\, dx_m\Big) \\ &\qquad = \Big(\sum_\sigma (\operatorname{sgn} \sigma)\, b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)}\Big)\, dx_1 \wedge \cdots \wedge dx_n \\ &\qquad = (\det B)\, dx_1 \wedge \cdots \wedge dx_n. \end{aligned} \tag{A.3.8}$$

Hence, if $F : \Omega \to \mathcal{O}$ is a C^1 map and α is an n-form on $\mathcal{O}$, as in (A.3.4), then

$$F^*\alpha = \det DF(x)\, A(F(x))\, dx_1 \wedge \cdots \wedge dx_n. \tag{A.3.9}$$

This formula is especially significant in light of the standard change of variable formula

$$\int\limits_{\mathcal{O}} A(x)\, dx = \int\limits_{\Omega} A(F(x))\, |\det DF(x)|\, dx, \tag{A.3.10}$$

when $F : \Omega \to \mathcal{O}$ is a C^1 diffeomorphism. The only difference between the right side of (A.3.10) and $\int_\Omega F^*\alpha$ is the absolute value sign around $\det DF(x)$. We say a C^1 map $F : \Omega \to \mathcal{O}$ is *orientation preserving* when $\det DF(x) > 0$ for all $x \in \Omega$. In such a case, (A.3.10) yields

Proposition A.3.1. *If $F : \Omega \to \mathcal{O}$ is a C^1 orientation-preserving diffeomorphism and α an integrable n-form on $\mathcal{O}$, then*

$$\int\limits_{\mathcal{O}} \alpha = \int\limits_{\Omega} F^*\alpha. \tag{A.3.11}$$

In addition to the pull-back, there are some other operations on differential forms. The wedge product of dx_ℓ's extends to a wedge product on forms as follows. If $\beta \in \Lambda^k(\mathcal{O})$ has the form (A.3.1) and if

$$\alpha = \sum_i a_i(x)\, dx_{i_1} \wedge \cdots \wedge dx_{i_\ell} \in \Lambda^\ell(\mathcal{O}), \tag{A.3.12}$$

define

$$\alpha \wedge \beta = \sum_{i,j} a_i(x) b_j(x)\, dx_{i_1} \wedge \cdots \wedge dx_{i_\ell} \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_k} \tag{A.3.13}$$

in $\Lambda^{k+\ell}(\mathcal{O})$. We retain the equivalences (A.3.3). It follows that

$$\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha. \tag{A.3.14}$$

It is also readily verified that

$$F^*(\alpha \wedge \beta) = (F^*\alpha) \wedge (F^*\beta). \tag{A.3.15}$$

Another important operator on forms is the *exterior derivative*:

$$d : \Lambda^k(\mathcal{O}) \longrightarrow \Lambda^{k+1}(\mathcal{O}), \tag{A.3.16}$$

defined as follows. If $\beta \in \Lambda^k(\mathcal{O})$ is given by (A.3.1), then

$$d\beta = \sum_{j,\ell} \frac{\partial b_j}{\partial x_\ell}\, dx_\ell \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_k}. \tag{A.3.17}$$

The antisymmetry $dx_m \wedge dx_\ell = -dx_\ell \wedge dx_m$, together with the identity $\partial^2 b_j/\partial x_\ell \partial x_m = \partial^2 b_j/\partial x_m \partial x_\ell$, implies

$$d(d\beta) = 0, \tag{A.3.18}$$

for any smooth differential form β. We also have a product rule:

$$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^j \alpha \wedge (d\beta), \quad \alpha \in \Lambda^j(\mathcal{O}),\ \beta \in \Lambda^k(\mathcal{O}). \tag{A.3.19}$$

The exterior derivative has the following important property under pull-backs:

$$F^*(d\beta) = dF^*\beta, \tag{A.3.20}$$

if $\beta \in \Lambda^k(\mathcal{O})$ and $F : \Omega \to \mathcal{O}$ is a smooth map. To see this, extending (A.3.19) to a formula for $d(\alpha \wedge \beta_1 \wedge \cdots \wedge \beta_\ell)$ and using this to apply d to $F^*\beta$, we have

(A.3.21)
$$\begin{aligned} dF^*\beta = &\sum_{j,\ell} \frac{\partial}{\partial x_\ell}\big(b_j \circ F(x)\big)\, dx_\ell \wedge \big(F^* dx_{j_1}\big) \wedge \cdots \wedge \big(F^* dx_{j_k}\big) \\ &+ \sum_{j,\nu} (\pm) b_j\big(F(x)\big)\big(F^* dx_{j_1}\big) \wedge \cdots \wedge d\big(F^* dx_{j_\nu}\big) \wedge \cdots \wedge \big(F^* dx_{j_k}\big). \end{aligned}$$

Now the definition (A.3.6)–(A.3.7) of pull-back gives directly that

$$F^* dx_i = \sum_\ell \frac{\partial F_i}{\partial x_\ell}\, dx_\ell = dF_i, \tag{A.3.22}$$

and hence $d(F^* dx_i) = ddF_i = 0$, so only the first sum in (A.3.21) contributes to $dF^*\beta$. Meanwhile,

$$F^* d\beta = \sum_{j,m} \frac{\partial b_j}{\partial x_m}\big(F(x)\big)\ (F^* dx_m) \wedge \big(F^* dx_{j_1}\big) \wedge \cdots \wedge \big(F^* dx_{j_k}\big), \tag{A.3.23}$$

so (A.3.20) follows from the identity

$$\sum_\ell \frac{\partial}{\partial x_\ell}\big(b_j \circ F(x)\big)\, dx_\ell = \sum_m \frac{\partial b_j}{\partial x_m}\big(F(x)\big)\, F^* dx_m,$$

which in turn follows from the chain rule.

Here is another important consequence of the chain rule. Suppose $F : \Omega \to \mathcal{O}$ and $\psi : \mathcal{O} \to U$ are smooth maps between open subsets of $\mathbb{R}^n$. We claim that for any form α of any degree,

$$\psi \circ F = \varphi \Longrightarrow \varphi^*\alpha = F^*\psi^*\alpha. \tag{A.3.24}$$

It suffices to check (A.3.24) for $\alpha = dx_j$. Then (A.3.7) gives the basic identity $\psi^*\ dx_j = \sum (\partial \psi_j/\partial x_\ell)\, dx_\ell$. Consequently,

$$F^*\psi^*\ dx_j = \sum_{\ell,m} \frac{\partial F_\ell}{\partial x_m}\frac{\partial \psi_j}{\partial x_\ell}\ dx_m, \quad \varphi^*\, dx_j = \sum_m \frac{\partial \varphi_j}{\partial x_m}\, dx_m; \tag{A.3.25}$$

but the identity of these forms follows from the chain rule:

$$D\varphi = (D\psi)(DF) \Longrightarrow \frac{\partial \varphi_j}{\partial x_m} = \sum_\ell \frac{\partial \psi_j}{\partial x_\ell} \frac{\partial F_\ell}{\partial x_m}. \tag{A.3.26}$$

One can define a k-form on an n-dimensional manifold M as follows. Say M is covered by open sets $\mathcal{O}_j$ and there are coordinate charts $F_j : \Omega_j \to \mathcal{O}_j$, with $\Omega_j \subset \mathbb{R}^n$ open. A collection of forms $\beta_j \in \Lambda^k(\Omega_j)$ is said to define a k-form on M provided the following compatibility condition holds. If $\mathcal{O}_i \cap \mathcal{O}_j \neq \emptyset$ and we consider $\Omega_{ij} = F_i^{-1}(\mathcal{O}_i \cap \mathcal{O}_j)$ and diffeomorphisms

$$\varphi_{ij} = F_j^{-1} \circ F_i : \Omega_{ij} \longrightarrow \Omega_{ji}, \tag{A.3.27}$$

we require

$$\varphi_{ij}^* \beta_j = \beta_i. \tag{A.3.28}$$

The fact that this is a consistent definition is a consequence of (A.3.24). For example, if $G : M \to R^m$ is a smooth map and γ is a k-form on $\mathbb{R}^m$, then there is a well-defined k-form $\beta = G^*\gamma$ on M, represented in such coordinate charts by $\beta_j = (G \circ F_j)^*\gamma$. Similarly, if β is a k-form on M as defined above and $G : U \to M$ is smooth, with $U \subset \mathbb{R}^m$ open, then $G^*\beta$ is a well-defined k-form on U.

We give an intrinsic definition of $\int_M \alpha$ when α is an n-form on M, provided M is *oriented*, i.e., there is a coordinate cover as above such that $\det D\varphi_{jk} > 0$. The object called an "orientation" on M can be identified as an equivalence class of nowhere vanishing n-forms on M, two such forms being equivalent if one is a multiple of another by a positive function in $C^\infty(\Omega)$. A member of this equivalence class, say ω, defines the orientation. The standard orientation on $\mathbb{R}^n$ is determined by $dx_1 \wedge \cdots \wedge dx_n$. The equivalence class of positive multiples $a(x)\omega$ is said to consist of "positive" forms. A smooth map $\psi : S \to M$ between oriented n-dimensional manifolds preserves orientation provided $\psi^*\sigma$ is positive on S whenever $\sigma \in \Lambda^n(M)$ is positive. We mention that there exist surfaces that cannot be oriented, such as the famous "Möbius strip."

We define the integral of an n-form over an oriented n-dimensional manifold as follows. First, if α is an n-form supported on an open set $\mathcal{O} \subset \mathbb{R}^n$, given by (A.3.4), then we define $\int_{\mathcal{O}} \alpha$ by (A.3.5).

More generally, if M is an n-dimensional manifold with an orientation, say the image of an open set $\mathcal{O} \subset \mathbb{R}^n$ by $\varphi : \mathcal{O} \to M$, carrying the natural orientation of $\mathcal{O}$, we can set

$$\int_M \alpha = \int_{\mathcal{O}} \varphi^*\alpha \tag{A.3.29}$$

for an n-form α on M. If it takes several coordinate patches to cover M, define $\int_M \alpha$ by writing α as a sum of forms, each supported on one patch.

We need to show that this definition of $\int_M \alpha$ is independent of the choice of coordinate system on M (as long as the orientation of M is respected). Thus, suppose $\varphi : \mathcal{O} \to U \subset M$ and $\psi : \Omega \to U \subset M$ are both coordinate patches, so that $F = \psi^{-1} \circ \varphi : \mathcal{O} \to \Omega$ is an orientation-preserving diffeomorphism. We need to check that, if α is an n-form on M, supported on U, then

$$\int\limits_{\mathcal{O}} \varphi^* \alpha = \int\limits_{\Omega} \psi^* \alpha. \tag{A.3.30}$$

To establish this, we use (A.3.24). This implies that the left side of (A.3.30) is equal to

$$\int\limits_{\mathcal{O}} F^*(\psi^* \alpha), \tag{A.3.31}$$

which is equal to the right side of (A.3.30), by (A.3.11) (with slightly altered notation). Thus the integral of an n-form over an oriented n-dimensional manifold is well defined.

A.4. Flows and vector fields

Let $U \subset \mathbb{R}^n$ be open. A vector field on U is a smooth map

$$X : U \longrightarrow \mathbb{R}^n. \tag{A.4.1}$$

Consider the corresponding ODE

$$\frac{dy}{dt} = X(y), \quad y(0) = x, \tag{A.4.2}$$

with $x \in U$. A curve $y(t)$ solving (A.4.2) is called an integral curve of the vector field X. It is also called an *orbit*. For fixed t, write

$$y = y(t,x) = \mathcal{F}_X^t(x). \tag{A.4.3}$$

The locally defined $\mathcal{F}_X^t$, mapping (a subdomain of) U to U, is called the *flow* generated by the vector field X.

The vector field X defines a differential operator on scalar functions, as follows:

$$\mathcal{L}_X f(x) = \lim_{h\to 0} h^{-1}\big[f(\mathcal{F}_X^h x) - f(x)\big] = \frac{d}{dt} f(\mathcal{F}_X^t x)\big|_{t=0}. \tag{A.4.4}$$

We also use the common notation

$$\mathcal{L}_X f(x) = Xf, \tag{A.4.5}$$

that is, we apply X to f as a first order differential operator.

Note that, if we apply the chain rule to (A.4.4) and use (A.4.2), we have

$$\mathcal{L}_X f(x) = X(x) \cdot \nabla f(x) = \sum a_j(x) \frac{\partial f}{\partial x_j}, \tag{A.4.6}$$

if $X = \sum a_j(x) e_j$, with $\{e_j\}$ the standard basis of $\mathbb{R}^n$. In particular, using the notation (A.4.5), we have

$$a_j(x) = X x_j. \tag{A.4.7}$$

In the notation (A.5),

$$X = \sum a_j(x) \frac{\partial}{\partial x_j}. \tag{A.4.8}$$

We note that X is a *derivation*, i.e., a map on $C^\infty(U)$, linear over $\mathbb{R}$, satisfying

$$X(fg) = (Xf)g + f(Xg). \tag{A.4.9}$$

Conversely, any derivation on $C^\infty(U)$ defines a vector field, i.e., has the form (A.4.8), as we now show.

Proposition A.4.1. *If X is a derivation on $C^\infty(U)$, then X has the form* (A.4.8).

Proof. Set $a_j(x) = Xx_j,\ X^\# = \sum a_j(x)\partial/\partial x_j$, and $Y = X - X^\#$. Then Y is a derivation satisfying $Yx_j = 0$ for each j; we aim to show that $Yf = 0$ for all f. Note that, whenever Y is a derivation

$$1 \cdot 1 = 1 \Rightarrow Y \cdot 1 = 2Y \cdot 1 \Rightarrow Y \cdot 1 = 0,$$

i.e., Y annihilates constants. Thus in this case Y annihilates all polynomials of degree ≤ 1.

Now we show $Yf(p) = 0$ for all $p \in U$. Without loss of generality, we can suppose $p = 0$, the origin. Then we can take $b_j(x) = \int_0^1 (\partial_j f)(tx)\,dt$, and write

$$f(x) = f(0) + \sum b_j(x)x_j.$$

It immediately follows that Yf vanishes at 0, so the proposition is proved. $\square$

If U is a manifold, it is natural to regard a vector field X as a section of the tangent bundle of U. Of course, the characterization given in Proposition A.4.1 makes good invariant sense on a manifold.

A fundamental fact about vector fields is that they can be "straightened out" near points where they do not vanish. To see this, suppose a smooth vector field X is given on U (open in $\mathbb{R}^n$) such that, for a certain $p \in U,\ X(p) \neq 0$. Then near p there is a hypersurface M which is nowhere tangent to X. We can choose coordinates near p so that p is the origin and M is given by $\{x_n = 0\}$. Thus we can take a neighborhood $\mathcal{O}$ of $0 \in \mathbb{R}^{n-1}$, and define a map

$$\Phi : \mathcal{O} \times (-t_0, t_0) \longrightarrow U \tag{A.4.10}$$

by

$$\Phi(u', t) = \mathcal{F}_X^t(\psi(u')), \quad \psi(u') = (u', 0). \tag{A.4.11}$$

This is C^∞ and has surjective derivative, so by the Inverse Function Theorem is a local diffeomorphism. Note that

$$\mathcal{F}_X^s \Phi(u', t) = \mathcal{F}_X^s \mathcal{F}_X^t(\psi(u')) = \mathcal{F}_X^{t+s}(\psi(u')) = \Phi(u', t+s). \tag{A.4.12}$$

If we set $u = (u', t),\ x = \Phi(u)$, we have the following result, known as the Straightening Lemma.

Theorem A.4.2. *If X is a smooth vector field on U with $X(p) \neq 0$, then there exists a coordinate system $(u_1, \dots, u_n)$ centered at p (so $u_j(p) = 0$) with respect to which*

$$X = \frac{\partial}{\partial u_n}. \tag{A.4.13}$$

A.5. Lie brackets

If $F : \Omega \to \mathcal{O}$ is a diffeomorphism between two open domains in $\mathbb{R}^n$, or between two smooth manifolds, and Y is a vector field on $\mathcal{O}$, we define a vector field $F_{\#}Y$ on Ω so that

$$\mathcal{F}^t_{F_{\#}Y} = F^{-1} \circ \mathcal{F}^t_Y \circ F, \tag{A.5.1}$$

or equivalently, by the chain rule,

$$F_{\#}Y(x) = \bigl(DF^{-1}\bigr)\bigl(F(x)\bigr)Y\bigl(F(x)\bigr). \tag{A.5.2}$$

In particular, if $U \subset \mathbb{R}^n$ is open and X is a vector field on U, defining a flow $\mathcal{F}^t$, then for a vector field Y, $\mathcal{F}^t_{\#}Y$ is defined on most of U, for $|t|$ small, and we can define the Lie derivative:

$$\mathcal{L}_X Y = \lim_{h\to 0} h^{-1}\bigl(\mathcal{F}^h_{\#}Y - Y\bigr) = \frac{d}{dt}\mathcal{F}^t_{\#}Y\bigr|_{t=0}, \tag{A.5.3}$$

as a vector field on U.

Another natural construction is the operator-theoretic bracket:

$$[X, Y] = XY - YX, \tag{A.5.4}$$

where the vector fields X and Y are regarded as first order differential operators on $C^\infty(U)$. One verifies that (A.5.4) defines a derivation on $C^\infty(U)$, hence a vector field on U. The basic elementary fact about the Lie bracket is the following.

Theorem A.5.1. *If X and Y are smooth vector fields, then*

$$\mathcal{L}_X Y = [X, Y]. \tag{A.5.5}$$

Proof. Let us first verify the identity in the special case

$$X = \frac{\partial}{\partial x_1}, \quad Y = \sum b_j(x)\frac{\partial}{\partial x_j}.$$

Then $\mathcal{F}^t_{\#}Y = \sum b_j(x+te_1)\partial/\partial x_j$. Hence, in this case $\mathcal{L}_X Y = \sum(\partial b_j/\partial x_1)\partial/\partial x_j$, and a straightforward calculation shows this is also the formula for $[X, Y]$, in this case.

Now we verify (A.5.5) in general, at any point $x_0 \in U$. First, if X is nonvanishing at x_0, we can choose a local coordinate system so the example above gives the identity. By continuity, we get the identity (A.5.5) on the closure of the set of points x_0 where $X(x_0) \neq 0$. Finally, if x_0 has a neighborhood where $X = 0$, clearly $\mathcal{L}_X Y = 0$ and $[X, Y] = 0$ at x_0. This completes the proof. $\square$

Corollary A.5.2. *If X and Y are smooth vector fields on U, then*

$$\frac{d}{dt}\mathcal{F}^t_{X\#}Y = \mathcal{F}^t_{X\#}[X, Y] \tag{A.5.6}$$

for all t.

Proof. Since locally $\mathcal{F}_X^{t+s} = \mathcal{F}_X^s \mathcal{F}_X^t$, we have the same identity for $\mathcal{F}_{X\#}^{t+s}$, which yields (A.5.6) upon taking the s-derivative. In more detail, $(d/dt)\mathcal{F}_{X\#}^t Y = (d/ds)\mathcal{F}_{X\#}^{t+s} Y|_{s=0} = (d/ds)\mathcal{F}_{X\#}^t Y|_{s=0} = \mathcal{F}_{X\#}^t \mathcal{L}_X Y$, and the last step follows from (A.5.5). □

We make some further comments about cases when one can explicitly integrate a vector field X in the plane, exploiting "symmetries" that might be apparent. In fact, suppose one has in hand a vector field Y such that

$$[X, Y] = 0. \tag{A.5.7}$$

By (A.5.6), this implies $\mathcal{F}_{Y\#}^t X = X$ for all t. Suppose one has an explicit hold on the flow generated by Y, so one can produce explicit local coordinates (u, v) with respect to which

$$Y = \frac{\partial}{\partial u}. \tag{A.5.8}$$

In this coordinate system, write $X = a(u,v)\partial/\partial u + b(u,v)\partial/\partial v$. The condition (A.5.7) implies $\partial a/\partial u = 0 = \partial b/\partial u$, so in fact we have

$$X = a(v)\frac{\partial}{\partial u} + b(v)\frac{\partial}{\partial v}. \tag{A.5.9}$$

Integral curves of (A.5.9) satisfy

$$u' = a(v), \quad v' = b(v) \tag{A.5.10}$$

and can be found explicitly in terms of integrals; one has

$$\int b(v)^{-1}\, dv = t + C_1, \tag{A.5.11}$$

and then

$$u = \int a(v(t))\, dt + C_2. \tag{A.5.12}$$

More generally than (A.5.7), we can suppose that, for some constant c,

$$[X, Y] = cX, \tag{A.5.13}$$

which by (A.5.6) is the same as

$$\mathcal{F}_{Y\#}^t X = e^{-ct} X. \tag{A.5.14}$$

An example would be

$$X = f(x,y)\frac{\partial}{\partial x} + g(x,y)\frac{\partial}{\partial y}, \tag{A.5.15}$$

where f and g satisfy "homogeneity" conditions of the form

$$f(r^a x, r^b y) = r^{a-c} f(x,y), \quad g(r^a x, r^b y) = r^{b-c} g(x,y), \tag{A.5.16}$$

for $r > 0$; in such a case one can take explicitly

$$\mathcal{F}_Y^t(x, y) = (e^{at}x, e^{bt}y). \tag{A.5.17}$$

Now, if one again has (B.8) in a local coordinate system (u, v), then X must have the form

$$X = e^{cu}\Big[a(v)\frac{\partial}{\partial u} + b(v)\frac{\partial}{\partial v}\Big] \tag{A.5.18}$$

which can be explicitly integrated, since

$$u' = e^{cu}a(v),\ v' = e^{cu}b(v) \Longrightarrow \frac{du}{dv} = \frac{a(v)}{b(v)}. \tag{A.5.19}$$

The hypothesis (A.5.13) implies that the linear span (over $\mathbb{R}$) of X and Y is a two dimensional solvable Lie algebra. Sophus Lie devoted a good deal of effort to examining when one could use constructions of solvable Lie algebras of vector fields to explicitly integrate vector fields; his investigations led to his foundation of the theory of Lie groups.

A.6. Frobenius' theorem

Let $G : U \to V$ be a diffeomorphism. Recall from §A.5 the action on vector fields:

$$G_{\#}Y(x) = DG(y)^{-1}Y(y), \quad y = G(x). \tag{A.6.1}$$

As noted there, an alternative characterization of $G_{\#}Y$ is given in terms of the flow it generates. One has

$$\mathcal{F}_Y^t \circ G = G \circ \mathcal{F}_{G_{\#}Y}^t. \tag{A.6.2}$$

The proof of this is a direct consequence of the chain rule. As a special case, we have the following

Proposition A.6.1. *If $G_{\#}Y = Y$, then $\mathcal{F}_Y^t \circ G = G \circ \mathcal{F}_Y^t$.*

From this, we derive the following condition for a pair of flows to commute. Let X and Y be vector fields on U.

Proposition A.6.2. *If X and Y commute as differential operators, i.e.,*

$$[X, Y] = 0, \tag{A.6.3}$$

then locally $\mathcal{F}_X^s$ and $\mathcal{F}_Y^t$ commute, i.e., for any $p_0 \in U$, there exists $\delta > 0$ such that, for $|s|, |t| < \delta$,

$$\mathcal{F}_X^s \mathcal{F}_Y^t p_0 = \mathcal{F}_Y^t \mathcal{F}_X^s p_0. \tag{A.6.4}$$

Proof. By Proposition A.6.1, it suffices to show that $\mathcal{F}_{X\#}^s Y = Y$. Clearly this holds at $s = 0$. But by (A.5.6), we have

$$\frac{d}{ds}\mathcal{F}_{X\#}^s Y = \mathcal{F}_{X\#}^s [X, Y],$$

which vanishes if (A.6.3) holds. This finishes the proof. $\square$

We have stated that, given (A.6.3), then (A.6.4) holds locally. If the flows generated by X and Y are not complete, this can break down globally. For example, consider $X = \partial/\partial x_1$, $Y = \partial/\partial x_2$ on $\mathbb{R}^2$, which satisfy (A.6.3) and generate commuting flows. These vector fields lift to vector fields on the universal covering surface $\tilde{M}$ of $\mathbb{R}^2 \setminus (0, 0)$, which continue to satisfy (A.6.3). The flows on $\tilde{M}$ do not commute globally. This phenomenon does not arise, for example, for vector fields on a compact manifold.

We now consider when a family of vector fields has a multidimensional integral manifold. Suppose $X_1, \dots, X_k$ are smooth vector fields on U which are linearly independent at each point of a $k-$dimensional surface $\Sigma \subset U$. If each X_j is tangent to Σ at each point, Σ is said to be an integral manifold of $(X_1, \dots, X_k)$.

Proposition A.6.3. *Suppose $X_1, \dots, X_k$ are linearly independent at each point of U and $[X_j, X_\ell] = 0$ for all j, ℓ. Then, for each $x_0 \in U$, there is a k-dimensional integral manifold of $(X_1, \dots, X_k)$ containing x_0.*

Proof. We define a map $F : V \to U$, V a neighborhood of 0 in $\mathbb{R}^k$, by

$$F(t_1, \dots, t_k) = \mathcal{F}_{X_1}^{t_1} \cdots \mathcal{F}_{X_k}^{t_k} x_0. \tag{A.6.5}$$

Clearly $(\partial/\partial t_1)F = X_1(F)$. Similarly, since $\mathcal{F}_{X_j}^{t_j}$ all commute, we can put any $\mathcal{F}_{X_j}^{t_j}$ first and get $(\partial/\partial t_j)F = X_j(F)$. This shows that the image of V under F is an integral manifold containing x_0. $\square$

We now derive a more general condition guaranteeing the existence of integral submanifolds. This important result is due to Frobenius. We say $(X_1, \dots, X_k)$ is *involutive* provided that, for each j, ℓ, there are smooth $b_m^{j\ell}(x)$ such that

$$[X_j, X_\ell] = \sum_{m=1}^{k} b_m^{j\ell}(x) X_m. \tag{A.6.6}$$

The following is Frobenius' Theorem.

Theorem A.6.4. *If $(X_1, \dots, X_k)$ are C^∞ vector fields on U, linearly independent at each point, and the involutivity condition* (A.6.6) *holds, then through each x_0 there is, locally, a unique integral manifold Σ, of dimension k.*

We will give two proofs of this result. First, let us restate the conclusion as follows. There exist local coordinates $(y_1, \dots, y_n)$ centered at x_0 such that

$$\text{span } (X_1, \dots, X_k) = \text{ span } \Big(\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_k}\Big). \tag{A.6.7}$$

First proof. The result is clear for $k = 1$. We will use induction on k. So let the set of vector fields $X_1, \dots, X_{k+1}$ be linearly independent at each point and involutive. Choose a local coordinate system so that $X_{k+1} = \partial/\partial u_1$. Now let

$$Y_j = X_j - (X_j u_1)\frac{\partial}{\partial u_1} \quad \text{for} \quad 1 \le j \le k, \quad Y_{k+1} = \frac{\partial}{\partial u_1}. \tag{A.6.8}$$

Since, in $(u_1, \dots, u_n)$ coordinates, no $Y_1, \dots, Y_k$ involves $\partial/\partial u_1$, neither does any Lie bracket, so

$$[Y_j, Y_\ell] \in \text{ span } (Y_1, \dots, Y_k), \quad j, \ell \le k.$$

Thus $(Y_1, \dots, Y_k)$ is involutive. The induction hypothesis implies there exist local coordinates $(y_1, \dots, y_n)$ such that

$$\text{span } (Y_1, \dots, Y_k) = \text{ span } \Big(\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_k}\Big).$$

Now let

$$Z = Y_{k+1} - \sum_{\ell=1}^{k} (Y_{k+1} y_\ell) \frac{\partial}{\partial y_\ell} = \sum_{\ell > k} (Y_{k+1} y_\ell) \frac{\partial}{\partial y_\ell}. \tag{A.6.9}$$

Since, in the $(u_1, \dots, u_n)$ coordinates, $Y_1, \dots, Y_k$ do not involve $\partial/\partial u_1$, we have

$$[Y_{k+1}, Y_j] \in \text{ span } (Y_1, \dots, Y_k).$$

Thus $[Z, Y_j] \in$ span $(Y_1, \dots, Y_k)$ for $j \le k$, while (A.6.9) implies that $[Z, \partial/\partial y_j]$ belongs to the span of $(\partial/\partial y_{k+1}, \dots, \partial/\partial y_n)$, for $j \le k$. Thus we have

$$\Big[Z, \frac{\partial}{\partial y_j}\Big] = 0, \quad j \le k.$$

Proposition A.6.3 implies span $(\partial/\partial y_1, \dots, \partial/\partial y_k, Z)$ has an integral manifold through each point, and since this span is equal to the span of $X_1, \dots, X_{k+1}$, the first proof is complete. $\square$

Second proof. Let $X_1, \dots, X_k$ be C^∞ vector fields, linearly independent at each point, and satisfying the condition (A.6.6). Choose an $n-k$ dimensional surface $\mathcal{O} \subset U$, transverse to $X_1, \dots, X_k$. For V a neighborhood of the origin in $\mathbb{R}^k$, define $\Phi : V \times \mathcal{O} \to U$ by

$$\Phi(t_1, \dots, t_k, x) = \mathcal{F}_{X_1}^{t_1} \cdots \mathcal{F}_{X_k}^{t_k} x. \tag{A.6.10}$$

We claim that, for x fixed, the image of V in U is a k dimensional surface Σ tangent to each X_j, at each point of Σ. Note that, since $\Phi(0, \dots, t_j, \dots, 0, x) = \mathcal{F}_{X_j}^{t_j} x$, we have

$$\frac{\partial}{\partial t_j} \Phi(0, \dots, 0, x) = X_j(x), \quad x \in \mathcal{O}. \tag{A.6.11}$$

To establish the claim, it suffices to show that $\mathcal{F}_{X_j \#}^{t} X_\ell$ is a linear combination with coefficients in $C^\infty(U)$ of $X_1, \dots, X_k$. This is accomplished by the following:

Lemma A.6.5. *Suppose $[Y, X_j] = \sum_\ell \lambda_{j\ell}(x) X_\ell$, with smooth coefficients $\lambda_{j\ell}(x)$. Then $\mathcal{F}_{Y\#}^t X_j$ is a linear combination of $X_1, \dots, X_k$, with coefficients in $C^\infty(U)$.*

Proof. Denote by Λ the matrix $(\lambda_{j\ell})$ and let $\Lambda(t) = \Lambda(t, x) = (\lambda_{j\ell}(\mathcal{F}_Y^t x))$. Now let $A(t) = A(t, x)$ be the unique solution to the ODE

$$\frac{d}{dt}A(t) = \Lambda(t)A(t), \quad A(0) = I. \tag{A.6.12}$$

Write $A = (\alpha_{j\ell})$. We claim that

$$\mathcal{F}_{Y\#}^t X_j = \sum_\ell \alpha_{j\ell}(t, x)X_\ell. \tag{A.6.13}$$

This formula will prove the lemma. Indeed, we have

$$\begin{aligned}
\frac{d}{dt}(\mathcal{F}_Y^t)_\# X_j &= (\mathcal{F}_Y^t)_\#[Y, X_j] \\
&= (\mathcal{F}_Y^t)_\# \sum_\ell \lambda_{j\ell} X_\ell \\
&= \sum_\ell (\lambda_{j\ell} \circ \mathcal{F}_Y^t)(\mathcal{F}_{Y\#}^t X_\ell).
\end{aligned}$$

Uniqueness of the solution to (A.6.12) gives (A.6.13), and we are done. $\square$

This completes the second proof of Frobenius' Theorem.

A.7. Variation of flows

We want to derive a formula for the variation of a flow as the vector field generating the flow is varied. It will be technically convenient to consider first how a solution to an ODE depends on the initial conditions. Consider a nonlinear system

$$\frac{dy}{dt} = F(y), \quad y(0) = x. \tag{A.7.1}$$

Suppose $F : U \to \mathbb{R}^n$ is smooth, $U \subset \mathbb{R}^n$ open; for simplicity we assume U is convex. Say $y = y(t, x)$. We want to examine smoothness in x.

Note that *formally* differentiating (A.7.1) with respect to x suggests that $W = D_x y(t, x)$ satisfies an ODE called the *linearization* of (A.7.1):

$$\frac{dW}{dt} = DF(y)W, \quad W(0) = I. \tag{A.7.2}$$

In other words, $w(t, x) = D_x y(t, x)w_0$ satisfies

$$\frac{dw}{dt} = DF(y)w, \quad w(0) = w_0. \tag{A.7.3}$$

To justify this, we want to compare $w(t)$ and

$$z(t) = y_1(t) - y(t) = y(t, x + w_0) - y(t, x). \tag{A.7.4}$$

It would be convenient to show that z satisfies an ODE similar to (A.7.3). Indeed, $z(t)$ satisfies

$$\frac{dz}{dt} = F(y_1) - F(y) = \Phi(y_1, y)z, \quad z(0) = w_0, \tag{A.7.5}$$

where

$$\Phi(y_1, y) = \int_0^1 DF\big(\tau y_1 + (1 - \tau)y\big)\, d\tau. \tag{A.7.6}$$

If we assume

$$\|DF(u)\| \le M \text{ for } u \in U, \tag{A.7.7}$$

then the solution operator $S(t, 0)$ of the linear ODE $d/dt - B(t)$, with $B(y) = \Phi(y_1(t), y(t))$, satisfies a bound $\|S(t, 0)\| \le e^{|t|M}$ as long as $y(t), y_1(t) \in U$. Hence

$$\|y_1(t) - y(t)\| \le e^{|t|M}\|w_0\|. \tag{A.7.8}$$

This establishes that $y(t, x)$ is *Lipschitz* in x.

To continue, since $\Phi(y, y) = DF(y)$, we rewrite (D.5) as

$$\frac{dz}{dt} = \Phi(y + z, y)z = DF(y)z + R(y, z), \quad w(0) = w_0. \tag{A.7.9}$$

where

$$\in C^1(U) \Longrightarrow \|R(y, z)\| = o(\|z\|) = o(\|w_0\|). \tag{A.7.10}$$

Now comparing the ODE (A.7.9) with (A.7.3), we have

$$\frac{d}{dt}(z-w) = DF(y)(z-w) + R(y,z), \quad (z-w)(0) = 0. \tag{A.7.11}$$

Then Duhamel's principle yields

$$z(t) - w(t) = \int_0^t S(t,s) R\big(y(s), z(s)\big)\, ds, \tag{A.7.12}$$

so by the bound $\|S(t,s)\| \le e^{|t-s|M}$ and (A.7.10) we have

$$z(t) - w(t) = o(\|w_0\|). \tag{A.7.13}$$

This is precisely what is required to show that $y(t,x)$ is differentiable with respect to x, with derivative $W = D_x y(t,x)$ satisfying (A.7.2). We state our first result.

Proposition A.7.1. *If $F \in C^1(U)$, and if solutions to* (A.7.1) *exist for $t \in (-T_0, T_1)$, then for each such t, $y(t,x)$ is C^1 in x, with derivative $D_x y(t,x) = W(t,x)$ satisfying* (A.7.2).

So far we have shown that $y(t,x)$ is both Lipschitz and differentiable in x, but the continuity of $W(t,x)$ in x follows easily by comparing the ODEs of the form (A.7.2) for $W(t,x)$ and $W(t,x+w_0)$, in the spirit of the analysis of (A.7.11).

If F possesses further smoothness, we can obtain higher differentiability of $y(t,x)$ in x by the following trick. *Couple* (A.7.1) and (A.7.2), to get an ODE for (y,W) :

$$\frac{dy}{dt} = F(y), \quad \frac{dW}{dt} = DF(y)W \tag{A.7.14}$$

with initial condition

$$y(0) = x, \quad W(0) = I. \tag{A.7.15}$$

We can reiterate the argument above, getting results on $D_x(y,W)$, i.e., on $D_x^2 y(t,x)$, and continue, proving:

Proposition A.7.2. *If $F \in C^k(U)$, then $y(t,x)$ is C^k in x.*

We now tackle our stated goal: to consider dependence of the solution to a system of the form

$$\frac{dy}{dt} = F(\tau, y), \quad y(0) = x \tag{A.7.16}$$

on a parameter τ, assuming F is smooth jointly in τ, y. This result can be deduced from the previous one by the following trick: consider the ODE

$$\frac{dy}{dt} = F(z,y), \ \frac{dz}{dt} = 0; \quad y(0) = x, \ z(0) = \tau. \tag{A.7.17}$$

Thus we get smoothness of $y(t, \tau, x)$ in (τ, x). Furthermore, $v(t, \tau, x) = \partial_\tau y(t, \tau, x)$ satisfies

$$\frac{dv}{dt} = D_y F(\tau, y)v + F_\tau(\tau, y), \quad v(0, \tau, x) = 0. \tag{A.7.18}$$

A.8. The Laplace-Beltrami operator

The Laplace operator Δ on an open set $\Omega \subset \mathbb{R}^n$ is given by

$$\Delta u(x) = \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2}. \tag{A.8.1}$$

One useful characterization of this operator is that, if $u \in C^2(\Omega)$ and $v \in C_0^2(\Omega)$, then

$$\int_\Omega \Delta u(x)\, v(x)\, dx = -\int_\Omega \nabla u(x) \cdot \nabla v(x)\, dx, \tag{A.8.2}$$

an identity that follows by integration by parts. We can use this to define the Laplace-Beltrami operator Δ on a Riemannian manifold M (equipped with a metric tensor (g_{jk})) as follows. Take $u \in C^2(M)$. Then we want Δu to satisfy

$$\int_M \Delta u(x)\, v(x)\, dV = -\int_M \langle \nabla u, \nabla v\rangle\, dV, \quad \forall\, v \in C_0^2(M), \tag{A.8.3}$$

where the inner product $\langle X, Y\rangle$ of the vector fields X and Y is determined by the metric tensor. We have, in local coordinates,

$$\langle \nabla u, \nabla v\rangle = g^{jk}(x)\, \partial_k u(x)\, \partial_j v(x), \tag{A.8.4}$$

using the summation convention, and taking (g^{jk}) to be the matrix inverse of (g_{jk}). Hence, if v is supported on one coordinate chart,

$$\begin{aligned} -\int_M \langle \nabla u, \nabla v\rangle\, dV &= -\int g^{jk} \partial_k u\, \partial_j v\, g^{1/2}\, dx \\ &= \int \partial_j \Big(g^{1/2} g^{jk} \partial_k u \Big) v\, dx \\ &= \int_M g^{-1/2} \partial_j \Big(g^{1/2} g^{jk} \partial_k u \Big) v\, dV, \end{aligned} \tag{A.8.5}$$

the second identity in (A.8.5) by integration by parts. Thus the Laplace-Beltrami operator Δ is given in local coordinates by

$$\Delta u(x) = g(x)^{-1/2} \partial_j \Big(g(x)^{1/2} g^{jk}(x) \partial_k u(x) \Big). \tag{A.8.6}$$

The formula (A.8.6) has the following important implication.

Proposition A.8.1. *Assume $\varphi : M \to M$ is a smooth isometry (i.e., it preserves the metric tensor). Take*

$$\varphi^* u(x) = u(\varphi(x)). \tag{A.8.7}$$

Then, for each $u \in C^2(M)$,

$$\Delta\varphi^* u = \varphi^* \Delta u. \tag{A.8.8}$$

Applications of these concepts and results arise in the use of spherical polar coordinates. In detail, let S^{n-1} denote the unit sphere in $\mathbb{R}^n$,

$$S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}. \tag{A.8.9}$$

Spherical polar coordinates on $\mathbb{R}^n$ are defined in terms of a smooth diffeomorphism

$$R : (0,\infty) \times S^{n-1} \longrightarrow \mathbb{R}^n \setminus 0, \quad R(r,\omega) = r\omega. \tag{A.8.10}$$

Let $(h_{\ell m})$ denote the metric tensor on S^{n-1} (induced from its inclusion in $\mathbb{R}^n$), with respect to some coordinate chart $\varphi : \mathcal{O} \to U \subset S^{n-1}$. Then we have a coordinate chart

$$\Phi : (0,\infty) \times \mathcal{O} \longrightarrow \mathcal{U} \subset \mathbb{R}^n, \quad \Phi(r,y) = r\varphi(y). \tag{A.8.11}$$

Take $y_0 = r$, $y = (y_1, \dots, y_{n-1})$. In this coordinate system Φ, the Euclidean metric tensor (e_{jk}) is give by

$$\begin{aligned} e_{00} &= \partial_0\Phi \cdot \partial_0\Phi = \varphi(y) \cdot \varphi(y) = 1, \\ e_{0j} &= \partial_0\Phi \cdot \partial_j\Phi = \varphi(y) \cdot \partial_j\varphi(y) = 0, \ 1 \le j \le n-1, \\ e_{jk} &= r^2 \partial_j\varphi \cdot \partial_k\varphi = r^2 h_{jk}, \quad 1 \le j,k \le n-1. \end{aligned} \tag{A.8.12}$$

The fact that $\varphi(y) \cdot \partial_j\varphi(y) = 0$ follows from applying $\partial/\partial y_j$ to the identity $\varphi(y) \cdot \varphi(y) \equiv 1$. To summarize,

$$(e_{jk}) = \begin{pmatrix} 1 & \\ & r^2 h_{\ell m} \end{pmatrix}. \tag{A.8.13}$$

One implication of (A.8.13) is that

$$\sqrt{e} = r^{n-1}\sqrt{h}, \tag{A.8.14}$$

so we have the following result for integrating a function in spherical polar coordinates:

$$\int\limits_{\mathbb{R}^n} f(x)\,dx = \int\limits_{S^{n-1}} \int_0^\infty f(r\omega) r^{n-1}\,dr\,dS(\omega). \tag{A.8.15}$$

Of primary significance to us now is that the Laplace operator (A.8.1) on $\mathbb{R}^n$ takes the following form in spherical polar coordinates:

$$\Delta u(r\omega) = \partial_r^2 u(r\omega) + \frac{n-1}{r}\partial_r u(r\omega) + \frac{1}{r^2}\Delta_S u(r\omega), \tag{A.8.16}$$

where Δ_S is the Laplace-Beltrami operator on S^{n-1}. In case u is a radial function, say $f(|x|)$, we obtain

$$u(x) = f(|x|) \Longrightarrow \Delta u(x) = f''(|x|) + \frac{n-1}{|x|} f'(|x|). \tag{A.8.17}$$

Appendix B

Linear algebra and multilinear algebra

This appendix treats some topics in linear algebra of use in the main text, particularly multilinear algebra, including tensor products and exterior algebra. This material plays an important role in the treatment of classes of representations of $\mathrm{U}(n)$, in Chapter 4, and similarly for representations of $\mathrm{SO}(n)$, and also for results on Clifford algebras in Chapter 7.

Section B.1 develops the basic theory of determinants of $n \times n$ matrices. Given $A \in M(n, \mathbb{F})$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, we analyze $\det A$ as a function of the columns of A that is linear in each column and changes sign when two columns are switched. That is, we regard

$$\det : \mathbb{F}^n \times \cdots \times \mathbb{F}^n \longrightarrow \mathbb{F}, \tag{B.0.1}$$

linear in each argument. This is a paradigm example of a multilinear map.

Section B.2 treats multilinear maps in general, bringing in

$$\mathcal{M}(V_1, \dots, V_\ell; W), \tag{B.0.2}$$

the space of maps $\beta : V_1 \times \cdots \times V_\ell \to W$ that are linear in each variable. As mentioned above, a prime example of such a map is the determinant, as in (B.0.1), with $\ell = n$, $V_j = \mathbb{F}^n$. In this case, $V_1 = \cdots = V_\ell = V$ (which equals $\mathbb{F}^n$), and the resulting special case of (B.0.2) is denoted

$$\mathcal{M}^\ell(V, W). \tag{B.0.3}$$

The determinant det provides an element of

$$\mathrm{Alt}^\ell(V, W), \tag{B.0.4}$$

with $\ell = n, V = \mathbb{F}^n, W = \mathbb{F}$, where an element of (B.0.4) is a multilinear map $\beta(v_1, \dots, v_\ell)$ that changes sign when two elements, e.g., v_j and v_k with $j \neq k$, are interchanged.

In §B.3 we treat tensor products. If V_j are finite-dimensional vector spaces, of dimension d_j, then $V_1 \otimes \cdots \otimes V_\ell$ is a vector space, of dimension $d_1 \cdots d_\ell$, for which we have a natural isomorphism

$$\mathcal{M}(V_1, \dots, V_\ell; W) \xrightarrow{\approx} \mathcal{L}(V_1 \otimes \cdots \otimes V_\ell, W), \tag{B.0.5}$$

for each vector space W. This tensor product construction ties together linear algebra and multilinear algebra.

Section B.4 treats exterior algebra, which provides a natural algebraic extension of the theory of the determinant. If V is a finite-dimensional vector space over $\mathbb{F}$, we set $\Lambda^0 V' = \mathbb{F},\ \Lambda^1 V' = V'$, and, generally,

$$\Lambda^k V' = \mathrm{Alt}^k(V, \mathbb{F}). \tag{B.0.6}$$

This sequence of vector spaces carries a wedge product,

$$\alpha \in \Lambda^k V',\ \beta \in \Lambda^\ell V' \Longrightarrow \alpha \wedge \beta \in \Lambda^{k+\ell} V'. \tag{B.0.7}$$

Topics treated in §B.4 include an approach to Cramer's formula, given here in terms of the structure of the exterior algebra. Section B.5 sketches an alternative approach to exterior algebra.

Section B.6 establishes the simplicity of the algebra $M(n, \mathbb{F})$, when $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, of use in the proof of Proposition 7.5.1. Section B.7 discusses the discriminant of an $n \times n$ matrix A and relates it to the behavior of $\operatorname{ad} A$.

B.1. Determinants

Determinants arise in the study of inverting a matrix. To take the 2×2 case, solving for x and y the system

$$\begin{aligned} ax + by &= u, \\ cx + dy &= v \end{aligned} \tag{B.1.1}$$

can be done by multiplying these equations by d and b, respectively, and subtracting, and by multiplying them by c and a, respectively, and subtracting, yielding

$$\begin{aligned} (ad - bc)x &= du - bv, \\ (ad - bc)y &= av - cu. \end{aligned} \tag{B.1.2}$$

The factor on the left is

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc, \tag{B.1.3}$$

and solving (B.1.2) for x and y leads to

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Longrightarrow A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \tag{B.1.4}$$

provided $\det A \neq 0$.

We now consider determinants of $n \times n$ matrices. Let $M(n, \mathbb{F})$ denote the set of $n \times n$ matrices with entries in $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. We write

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = (a_1, \dots, a_n), \tag{B.1.5}$$

where

$$a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \tag{B.1.6}$$

is the jth column of A. The determinant is defined as follows.

Proposition B.1.1. *There is a unique function*

$$\vartheta : M(n, \mathbb{F}) \longrightarrow \mathbb{F}, \tag{B.1.7}$$

satisfying the following three properties:

(a) ϑ *is linear in each column* a_j *of* A,
(b) $\vartheta(\widetilde{A}) = -\vartheta(A)$ *if* $\widetilde{A}$ *is obtained from* A *by interchanging two columns,*
(c) $\vartheta(I) = 1$.

This defines the determinant:

$$\vartheta(A) = \det A. \tag{B.1.8}$$

If (c) is replaced by

(c′) $\vartheta(I) = r$,

then

$$\vartheta(A) = r \det A. \tag{B.1.9}$$

The proof will involve constructing an explicit formula for $\det A$ by following the rules (a)–(c). We start with the case $n = 3$. We have

$$\det A = \sum_{j=1}^{3} a_{j1} \det(e_j, a_2, a_3), \tag{B.1.10}$$

by applying (a) to the first column of A, $a_1 = \sum_j a_{j1} e_j$. Here and below, $\{e_j : 1 \le j \le n\}$ denotes the standard basis of $\mathbb{F}^n$, so e_j has a 1 in the jth slot and 0s elsewhere. Applying (a) to the second and third columns gives

$$\begin{aligned} \det A &= \sum_{j,k=1}^{3} a_{j1} a_{k2} \det(e_j, e_k, a_3) \\ &= \sum_{j,k,\ell=1}^{3} a_{j1} a_{k2} a_{\ell 3} \det(e_j, e_k, e_\ell). \end{aligned} \tag{B.1.11}$$

This is a sum of 27 terms, but most of them are 0. Note that rule (b) implies

$$\det B = 0 \ \text{ whenever } \ B \ \text{ has two identical columns.} \tag{B.1.12}$$

Hence $\det(e_j, e_k, e_\ell) = 0$ unless j, k, and ℓ are distinct, that is, unless (j, k, ℓ) is a *permutation* of $(1, 2, 3)$. Now rule (c) says

$$\det(e_1, e_2, e_3) = 1, \tag{B.1.13}$$

and we see from rule (b) that $\det(e_j, e_k, e_\ell) = 1$ if one can convert (e_j, e_k, e_ℓ) to (e_1, e_2, e_3) by an even number of column interchanges, and $\det(e_j, e_k, e_\ell) = -1$ if it takes an odd number of interchanges. Explicitly,

$$\begin{aligned} &\det(e_1, e_2, e_3) = 1, \quad \det(e_1, e_3, e_2) = -1, \\ &\det(e_2, e_3, e_1) = 1, \quad \det(e_2, e_1, e_3) = -1, \\ &\det(e_3, e_1, e_2) = 1, \quad \det(e_3, e_2, e_1) = -1. \end{aligned} \tag{B.1.14}$$

Consequently (B.1.11) yields

$$\begin{aligned} \det A = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} \\ + a_{21}a_{32}a_{13} - a_{21}a_{12}a_{33} \\ + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}. \end{aligned} \tag{B.1.15}$$

Note that the second indices occur in $(1, 2, 3)$ order in each product. We can rearrange these products so that the *first* indices occur in $(1, 2, 3)$ order:

$$\begin{aligned} \det A = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} \\ + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} \\ + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}. \end{aligned} \tag{B.1.16}$$

We now tackle the case of general n. Parallel to (B.1.10)–(B.1.11), we have

$$\begin{aligned} \det A &= \sum_j a_{j1} \det(e_j, a_2, \dots, a_n) = \cdots \\ &= \sum_{j_1,\dots,j_n} a_{j_1 1} \cdots a_{j_n n} \det(e_{j_1}, \dots e_{j_n}), \end{aligned} \tag{B.1.17}$$

by applying rule (a) to each of the n columns of A. As before, (B.1.12) implies $\det(e_{j_1}, \dots, e_{j_n}) = 0$ unless $(j_1, \dots, j_n)$ are all distinct, that is, unless $(j_1, \dots, j_n)$ is a permutation of the set $(1, 2, \dots, n)$. We set

$$S_n = \text{ set of permutations of } (1, 2, \dots, n). \tag{B.1.18}$$

That is, S_n consists of elements σ, mapping the set $\{1, \dots, n\}$ to itself,

$$\sigma : \{1, 2, \dots, n\} \longrightarrow \{1, 2, \dots, n\}, \tag{B.1.19}$$

that are one-to-one and onto. We can compose two such permutations, obtaining the product $\sigma\tau \in S_n$, given σ and τ in S_n. A permutation that interchanges just two elements of $\{1, \dots, n\}$, say j and k $(j \neq k)$, is called a *transposition*, and labeled (jk). It is easy to see that each permutation of $\{1, \dots, n\}$ can be achieved by successively transposing pairs of elements of this set. That is, each element $\sigma \in S_n$ is a product of transpositions. We claim that

$$\det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = (\operatorname{sgn} \sigma) \det(e_1, \dots, e_n) = \operatorname{sgn} \sigma, \tag{B.1.20}$$

where

(B.1.21)

$$\begin{aligned} \operatorname{sgn} \sigma = \ & 1 \quad \text{if } \sigma \text{ is a product of an even number of transpositions,} \\ & -1 \quad \text{if } \sigma \text{ is a product of an odd number of transpositions.} \end{aligned}$$

In fact, the first identity in (B.1.20) follows from rule (b) and the second identity from rule (c).

There is one point to be checked here. Namely, we claim that a given $\sigma \in S_n$ cannot simultaneously be written as a product of an even number of transpositions and an odd number of transpositions. If σ could be so written, $\operatorname{sgn} \sigma$ would not be well defined, and it would be impossible to satisfy condition (b), so Proposition B.1.1 would fail. One neat way to see that $\operatorname{sgn} \sigma$ is well defined is the following. Let $\sigma \in S_n$ act on functions of n variables by

$$(\sigma f)(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)}). \tag{B.1.22}$$

It is readily verified that if also $\tau \in S_n$,

$$g = \sigma f \Longrightarrow \tau g = (\tau\sigma) f. \tag{B.1.23}$$

Now, let P be the polynomial

$$P(x_1, \dots, x_n) = \prod_{1 \le j < k \le n} (x_j - x_k). \tag{B.1.24}$$

One readily has

$$(\sigma P)(x) = -P(x), \quad \text{whenever} \quad \sigma \quad \text{is a transposition}, \tag{B.1.25}$$

and hence, by (B.1.23),

$$(\sigma P)(x) = (\operatorname{sgn} \sigma) P(x), \quad \forall\, \sigma \in S_n, \tag{B.1.26}$$

and $\operatorname{sgn} \sigma$ is well defined.

The proof of (B.1.20) is complete, and substitution into (B.1.17) yields the formula

$$\det A = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}. \tag{B.1.27}$$

It is routine to check that this satisfies the properties (a)–(c). Regarding (b), note that if $\vartheta(A)$ denotes the right side of (B.1.27) and $\widetilde{A}$ is obtained from A by applying a permutation τ to the columns of A, so $\widetilde{A} = (a_{\tau(1)}, \dots, a_{\tau(n)})$, then

$$\begin{aligned} \vartheta(\widetilde{A}) &= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{\sigma(1)\tau(1)} \cdots a_{\sigma(n)\tau(n)} \\ &= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{\sigma\tau^{-1}(1)1} \cdots a_{\sigma\tau^{-1}(n)n} \\ &= \sum_{\omega \in S_n} (\operatorname{sgn} \omega\tau) a_{\omega(1)1} \cdots a_{\omega(n)n} \\ &= (\operatorname{sgn} \tau) \vartheta(A), \end{aligned} \tag{B.1.28}$$

the last identity because

$$\operatorname{sgn} \omega\tau = (\operatorname{sgn} \omega)(\operatorname{sgn} \tau), \quad \forall\, \omega, \tau \in S_n. \tag{B.1.29}$$

As for the final part of Proposition B.1.1, if (c) is replaced by (c′), then (B.1.20) is replaced by

$$\vartheta(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = r(\operatorname{sgn}\sigma), \tag{B.1.30}$$

and (B.1.9) follows.

Remark. Some authors take (B.1.27) as a definition of the determinant. Our perspective is that, while (B.1.27) is a useful *formula* for the determinant, it is a bad *definition*, indeed one that has perhaps led to a bit of fear and loathing among math students.

Remark. Here is another formula for $\operatorname{sgn}\sigma$, which the reader is invited to verify. If $\sigma \in S_n$,

$$\operatorname{sgn}\sigma = (-1)^{\kappa(\sigma)}, \tag{B.1.31}$$

where

$$\begin{aligned} \kappa(\sigma) = {}&\text{number of pairs } (j,k) \text{ such that } 1 \le j < k \le n, \\ &\text{but } \sigma(j) > \sigma(k). \end{aligned} \tag{B.1.32}$$

Note that

$$a_{\sigma(1)1} \cdots a_{\sigma(n)n} = a_{1\tau(1)} \cdots a_{n\tau(n)}, \quad \text{with} \quad \tau = \sigma^{-1}, \tag{B.1.33}$$

and $\operatorname{sgn}\sigma = \operatorname{sgn}\sigma^{-1}$, so, parallel to (B.1.16), we also have

$$\det A = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \tag{B.1.34}$$

Comparison with (B.1.27) gives

$$\det A = \det A^t, \tag{B.1.35}$$

where $A = (a_{jk}) \Rightarrow A^t = (a_{kj})$. Note that the jth column of A^t has the same entries as the jth row of A. In light of this, we have:

Corollary B.1.2. *In Proposition B.1.1, one can replace "columns" by "rows."*

The following is a key property of the determinant.

Proposition B.1.3. *Given A and B in $M(n, \mathbb{F})$,*

$$\det(AB) = (\det A)(\det B). \tag{B.1.36}$$

Proof. For fixed A, apply Proposition B.1.1 to

$$\vartheta_1(B) = \det(AB). \tag{B.1.37}$$

If $B = (b_1, \dots, b_n)$, with jth column b_j, then

$$AB = (Ab_1, \dots, Ab_n). \tag{B.1.38}$$

Clearly rule (a) holds for ϑ_1. Also, if $\widetilde{B} = (b_{\sigma(1)}, \dots, b_{\sigma(n)})$ is obtained from B by permuting its columns, then $A\widetilde{B}$ has columns $(Ab_{\sigma(1)}, \dots, Ab_{\sigma(n)})$, obtained by permuting the columns of AB in the same fashion. Hence rule (b) holds for ϑ_1. Finally, rule (c′) holds for ϑ_1, with $r = \det A$, and (B.1.36) follows. $\square$

Corollary B.1.4. *If $A \in M(n, \mathbb{F})$ is invertible, then $\det A \neq 0$.*

Proof. If A is invertible, there exists $B \in M(n, \mathbb{F})$ such that $AB = I$. Then, by (B.1.36), $(\det A)(\det B) = 1$, so $\det A \neq 0$. $\square$

The converse of Corollary B.1.4 also holds. Before proving it, it is convenient to show that the determinant is invariant under a certain class of column operations, given as follows.

Proposition B.1.5. *If $\widetilde{A}$ is obtained from $A = (a_1, \dots, a_n) \in M(n, \mathbb{F})$ by adding ca_ℓ to a_k for some $c \in \mathbb{F}$, $\ell \neq k$, then*

$$\det \widetilde{A} = \det A. \tag{B.1.39}$$

Proof. By rule (a), $\det \widetilde{A} = \det A + c \det A^b$, where A^b is obtained from A by replacing the column a_k by a_ℓ. Hence A^b has two identical columns, so $\det A^b = 0$, and (B.1.39) holds. $\square$

We now extend Corollary B.1.4.

Proposition B.1.6. *If $A \in M(n, \mathbb{F})$, then A is invertible if and only if $\det A \neq 0$.*

Proof. We have half of this from Corollary B.1.4. To finish, assume A is not invertible. This implies the columns $a_1, \dots, a_n$ of A are linearly dependent. Hence, for some k,

$$a_k + \sum_{\ell \neq k} c_\ell a_\ell = 0, \tag{B.1.40}$$

with $c_\ell \in \mathbb{F}$. Now we can apply Proposition B.1.5 to obtain $\det A = \det \widetilde{A}$, where $\widetilde{A}$ is obtained by adding $\sum c_\ell a_\ell$ to a_k. But then the kth column of $\widetilde{A}$ is 0, so $\det A = \det \widetilde{A} = 0$. This finishes the proof of Proposition B.1.6. $\square$

B.2. Multilinear mappings

If $V_1, \dots, V_\ell$ and W are vector spaces over $\mathbb{F}$, we set

$$\begin{aligned}\mathcal{M}(V_1, \dots, V_\ell; W) = \text{set of mappings } \ \beta : V_1 \times \cdots \times V_\ell \to W \\ \text{that are linear in each variable.}\end{aligned} \tag{B.2.1}$$

That is, for each $j \in \{1, \dots, \ell\}$,

$$\begin{aligned} &v_j, w_j \in V_j,\ a, b \in \mathbb{F} \Longrightarrow \\ &\qquad \beta(u_1, \dots, av_j + bw_j, \dots, u_\ell) \\ &\qquad = a\beta(u_1, \dots, v_j, \dots, u_\ell) + b\beta(u_1, \dots, w_j, \dots, u_\ell). \end{aligned} \tag{B.2.2}$$

This has the natural structure of a vector space, and one readily computes that

$$\dim \mathcal{M}(V_1, \dots, V_\ell; W) = (\dim V_1) \cdots (\dim V_\ell)(\dim W). \tag{B.2.3}$$

If $\{e_{j,1}, \dots, e_{j,d_j}\}$ is a basis of V_j (of dimension d_j), then β is uniquely determined by the elements

$$\begin{aligned} &b_j \in W, \quad b_j = \beta(e_{1,j_1}, \dots, e_{\ell,j_\ell}\}, \\ &j = (j_1, \dots, j_\ell), \quad 1 \le j_\nu \le d_\nu. \end{aligned} \tag{B.2.4}$$

In many cases of interest, all the V_j are the same. Then we set

$$\mathcal{M}^\ell(V, W) = \mathcal{M}(V_1, \dots, V_\ell; W), \quad V_1 = \cdots = V_\ell = V. \tag{B.2.5}$$

This is the space of ℓ-linear maps from V to W. It has two distinguished subspaces,

$$\operatorname{Sym}^\ell(V, W), \quad \operatorname{Alt}^\ell(V, W), \tag{B.2.6}$$

where, given $\beta \in \mathcal{M}^\ell(V, W)$,

$$\begin{aligned} &\beta \in \operatorname{Sym}^\ell(V, W) \Longleftrightarrow \\ &\beta(v_1, \dots, v_j, \dots, v_k, \dots, v_\ell) = \beta(v_1, \dots, v_k, \dots, v_j, \dots, v_\ell), \\ &\beta \in \operatorname{Alt}^\ell(V, W) \Longleftrightarrow \\ &\beta(v_1, \dots, v_j, \dots, v_k, \dots, v_\ell) = -\beta(v_1, \dots, v_k, \dots, v_j, \dots, v_\ell), \end{aligned} \tag{B.2.7}$$

whenever $1 \le j < k \le \ell$.

Here are some basic examples of multilinear maps. There is $\vartheta = \det : M(n, \mathbb{F}) \to \mathbb{F}$ as an element

$$\vartheta \in \operatorname{Alt}^n(\mathbb{F}^n, \mathbb{F}). \tag{B.2.8}$$

For $A \in M(n, \mathbb{F})$, $\det A$ is linear in each column of A and changes sign upon switching any two columns. Another example is the cross product

$$\kappa \in \operatorname{Alt}^2(\mathbb{R}^3, \mathbb{R}^3), \quad \kappa(u, v) = u \times v. \tag{B.2.9}$$

Other examples of multilinear maps include the matrix product

$$\Pi \in \mathcal{M}^2(M(n,\mathbb{F}), M(n,\mathbb{F})), \quad \Pi(A,B) = AB, \tag{B.2.10}$$

and the matrix commutator,

$$\mathcal{C} \in \mathrm{Alt}^2(M(n,\mathbb{F}), M(n,\mathbb{F})), \quad \mathcal{C}(A,B) = AB - BA, \tag{B.2.11}$$

and anticommutator,

$$\mathcal{A} \in \mathrm{Sym}^2(M(n,\mathbb{F}), M(n,\mathbb{F})), \quad \mathcal{A}(A,B) = AB + BA. \tag{B.2.12}$$

Considerations of multilinear maps lead naturally to material treated in the next two sections, namely tensor products and exterior algebra. In §B.3 we define the tensor product $V_1 \otimes \cdots \otimes V_\ell$ of finite-dimensional vector spaces and describe a natural isomorphism

$$\mathcal{M}(V_1, \dots, V_\ell; W) \approx \mathcal{L}(V_1 \otimes \cdots \otimes V_\ell, W). \tag{B.2.13}$$

In §B.4 we discuss spaces $\Lambda^k V$ and describe a natural isomorphism

$$\mathrm{Alt}^k(V, W) \approx \mathcal{L}(\Lambda^k V, W). \tag{B.2.14}$$

B.3. Tensor products

Here all vector spaces will be finite-dimensional vector spaces over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$).

Definition. Given vector spaces $V_1, \dots, V_\ell$, the tensor product $V_1 \otimes \cdots \otimes V_\ell$ is the space of ℓ-linear maps

$$\beta : V_1' \times \cdots \times V_\ell' \longrightarrow \mathbb{F}. \tag{B.3.1}$$

Given $v_j \in V_j$, we define $v_1 \otimes \cdots \otimes v_\ell \in V_1 \otimes \cdots \otimes V_\ell$ by

$$(v_1 \otimes \cdots \otimes v_\ell)(w_1, \dots, w_\ell) = \langle v_1, w_1 \rangle \cdots \langle v_\ell, w_\ell \rangle, \quad w_j \in V_j'. \tag{B.3.2}$$

If $\{e_{j,1}, \dots, e_{j,d_j}\}$ is a basis of V_j (of dimension d_j), with dual basis $\{\varepsilon_{j,1}, \dots, \varepsilon_{j,d_j}\}$ for V_j', then β in (B.3.1) is uniquely determined by the numbers

$$b_j = \beta(\varepsilon_{1,j_1}, \dots, \varepsilon_{\ell,j_\ell}\}, \quad j = (j_1, \dots, j_\ell),\ 1 \le j_\nu \le d_\nu. \tag{B.3.3}$$

It follows that

$$\dim V_1 \otimes \cdots \otimes V_\ell = d_1 \cdots d_\ell, \tag{B.3.4}$$

and a basis of $V_1 \otimes \cdots \otimes V_\ell$ is given by

$$e_{1,j_1} \otimes \cdots \otimes e_{\ell,j_\ell}, \quad 1 \le j_\nu \le d_\nu. \tag{B.3.5}$$

The following is a universal property for the tensor product.

Proposition B.3.1. *Given vector spaces V_j and W, there is a natural isomorphism*

$$\Phi : \mathcal{M}(V_1, \dots, V_\ell; W) \xrightarrow{\approx} \mathcal{L}(V_1 \otimes \cdots \otimes V_\ell, W). \tag{B.3.6}$$

Proof. Given an ℓ-linear map

$$\alpha : V_1 \times \cdots \times V_\ell \longrightarrow W, \tag{B.3.7}$$

the map $\Phi\alpha : V_1 \otimes \cdots \otimes V_\ell \to W$ should satisfy

$$\Phi\alpha(v_1 \otimes \cdots \otimes v_\ell) = \alpha(v_1, \dots, v_\ell), \quad v_j \in V_j. \tag{B.3.8}$$

In fact, in terms of the basis (B.3.5) of $V_1 \otimes \cdots \otimes V_\ell$, we can specify that

$$\Phi\alpha(e_{1,j_1} \otimes \cdots \otimes e_{\ell,j_\ell}) = \alpha(e_{1,j_1}, \dots, e_{\ell,j_\ell}), \quad 1 \le j_\nu \le d_\nu, \tag{B.3.9}$$

and then extend $\Phi\alpha$ by linearity. Such an extension uniquely defines $\Phi\alpha \in \mathcal{L}(V_1 \otimes \cdots \otimes V_\ell, W)$, and it satisfies (B.3.8). In light of this, it follows that the construction of $\Phi\alpha$ is independent of the choice of bases of $V_1, \dots, V_\ell$. We see that Φ is then injective. In fact, if $\Phi\alpha = 0$, then (B.3.9) is identically 0, so $\alpha = 0$. Since $\mathcal{M}(V_1, \dots, V_\ell; W)$ and $\mathcal{L}(V_1 \otimes \cdots \otimes V_\ell, W)$ both have dimension $d_1 \cdots d_\ell (\dim W)$, the isomorphism property of Φ follows. $\square$

We next note that linear maps $A_j : V_j \to W_j$ naturally induce a linear map

$$A_1 \otimes \cdots \otimes A_\ell : V_1 \otimes \cdots \otimes V_\ell \longrightarrow W_1 \otimes \cdots \otimes W_\ell, \tag{B.3.10}$$

as follows. If $\omega_j \in W_j'$, and $\beta : V_1' \times \cdots \times V_\ell' \to \mathbb{F}$ defines $\beta \in V_1 \otimes \cdots \otimes V_\ell$, then

$$(A_1 \otimes \cdots \otimes A_\ell)\beta(\omega_1, \dots, \omega_\ell) = \beta(A_1^t \omega_1, \dots, A_\ell^t \omega_\ell), \tag{B.3.11}$$

with $A_j^t \omega_j \in V_j'$. One sees readily that, for $v_j \in V_j$,

$$(A_1 \otimes \cdots \otimes A_\ell)(v_1 \otimes \cdots \otimes v_\ell) = (A_1 v_1) \otimes \cdots \otimes (A_\ell v_\ell). \tag{B.3.12}$$

For notational simplicity, we now restrict attention to the case $\ell = 2$, i.e., to tensor products of two vector spaces. The following is straightforward.

Proposition B.3.2. *Given* $A \in \mathcal{L}(V)$, $B \in \mathcal{L}(W)$, *inducing* $A \otimes B \in \mathcal{L}(V \otimes W)$, *suppose* $\operatorname{Spec} A = \{\lambda_j\}$ *and* $\operatorname{Spec} B = \{\mu_k\}$. *Then*

$$\operatorname{Spec} A \otimes B = \{\lambda_j \mu_k\}. \tag{B.3.13}$$

Also,

$$\begin{aligned} \mathcal{E}(A \otimes B, \sigma) = \operatorname{Span}\{v \otimes w : v \in \mathcal{E}(A, \lambda_j), \\ w \in \mathcal{E}(B, \mu_k),\ \sigma = \lambda_j \mu_k\}, \end{aligned} \tag{B.3.14}$$

and

$$\begin{aligned} \mathcal{GE}(A \otimes B, \sigma) = \operatorname{Span}\{v \otimes w : v \in \mathcal{GE}(A, \lambda_j), \\ w \in \mathcal{GE}(B, \mu_k),\ \sigma = \lambda_j \mu_k\}. \end{aligned} \tag{B.3.15}$$

Furthermore,

$$\operatorname{Spec}(A \otimes I + I \otimes B) = \{\lambda_j + \mu_k\}, \tag{B.3.16}$$

and we have

$$\begin{aligned} \mathcal{E}(A \otimes I + I \otimes B, \tau) = \operatorname{Span}\{v \otimes w : v \in \mathcal{E}(A, \lambda_j), \\ w \in \mathcal{E}(B, \mu_k),\ \tau = \lambda_j + \mu_k\}, \end{aligned} \tag{B.3.17}$$

and

$$\begin{aligned} \mathcal{GE}(A \otimes I + I \otimes B, \tau) = \operatorname{Span}\{v \otimes w : v \in \mathcal{GE}(A, \lambda_j), \\ w \in \mathcal{GE}(B, \mu_k),\ \tau = \lambda_j + \mu_k\}. \end{aligned} \tag{B.3.18}$$

B.4. Exterior algebra

Let V be a finite dimensional vector space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$), with dual V'. We define the spaces $\Lambda^k V'$ as follows:

$$\Lambda^0 V' = \mathbb{F}, \quad \Lambda^1 V' = V', \tag{B.4.1}$$

and, for $k \geq 2$,

$$\begin{aligned} \Lambda^k V' = \text{set of } k\text{-linear maps } \alpha : V \times \cdots \times V \to \mathbb{F} \\ \text{that are anti-symmetric,} \end{aligned} \tag{B.4.2}$$

i.e.,

$$\alpha(v_1, \dots, v_j, \dots, v_\ell, \dots, v_k) = -\alpha(v_1, \dots, v_\ell, \dots, v_j, \dots, v_k), \tag{B.4.3}$$

for $v_1, \dots, v_k \in V$, $1 \leq j < \ell \leq k$. Another way to picture such α is as a map

$$\alpha : M(k \times n, \mathbb{F}) \longrightarrow \mathbb{F} \tag{B.4.4}$$

that is linear in each column $v_1, \dots, v_k$ of $A = (v_1, \dots, v_k) \in M(k \times n, \mathbb{F})$, and satisfies the anti-symmetry condition (B.4.3), if

$$n = \dim V, \quad \text{so} \quad V \approx \mathbb{F}^n. \tag{B.4.5}$$

In case $k = n$, the theory of the determinant implies that any such $\alpha : M(n \times n, \mathbb{F}) \to \mathbb{F}$ must be a multiple of the determinant. We have

Proposition B.4.1. *Given* (B.4.5),

$$\dim \Lambda^n V' = 1. \tag{B.4.6}$$

Before examining $\dim \Lambda^k V'$ for other values of k, let us look into the following. Pick a basis $\{e_1, \dots, e_n\}$ of V, and let $\{\varepsilon_1, \dots, \varepsilon_n\}$ denote the dual basis of V'. Clearly an element $\alpha \in \Lambda^k V'$ is uniquely determined by its values

$$a_j = \alpha(e_{j_1}, \dots, e_{j_k}), \quad j = (j_1, \dots, j_k), \tag{B.4.7}$$

as j runs over the set of k-tuples $(j_1, \dots, j_k)$, with $1 \leq j_\nu \leq n$. Now, α satisfies the anti-symmetry condition (B.4.3) if and only if

$$a_{j_1 \cdots j_k} = (\operatorname{sgn} \sigma) a_{j_{\sigma(1)} \cdots j_{\sigma(k)}}, \tag{B.4.8}$$

for each $\sigma \in S_k$, i.e., for each permutation σ of $\{1, \dots, k\}$. In particular,

$$j_\mu = j_\nu \text{ for some } \mu \neq \nu \Longrightarrow \alpha(e_{j_1}, \dots, e_{j_k}) = 0. \tag{B.4.9}$$

Applying this observation to $k > n$ yields the following:

Proposition B.4.2. *In the setting of Proposition B.4.1,*

$$k > n \Longrightarrow \Lambda^k V' = 0. \tag{B.4.10}$$

Meanwhile, if $1 \le k \le n$, an element α of $\Lambda^k V'$ is uniquely determined by its values

$$a_j = \alpha(e_{j_1}, \dots, e_{j_k}), \quad 1 \le j_1 < \cdots < j_k \le n. \tag{B.4.11}$$

There are $\binom{n}{k}$ such multi-indices, so we have the following (which contains Proposition B.4.1).

Proposition B.4.3. *In the setting of Proposition B.4.1,*

$$1 \le k \le n \Longrightarrow \dim \Lambda^k V' = \binom{n}{k}. \tag{B.4.12}$$

Here is some useful notation. Given the dual basis $\{\varepsilon_1, \dots, \varepsilon_n\}$, we define

$$\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k} \in \Lambda^k V', \tag{B.4.13}$$

for $j_\nu \in \{1, \dots, n\}$, all distinct, by

$$\begin{aligned} &(\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k})(e_{j_1}, \dots, e_{j_k}) = 1, \\ &(\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k})(e_{\ell_1}, \dots, e_{\ell_k}) = 0, \quad \text{if} \quad \{\ell_1, \dots, \ell_k\} \ne \{j_1, \dots, j_k\}. \end{aligned} \tag{B.4.14}$$

The anti-symmetry condition then specifies

$$(\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k})(e_{j_{\sigma(1)}}, \dots, e_{j_{\sigma(k)}}) = \operatorname{sgn} \sigma, \quad \text{for} \quad \sigma \in S_k. \tag{B.4.15}$$

Note that

$$\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k} = (\operatorname{sgn} \sigma)\, \varepsilon_{j_{\sigma(1)}} \wedge \cdots \wedge \varepsilon_{j_{\sigma(k)}}, \tag{B.4.16}$$

if $\sigma \in S_k$. In light of this, if not all $\{j_1, \dots, j_k\}$ are distinct, i.e., if $j_\mu = j_\nu$ for some $\mu \ne \nu$, we say (B.4.16) vanishes, i.e.,

$$\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k} = 0 \ \text{ if } \ j_\mu = j_\nu \text{ for some } \mu \ne \nu. \tag{B.4.17}$$

Then, for arbitrary $\alpha \in \Lambda^k V'$, we can write

$$\alpha = \frac{1}{k!} \sum_j a_j\, \varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k}, \tag{B.4.18}$$

as j runs over all k-tuples, and a_j is as in (B.4.7). Alternatively, we can write

$$\alpha = \sum_{1 \le j_1 < \cdots < j_k \le n} a_j\, \varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k}, \tag{B.4.19}$$

with a_j as in (B.4.7). Proposition B.4.3 has the following more explicit form.

Proposition B.4.4. *In the setting of Proposition B.4.3, if* $1 \le k \le n$,

$$\{\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k} : 1 \le j_1 < \cdots < j_k \le n\} \ \ \textit{is a basis of} \ \ \Lambda^k V'. \tag{B.4.20}$$

The products arising in (B.4.13)–(B.4.20) are called *wedge products*. As these formulas suggest, it is useful to define wedge products as bilinear maps

$$w : \Lambda^k V' \times \Lambda^\ell V' \longrightarrow \Lambda^{k+\ell} V', \quad w(\alpha, \beta) = \alpha \wedge \beta, \tag{B.4.21}$$

such that

$$(\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k}) \wedge (\varepsilon_{m_1} \wedge \cdots \wedge \varepsilon_{m_\ell}) = \varepsilon_{j_1} \wedge \cdots \varepsilon_{j_k} \wedge \varepsilon_{m_1} \wedge \cdots \wedge \varepsilon_{m_\ell}, \tag{B.4.22}$$

with equivalencies as in (B.4.16)–(B.4.17). We also want to define (B.4.21) in a fashion that does not depend on the choice of basis of V (and associated dual basis of V'). The following result gives a clue as to how to do this.

Proposition B.4.5. *If* $\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k} \in \Lambda^k V'$ *is specified by* (B.4.14)–(B.4.17), *then, for* $v_1, \dots, v_k \in V$,

$$(\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k})(v_1, \dots, v_k) = \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \varepsilon_{j_{\sigma(1)}}(v_1) \cdots \varepsilon_{j_{\sigma(k)}}(v_k). \tag{B.4.23}$$

Proof. We set

$$v_\ell = \sum_{j=1}^n a_{j\ell} e_j, \quad a_{j\ell} = \varepsilon_j(v_\ell), \tag{B.4.24}$$

and substitute into the left side of (B.4.23), obtaining

$$\sum_{\ell_1, \dots, \ell_k = 1}^n \varepsilon_{\ell_1}(v_1) \cdots \varepsilon_{\ell_k}(v_k) (\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k})(e_{\ell_1}, \dots, e_{\ell_k}), \tag{B.4.25}$$

and (B.4.14)–(B.4.17) gives

$$(\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k})(e_{\ell_1}, \dots, e_{\ell_k}) = 0, \tag{B.4.26}$$

unless $\{j_1, \dots, j_k\} = \{\ell_1, \dots, \ell_k\}$, and the k numbers are all distinct, in which case $\ell_\nu = j_{\sigma(\nu)}$ for some $\sigma \in S_k$, and we get $\operatorname{sgn} \sigma$ in (B.4.26). Thus (B.4.25) is converted to the right side of (B.4.23). (Both sides of (B.4.23) vanish if the numbers $j_1, \dots, j_k$ are not all distinct.) $\square$

Remark. In case $n = k$, we obtain the fundamental result on the determinant. Note also that the right side of (B.4.23) is equal to

$$\sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \varepsilon_{j_1}(v_{\sigma(1)}) \cdots \varepsilon_{j_k}(v_{\sigma(k)}). \tag{B.4.27}$$

As a further preparation for defining $\alpha \wedge \beta$ in (B.4.21), note that

$$\alpha \in \Lambda^k V' \Rightarrow \alpha(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)}). \tag{B.4.28}$$

We now define the wedge product:

Definition. If $\alpha \in \Lambda^k V'$ and $\beta \in \Lambda^\ell V'$, then $\alpha \wedge \beta \in \Lambda^{k+\ell} V'$ is given by

$$\begin{aligned}(\alpha \wedge \beta)&(v_1, \dots, v_{k+\ell}) \\ &= \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn} \sigma) \alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \cdot \beta(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)}).\end{aligned} \tag{B.4.29}$$

Our first task is to check the fundamental identity (B.4.22).

Proposition B.4.6. *With $\alpha \wedge \beta$ defined as in* (B.4.29), *the identity* (B.4.22) *holds.*

Proof. With $\alpha = \varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k}$ and $\beta = \varepsilon_{m_1} \wedge \cdots \wedge \varepsilon_{m_\ell}$, we have

$$\begin{aligned}(\alpha \wedge \beta)&(v_1, \dots, v_{k+\ell}) \\ &= \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn} \sigma)(\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k})(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \\ &\qquad\qquad \cdot (\varepsilon_{m_1} \wedge \cdots \wedge \varepsilon_{m_\ell})(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)}),\end{aligned} \tag{B.4.30}$$

which expands out to

$$\begin{aligned}&\frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \sum_{\tau \in S_k} \sum_{\rho \in S_\ell} (\operatorname{sgn} \sigma)(\operatorname{sgn} \tau)(\operatorname{sgn} \rho) \\ &\qquad \cdot \varepsilon_{j_1}(v_{\sigma\tau(1)}) \cdots \varepsilon_{j_k}(v_{\sigma\tau(k)}) \cdot \varepsilon_{m_1}(v_{\sigma\rho(k+1)}) \cdots \varepsilon_{m_\ell}(v_{\sigma\rho(k+\ell)}).\end{aligned} \tag{B.4.31}$$

Here, σ permutes $\{1, \dots, k+\ell\}$, τ permutes $\{1, \dots, k\}$, and ρ permutes $\{k+1, \dots, k+\ell\}$. Note that such σ, τ, ρ yield $\gamma(\sigma, \tau, \rho) \in S_{k+\ell}$, with

$$\begin{aligned}\gamma(\sigma, \tau, \rho)(\nu) &= \sigma\tau(\nu) \quad \text{for} \quad 1 \le \nu \le k, \\ &\ \sigma\rho(\nu) \quad \text{for} \quad k+1 \le \nu \le k+\ell,\end{aligned} \tag{B.4.32}$$

and $\operatorname{sgn} \gamma(\sigma, \tau, \rho) = (\operatorname{sgn} \sigma)(\operatorname{sgn} \tau)(\operatorname{sgn} \rho)$. Also, for each fixed $\tau \in S_k,\ \rho \in S_\ell$ $\gamma(\sigma, \tau, \rho)$ runs over $S_{k+\ell}$ once as σ runs over $S_{k+\ell}$. Hence, if we fix τ and ρ in (B.4.31) and just sum over σ, we get

$$\begin{aligned}&\sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn} \gamma(\sigma, \tau, \rho)) \varepsilon_{j_1}(v_{\gamma(\sigma,\tau,\rho)(1)}) \cdots \varepsilon_{j_k}(v_{\gamma(\sigma,\tau,\rho)(k)}) \\ &\qquad\qquad \cdot \varepsilon_{m_1}(v_{\gamma(\sigma,\tau,\rho)(k+1)}) \cdots \varepsilon_{m_\ell}(v_{\gamma(\sigma,\tau,\rho)(k+\ell)}),\end{aligned} \tag{B.4.33}$$

and each such sum is equal to

$$(\varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k} \wedge \varepsilon_{m_1} \wedge \cdots \wedge \varepsilon_{m_\ell})(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+\ell}). \tag{B.4.34}$$

Then summing over $\tau \in S_k$ and $\rho \in S_\ell$ and dividing by $k!\ell!$ also yields (B.4.34), as desired. $\square$

From here, the following is straightforward.

Proposition B.4.7. *The wedge product* $\alpha\wedge\beta$*, defined by* (B.4.29)*, produces a well defined bilinear map* $\Lambda^k V' \times \Lambda^\ell V' \to \Lambda^{k+\ell} V'$*. Furthermore, given* $\alpha \in \Lambda^k V'$ *and* $\beta \in \Lambda^\ell V'$*,*

$$\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha, \tag{B.4.35}$$

and, if also $\gamma \in \Lambda^m V'$*,*

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma). \tag{B.4.36}$$

The wedge product gives us an algebra. We define the *exterior algebra* $\Lambda^* V'$ to be

$$\Lambda^* V' = \bigoplus_{k \geq 0} \Lambda^k V', \tag{B.4.37}$$

keeping in mind that the summands on the right are nonvanishing only for $k \leq n = \dim V$. Proposition B.4.7 says this is an algebra. The element $1 \in \mathbb{F} = \Lambda^0 V' \subset \Lambda^* V'$ acts as the unit in this algebra. The identity (B.4.36) is the associative law for the wedge product. By (B.4.35), this is not a commutative algebra (if $n > 1$).

We next consider the action a linear map on V induces on $\Lambda^* V'$. A linear map $A : V \to V$ induces a linear map

$$\Lambda^k A^t : \Lambda^k V' \longrightarrow \Lambda^k V', \tag{B.4.38}$$

via

$$(\Lambda^k A^t)\alpha(v_1, \dots, v_k) = \alpha(Av_1, \dots, Av_k). \tag{B.4.39}$$

In particular, $\Lambda^1 A^t = A^t : V' \to V'$. A straightforward calculation from (B.4.29) yields

$$\begin{aligned} \alpha \in& \Lambda^k V',\ \beta \in \Lambda^\ell V',\ A \in \mathcal{L}(V) \\ &\Longrightarrow (\Lambda^{k+\ell} A^t)(\alpha \wedge \beta) = (\Lambda^k A^t)\alpha \wedge (\Lambda^\ell A^t)\beta. \end{aligned} \tag{B.4.40}$$

Here is a natural extension of the identity $(AB)^t = B^t A^t$.

Proposition B.4.8. *If* $A, B \in \mathcal{L}(V)$*, then*

$$\Lambda^k (AB)^t = (\Lambda^k B^t)\,(\Lambda^k A^t). \tag{B.4.41}$$

Proof. We have

$$\begin{aligned} \Lambda^k (AB)^t \alpha(v_1, \dots, v_k) &= \alpha(ABv_1, \dots, ABv_k) \\ &= (\Lambda^k A^t)\alpha(Bv_1, \dots, Bv_k) \\ &= (\Lambda^k B^t)(\Lambda^k A^t)\alpha(v_1, \dots, v_k). \end{aligned} \tag{B.4.42}$$

□

We now return to determinants.

Proposition B.4.9. *If $A \in \mathcal{L}(V)$ and $n = \dim V$, then, for $\omega \in \Lambda^n V'$,*

$$(\Lambda^n A^t)\omega = (\det A)\omega. \tag{B.4.43}$$

Proof. We may as well take $\omega = \varepsilon_1 \wedge \cdots \wedge \varepsilon_n$. Then an iteration of (B.4.40) gives

$$(\Lambda^n A^t)\omega = (A^t \varepsilon_1) \wedge \cdots \wedge (A^t \varepsilon_n). \tag{B.4.44}$$

If $A = (a_{jk})$ with respect to the basis $\{e_j\}$, then $A^t \varepsilon_j = \sum_k a_{jk} \varepsilon_k$, so

$$\begin{aligned} (\Lambda^n A^t)\omega &= \sum_{1 \le k_\nu \le n} a_{1k_1} \cdots a_{nk_n}\, \varepsilon_{k_1} \wedge \cdots \wedge \varepsilon_{k_n} \\ &= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}\, \varepsilon_1 \wedge \cdots \wedge \varepsilon_n \\ &= (\det A)\, \varepsilon_1 \wedge \cdots \wedge \varepsilon_n, \end{aligned} \tag{B.4.45}$$

the last identity constituting a fundamental formula for the determinant. □

Combining Propositions B.4.8 and B.4.9 yields the following multiplicative property for the determinant.

Corollary B.4.10. *If $A, B \in \mathcal{L}(V)$, then*

$$\det(AB) = (\det A)(\det B). \tag{B.4.46}$$

Interior products

We next define the interior product

$$\iota_v : \Lambda^k V' \longrightarrow \Lambda^{k-1} V', \quad \text{for} \quad v \in V, \tag{B.4.47}$$

$k \ge 1$, as follows. If $\alpha \in \Lambda^k V'$, then $\iota_v \alpha \in \Lambda^{k-1} V'$ is defined by

$$(\iota_v \alpha)(v_1, \dots, v_{k-1}) = \alpha(v, v_1, \dots, v_{k-1}). \tag{B.4.48}$$

From this we can compute that, if $\{e_1, \dots, e_n\}$ is a basis of V, with dual basis $\{\varepsilon_1, \dots, \varepsilon_n\}$ for V', then, if $j_1, \dots, j_k$ are distinct,

$$\alpha = \varepsilon_{j_1} \wedge \cdots \wedge \varepsilon_{j_k} \Rightarrow \iota_{e_{j_\ell}} \alpha = (-1)^{\ell-1} \varepsilon_{j_1} \wedge \cdots \wedge \widehat{\varepsilon}_{j_\ell} \wedge \cdots \wedge \varepsilon_{j_k}, \tag{B.4.49}$$

where $\widehat{\varepsilon}_{j_\ell}$ denotes removing the factor ε_{j_ℓ}. Furthermore, for such α,

$$m \notin \{j_1, \dots, j_k\} \Longrightarrow \iota_{e_m} \alpha = 0. \tag{B.4.50}$$

By convention, $\iota_v \alpha = 0$ if $\alpha \in \Lambda^0 V'$.

We make use of the operators $\wedge_k$ and ι_k on $\Lambda^* V'$:

$$\wedge_k \alpha = \varepsilon_k \wedge \alpha, \quad \iota_k \alpha = \iota_{e_k} \alpha. \tag{B.4.51}$$

There is the following useful anticommutation relation:

Proposition B.4.11. *With the notation (m21.51),*

$$\wedge_k \iota_\ell + \iota_\ell \wedge_k = \delta_{k\ell}, \tag{B.4.52}$$

where $\delta_{k\ell} = 1$ *if* $k = \ell$, 0 *otherwise.*

The proof is an exercise. We also have

$$\wedge_j \wedge_k + \wedge_k \wedge_j = 0, \quad \iota_j \iota_k + \iota_k \iota_j = 0. \tag{B.4.53}$$

We mention that (B.4.52) implies the following.

$$(\wedge_w \iota_v + \iota_v \wedge_w)\alpha = \langle v, w\rangle \alpha, \tag{B.4.54}$$

given $\alpha \in \Lambda^k V'$, $w \in V'$, $v \in V$, with the notation

$$\wedge_w \alpha = w \wedge \alpha. \tag{B.4.55}$$

Cramer's formula

Cramer's formula computes a matrix inverse A^{-1} in terms of $\det A$ and the $(n-1) \times (n-1)$ minors of A (or better, of A^t). We present a derivation of such a formula here, using exterior algebra.

Let V be n-dimensional, with dual V'. Let $A \in \mathcal{L}(V)$, with transpose $A^t \in \mathcal{L}(V')$. We bring in the isomorphism

$$\kappa : V \otimes \Lambda^n V' \xrightarrow{\approx} \Lambda^{n-1} V', \tag{B.4.56}$$

given by

$$\kappa(u \otimes \omega)(v_1, \dots, v_{n-1}) = \omega(u, v_1, \dots, v_{n-1}). \tag{B.4.57}$$

We aim to prove the following version of Cramer's formula.

Proposition B.4.12. *If* $A \in \mathcal{L}(V)$ *is invertible, then*

$$(\det A)\, A^{-1} \otimes I = \kappa^{-1} \circ \Lambda^{n-1} A^t \circ \kappa, \tag{B.4.58}$$

in $\mathcal{L}(V \otimes \Lambda^n V')$.

Proof. Since $\Lambda^n A^t = (\det A) I$ in $\mathcal{L}(\Lambda^n V')$, the desired identity (B.4.58) is equivalent to

$$(\Lambda^{n-1} A^t) \circ \kappa = \kappa \circ (A^{-1} \otimes \Lambda^n A^t), \tag{B.4.59}$$

in $\mathcal{L}(V \otimes \Lambda^n V', \Lambda^{n-1} V')$. Recall that $\Lambda^{n-1} A^t \in \mathcal{L}(\Lambda^{n-1} V')$ is defined by

$$(\Lambda^{n-1} A^t)\beta(v_1, \dots, v_{n-1}) = \beta(Av_1, \dots, Av_{n-1}). \tag{B.4.60}$$

Hence if we take $u \otimes \omega \in V \otimes \Lambda^n V'$, we get

$$\begin{aligned} (\Lambda^{n-1} A^t) \circ \kappa(u \otimes \omega)(v_1, \dots, v_{n-1}) &= \kappa(u \otimes \omega)(Av_1, \dots, Av_{n-1}) \\ &= \omega(u, Av_1, \dots, Av_{n-1}). \end{aligned} \tag{B.4.61}$$

On the other hand, since

$$(A^{-1} \otimes \Lambda^n A^t)(u \otimes \omega) = A^{-1}u \otimes \Lambda^n A^t \omega, \tag{B.4.62}$$

we have

$$\begin{aligned} \kappa \circ (A^{-1} \otimes \Lambda^n A^t)&(u \otimes \omega)(v_1, \dots, v_{n-1}) \\ &= \kappa(A^{-1}u \otimes \Lambda^n A^t \omega)(v_1, \dots, v_{n-1}) \\ &= (\Lambda^n A^t \omega)(A^{-1}u, v_1, \dots, v_{n-1}) \\ &= \omega(u, Av_1, \dots, Av_{n-1}), \end{aligned} \tag{B.4.63}$$

which agrees with the right side of (B.4.61). This completes the proof. $\square$

The exterior algebra $\Lambda^* V$

If V is an n-dimensional space, we define $\Lambda^k V$ in a fashion to the definition of $\Lambda^k V'$, simply by switching V and V', using the natural isomorphism $V \approx (V')'$. Thus we set $\Lambda^0 V = \mathbb{F}$, $\Lambda^1 V = V$, and, for $k \geq 2$,

$$\begin{aligned} \Lambda^k V = \text{set of } k\text{-linear maps } \beta : V' \times \cdots \times V' \to \mathbb{F} \\ \text{that are anti-symmetric.} \end{aligned} \tag{B.4.64}$$

All the results from the early part of this section go through, with the roles of V and V', and also of the bases $\{e_1, \dots, e_n\}$ and $\{\varepsilon_1, \dots, \varepsilon_n\}$, interchanged. For example, for $1 \leq k \leq n$,

$$\{e_{j_1} \wedge \cdots \wedge e_{j_k} : 1 \leq j_1 < \cdots < j_k \leq n\} \ \text{ is a basis of } \ \Lambda^k V. \tag{B.4.65}$$

With these facts in mind, we can pass from $A \in \mathcal{L}(V)$ to $A^t \in \mathcal{L}(V')$ to

$$\Lambda^k A : \Lambda^k V \longrightarrow \Lambda^k V, \tag{B.4.66}$$

and, parallel to (B.4.40),

$$\begin{aligned} \alpha \in \Lambda^k V, \ \beta \in \Lambda^\ell V, \ A \in \mathcal{L}(V) \\ \Longrightarrow (\Lambda^{k+\ell} A)(\alpha \wedge \beta) = (\Lambda^k A)\alpha \wedge (\Lambda^\ell A)\beta. \end{aligned} \tag{B.4.67}$$

Consequently,

$$(\Lambda^k A)(e_{j_1} \wedge \cdots \wedge e_{j_k}) = Ae_{j_1} \wedge \cdots \wedge Ae_{j_k}. \tag{B.4.68}$$

We now mention a "universal property" possessed by $\Lambda^k V$. Let W be another finite-dimensional vector space over $\mathbb{F}$, and set

$$\begin{aligned} \mathrm{Alt}^k(V, W) = \text{set of } k\text{-linear maps } V \times \cdots \times V \to W \\ \text{that are anti-symmetric.} \end{aligned} \tag{B.4.69}$$

This has the structure of a finite-dimensional vector space.

Proposition B.4.13. *There is a natural linear isomorphism*

$$\Phi : \mathrm{Alt}^k(V, W) \xrightarrow{\approx} \mathcal{L}(\Lambda^k V, W). \tag{B.4.70}$$

One way to describe Φ is with the aid of a basis $\{e_1, \dots, e_n\}$ of V, leading, as mentioned, to the basis (B.4.65) of $\Lambda^k V$. Given $\alpha \in \mathrm{Alt}^k(V, W)$, hence

$$\alpha : V \times \cdots \times V \longrightarrow W, \tag{B.4.71}$$

we can define $\Phi\alpha : \Lambda^k V \to W$ by

$$(\Phi\alpha)(e_{j_1} \wedge \cdots \wedge e_{j_k}) = \alpha(e_{j_1}, \dots, e_{j_k}). \tag{B.4.72}$$

It is clear that this defines a linear map $\Phi : \mathrm{Alt}^k(V, W) \to \mathcal{L}(\Lambda^k V, W)$. One needs to show that this is an isomorphism and that it is independent of the choice of basis $\{e_j\}$ of V. We leave these tasks to the enthusiastic reader.

Now that, in case $W = \mathbb{F}$, we have

$$\mathrm{Alt}^k(V, \mathbb{F}) = \Lambda^k V', \quad \mathcal{L}(\Lambda^k V, \mathbb{F}) = (\Lambda^k V)', \tag{B.4.73}$$

and Proposition B.4.13 implies that there is a natural isomorphism

$$\Lambda^k V' \approx (\Lambda^k V)'. \tag{B.4.74}$$

B.5. Second perspective on exterior algebra

Let V be an n-dimensional vector space (over $\mathbb{R}$ or $\mathbb{C}$), with basis $\{e_1, \dots, e_n\}$. We define $\Lambda^* V$ by

$$\Lambda^* V = \bigotimes{}^* V / \mathcal{I}, \tag{B.5.1}$$

where $\otimes^* V = \mathbb{R} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots$ is the tensor algebra and

$$\mathcal{I} \ \text{ is the 2-sided ideal generated by } \ \{u \otimes v + v \otimes u : u, v \in V\}. \tag{B.5.2}$$

Equivalently, $\mathcal{I}$ is the 2-sided ideal generated by $\{e_j \otimes e_k + e_k \otimes e_j : 1 \le j, k \le n\}$. We denote the product of $\varphi, \psi \in \Lambda^* V$ by $\varphi \wedge \psi$. Note that

$$u, v \in V \Longrightarrow u \wedge v = -v \wedge u. \tag{B.5.3}$$

We see that

$$\Lambda^* V = \bigoplus_{k=0}^{n} \Lambda^k V, \tag{B.5.4}$$

where $\Lambda^0 V = \mathbb{R}$ (or $\mathbb{C}$) and $\Lambda^k V$ is spanned by

$$\{e_{j_1} \wedge \cdots \wedge e_{j_k} : 1 \le j_1 < \cdots < j_k \le n\}. \tag{B.5.5}$$

Our goal in this appendix is to prove the following result, which was used in §7.3; cf. (7.3.14).

Proposition B.5.1. *The set* (B.5.5) *is linearly independent, hence a basis of $\Lambda^k V$, for each $k \in \{1, \dots, n\}$.*

Proof. We start with $k = n$, where the assertion is that

$$e_1 \wedge \cdots \wedge e_n \neq 0, \tag{B.5.6}$$

or, equivalently,

$$e_1 \wedge \cdots \wedge e_n \neq -e_1 \wedge \cdots \wedge e_n. \tag{B.5.7}$$

Note from (B.5.3) that if $\sigma \in S_n$, i.e., σ is a permutation of $\{1, \dots, n\}$, then

$$e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} = (\operatorname{sgn} \sigma)\, \sigma_1 \wedge \cdots \wedge e_n. \tag{B.5.8}$$

The content of (B.5.7) is that $\operatorname{sgn} \sigma$ is well defined, as a one-dimensional representation of S_n. One way to see this is to represent S_n on the space of functions on $\mathbb{R}^n \approx V$:

$$R(\sigma) f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)}), \tag{B.5.9}$$

and note that $R(\sigma)$ leaves invariant

$$W = \operatorname{Span} P(x), \quad P(x) = \prod_{1 \le j < k \le n} (x_j - x_k). \tag{B.5.10}$$

Furthermore,

$$R(\sigma) P(x) = (\operatorname{sgn} \sigma)\, P(x), \tag{B.5.11}$$

showing that $\operatorname{sgn} \sigma$ is well defined.

Having the result for $k = n$, we proceed by induction. Let $\ell < n$ and suppose we have the result for all $k > \ell$. To establish independence of (B.5.5) with $k = \ell$, suppose

$$\sum_{1 \le j_1 < \cdots < j_\ell \le n} a_{j_1 \cdots j_\ell}\, e_{j_1} \wedge \cdots \wedge e_{j_\ell} = 0. \tag{B.5.12}$$

Then, for each $m \in \{1, \dots, n\}$, wedge this with e_m on the left to get

$$\sum_{m \notin \{j_1, \dots, j_\ell\}} a_{j_1 \cdots j_\ell}\, e_m \wedge e_{j_1} \wedge \cdots \wedge e_{j_\ell} = 0. \tag{B.5.13}$$

One can reorder $(m, j_1, \dots, j_\ell)$ to express (B.5.13) as a linear combination of monomials of the form (B.5.5) with $k = \ell + 1$. The inductive hypothesis yields

$$a_{j_1 \cdots j_\ell} = 0, \tag{B.5.14}$$

for all multi-indices $(j_1, \dots, j_\ell)$ not containing m, for each m, hence (B.5.14) holds for all multi-indices $(j_1, \dots, j_\ell)$. This completes the inductive argument. $\square$

B.6. Simplicity of $\mathrm{M}(n, \mathbb{F})$

The following result (in the case $\mathbb{F} = \mathbb{C}$) was useful in the proof of Proposition 7.5.1.

Proposition B.6.1. *If* $\mathbb{F}$ *is a field, then the associative algebra* $\mathrm{M}(n, \mathbb{F})$ *of* $n \times n$ *matrices with entries in* $\mathbb{F}$ *is simple, i.e., it has no proper two-sided ideal.*

Proof. Suppose $\mathcal{I} \subset \mathrm{M}(n, \mathbb{F})$ is a two-sided ideal, i.e., $A \in \mathcal{I},\ X \in \mathrm{M}(n, \mathbb{F}) \Rightarrow AX \in \mathcal{I}$ and $XA \in \mathcal{I}$. Suppose $\mathcal{I}$ contains a nonzero element A. Say $A = (a_{jk})$ and $a_{\ell m} \neq 0$. Denote by E_{jk} the element of $\mathrm{M}(n, \mathbb{F})$ with a 1 in the jth row and kth column and zeros elsewhere. Thus, if $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{F}^n$,

$$E_{jk} e_\ell = \delta_{k\ell} e_j. \tag{B.6.1}$$

A calculation gives

$$E_{j\ell} A E_{mk} = a_{\ell m} E_{jk}, \tag{B.6.2}$$

for each $j, k, \ell, m \in \{1, \dots, n\}$. Hence, if $a_{\ell m} \neq 0$ it follows that $E_{jk} \in \mathcal{I}$ for each $j, k \in \{1, \dots, n\}$, so $\mathcal{I} = \mathrm{M}(n, \mathbb{F})$. $\square$

B.7. The discriminant of a matrix

Take $A \in \operatorname{End}(\mathbb{C}^n)$. Say $\operatorname{Spec} A = \{\lambda_1, \dots, \lambda_n\}$, counting multiplicities. Then

$$L_A, R_A : \operatorname{End}(\mathbb{C}^n) \longrightarrow \operatorname{End}(\mathbb{C}^n), \quad L_A X = AX, \quad R_A X = XA, \tag{B.7.1}$$

have the same spectrum, with n-fold increases in multiplicity. Since L_A and R_A commute, we can say about $\operatorname{ad} A = L_A - R_A$ that

$$\operatorname{Spec} \operatorname{ad} A = \{\lambda_j - \lambda_k : 1 \le j, k \le n\}. \tag{B.7.2}$$

We thus have

$$\begin{aligned} \det(sI - \operatorname{ad} A) &= \prod_{j,k} \bigl[s - (\lambda_j - \lambda_k)\bigr] \\ &= s^n \prod_{j<k} \bigl[s^2 - (\lambda_j - \lambda_k)^2\bigr] \\ &= (-1)^{n(n-1)/2}\, s^n D(A) + O(s^{n+1}), \end{aligned} \tag{B.7.3}$$

as $s \to 0$, where $D(A)$ is the discriminant of A:

$$D(A) = \prod_{j<k} (\lambda_j - \lambda_k)^2. \tag{B.7.4}$$

It follows that

$$D(A) = \frac{(-1)^{n(n-1)/2}}{n!} \frac{d^n}{ds^n} \det(sI - \operatorname{ad} A)\big|_{s=0}. \tag{B.7.5}$$

Suppose A is diagonal, say $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Let E_{jk} denote the $n \times n$ matrix with a 1 in row j, column k, zeroes elsewhere. We have

$$[A, E_{jk}] = (\lambda_j - \lambda_k) E_{jk}. \tag{B.7.6}$$

It follows readily from (B.7.6) that, when A is diagonal,

$$D(A) = \det \operatorname{ad} A\big|_{\operatorname{End}(\mathbb{C}^n)/\mathcal{D}}, \tag{B.7.7}$$

where $\mathcal{D}$ is the space of complex diagonal matrices. This yields

$$D(A) = \det \operatorname{ad} A\big|_{\mathfrak{u}/\mathfrak{t}}, \tag{B.7.8}$$

when $A \in \mathfrak{t}$, $\mathfrak{u} =$ set of skew-adjoint operators on $\mathbb{C}^n$, $\mathfrak{t} =$ space of diagonal matrices with purely imaginary diagonal entries.

Appendix C

Functional analysis background

This appendix provides background on topics in functional analysis that arise in the text, starting in Chapter 2. One guiding principle is to present enough functional analysis background to support the arguments proving the Peter-Weyl theorem in that chapter. This background is useful for the understanding of other analytical results, such as those on analysis on $U(n)$ presented in Chapter 5, those on spherical harmonics in Chapter 8, and those on analysis on homogeneous spaces in Chapter 9.

Section C.1 introduces the class of Banach spaces and establishes some basic results. Fundamental examples include the space $C(X)$ of continuous functions on a compact space X, and the Lebesgue spaces $L^p(X, \mu)$ of pth power integrable functions on a measure space (X, μ). For $p = 2$, we get the Hilbert space $L^2(X, \mu)$. Other results reviewed in §C.1 include the Stone-Weierstrass theorem, giving sufficient conditions that an algebra $\mathcal{A}$ of continuous functions on a compact space X be dense in $C(X)$. Our first major use of this result is given in Chapter 2, in the first version of the Peter-Weyl theorem.

Section C.2 is devoted to basic results on a Hilbert space H, with emphasis on the notion of an orthonormal basis $\{u_k\}$ of H, and the expansion of an element $v \in H$ in terms of this basis, as a sequence $S_N v$, that converges to v in H-norm.

Section C.3 considers the space $\mathcal{L}(V, W)$ of bounded linear operators from a Banach space V to W. We endow this space with an operator norm, and consider various convergence issues. Section C.4 specializes to the subspace $\mathcal{K}(V, W)$ of compact linear operators. The main result here is the spectral theorem for a compact, self-adjoint operator A on a Hilbert space H, which states that H has an orthonormal basis of eigenvectors of A, with real eigenvalues $\lambda_k \to 0$. This is a key functional analytical tool in the proof of the general Peter-Weyl theorem in Chapter 2.

What we present here is simply a "bare bones" introduction to functional analysis. A more substantial introduction to the subject is given in Appendix A of [**39**] (Outline of functional analysis). Beyond that, one can find full-blown treatments of functional analysis in various texts, such as [**48**].

C.1. Banach spaces

A Banach space V is a complete, normed linear space. A norm on V is a positive function $\|v\|$ having the properties

$$\begin{aligned} \|av\| &= |a| \cdot \|v\|, \quad \forall\, v \in V,\ a \in \mathbb{C} \text{ (or } \mathbb{R}), \\ \|v\| &> 0 \text{ unless } v = 0, \\ \|v + w\| &\le \|v\| + \|w\|. \end{aligned} \tag{C.1.1}$$

The last of these conditions is called the triangle inequality. Given a norm on V, we have the distance function $d(u, v) = \|u - v\|$, making V a metric space.

A metric space is a set X, with distance function $d : X \times X \to \mathbb{R}^+$, satisfying

$$\begin{aligned} &d(u, v) = d(v, u), \\ &d(u, v) > 0 \text{ unless } u = v, \\ &d(u, v) \le d(u, w) + d(w, v). \end{aligned} \tag{C.1.2}$$

A sequence (v_j) is Cauchy provided $d(v_n, v_m) \to 0$ as $m, n \to \infty$. Completeness is the property that each Cauchy sequence converges.

We list some examples of Banach spaces. First let X be a compact metric space, that is, a metric space with the property that each sequence (x_n) has a convergent subsequence. Then $C(X)$, the space of continuous (real or complex valued) functions on X, is a Banach space, with norm

$$\|u\|_{\sup} = \sup\{|u(x)| : x \in X\}. \tag{C.1.3}$$

Also, given $\alpha \in [0, 1]$, we set

$$\mathrm{Lip}^\alpha(X) = \{u \in C(X) : |u(x) - u(y)| \le C d(x, y)^\alpha,\ \forall\, x, y \in X\}. \tag{C.1.4}$$

This is a Banach space, with norm

$$\|u\|_\alpha = \|u\|_{\sup} + \sup_{x,y \in X} \frac{|u(x) - u(y)|}{d(x, y)^\alpha}. \tag{C.1.5}$$

$\mathrm{Lip}^0(X) = C(X)$. The space $\mathrm{Lip}^1(X)$ is typically denoted $\mathrm{Lip}(X)$. For $\alpha \in (0, 1)$, $\mathrm{Lip}^\alpha(X)$ is frequently denoted $C^\alpha(X)$. It is straightforward to verify the conditions (C.1.1) on the proposed norms and to establish completeness.

More subtle examples of Banach spaces are the L^p-spaces, defined as follows. First take $p = 1$. Let (X, μ) be a measure space. We say a measurable function f belongs to $\mathcal{L}^1(X, \mu)$ provided

$$\int_X |f(x)|\, d\mu(x) < \infty. \tag{C.1.6}$$

Elements of $L^1(X,\mu)$ consist of equivalence classes of elements of $\mathcal{L}^1(X,\mu)$, where we say

$$f \sim \tilde{f} \Leftrightarrow f(x) = \tilde{f}(x), \text{ for } \mu\text{-a.e. } x. \tag{C.1.7}$$

Abusing notation, we denote by f both a measurable function in $\mathcal{L}^1(X,\mu)$ and its equivalence class in $L^1(X,\mu)$.

The norm $\|f\|_{L^1}$ is given by (C.1.6). It is easy to see that this norm has the properties (C.1.1). The proof of completeness makes use of the monotone convergence theorem and dominated convergence theorem of Lebesgue theory. We refer to standard books on measure theory (such as [**41**]) for details.

Continuing with a description of L^p-spaces, we define $\mathcal{L}^\infty(X,\mu)$ to consist of bounded, measurable functions, $L^\infty(X,\mu)$ to consist of equivalence classes of such functions, via (C.1.7), and we define $\|f\|_{L^\infty}$ to be the smallest sup of $\tilde{f} \sim f$. It is easy to show that $L^\infty(X,\mu)$ is a Banach space.

For $p \in (1,\infty)$, we define $\mathcal{L}^p(X,\mu)$ to consist of measurable functions f such that

$$\Big[\int_X |f(x)|^p \, d\mu(x)\Big]^{1/p} \tag{C.1.8}$$

is finite. $L^p(X,\mu)$ consists of equivalence classes, via (C.1.7), and the L^p-norm $\|f\|_{L^p}$ is given by (C.1.8). This time it takes some work to verify the triangle inequality, i.e.,

$$\|f+g\|_{L^p} \le \|f\|_{L^p} + \|g\|_{L^p}, \tag{C.1.9}$$

known as Minkowski's inequality. One way to establish this is by the following characterization of the L^p-norm. Suppose p and q are related by

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{C.1.10}$$

Then, if $f \in L^p(X,\mu)$,

$$\|f\|_{L^p} = \sup\,\{\|fh\|_{L^1} : h \in L^q(X,\mu),\ \|h\|_{L^q} \le 1\}. \tag{C.1.11}$$

We can apply this to $f+g$, which belongs to $L^p(X,\mu)$ if f and g do, since $|f+g|^p \le 2^p(|f|^p+|g|^p)$. Given this, (C.1.9) follows easily from the inequality $\|(f+g)h\|_{L^1} \le \|fh\|_{L^1}+\|gh\|_{L^1}$. We again refer to standard measure theory texts for a proof of (C.1.11). Details on the case $L^2(X,\mu)$, which is a Hilbert space, are given in the following section.

There are occasions where it is important to know that one space of functions is a dense linear subspace of another. We mention a couple of particularly significant cases here.

Proposition C.1.1. *If μ is a finite Borel measure on a compact metric space X, then $C(X)$ is dense in $L^p(X,\mu)$ for each $p \in [1,\infty)$.*

The following result is known as the Stone-Weierstrass theorem.

Theorem C.1.2. *Let X be a compact metric space and $\mathcal{A}$ a subalgebra of $C_{\mathbb{R}}(X)$, the algebra of real-valued continuous functions on X. Suppose that $1 \in \mathcal{A}$ and that $\mathcal{A}$ separates the points of X, that is, given distinct $p, q \in X$, there exists $h_{pq} \in \mathcal{A}$ such that $h_{pq}(p) \neq h_{pq}(q)$. Then $\mathcal{A}$ is dense in $C_{\mathbb{R}}(X)$.*

If instead $\mathcal{A}$ is such a subalgebra of $C_{\mathbb{C}}(X)$ and if also $u \in \mathcal{A} \Rightarrow \overline{u} \in \mathcal{A}$, then $\mathcal{A}$ is dense in $C_{\mathbb{C}}(X)$.

Again, proofs can be found in [**41**].

C.2. Hilbert spaces

A Hilbert space H is a Banach space equipped with an inner product, i.e., an assignment $(u,v) \in \mathbb{C}$ to each $u,v \in H$, satisfying

$$\begin{aligned}
&(a_1u_1 + a_2u_2, v) = a_1(u_1,v) + a_2(u_2,v), \\
&\qquad\qquad\qquad\qquad \forall\, a_j \in \mathbb{C},\ u_j, v \in H, \\
&(u,v) = \overline{(v,u)}, \quad \forall\, u,v, \in H, \\
&(u,u) > 0, \quad \text{unless} \quad u = 0.
\end{aligned} \tag{C.2.1}$$

In such a case we define a norm on H by

$$\|u\|^2 = (u,u). \tag{C.2.2}$$

We need to verify that this norm satisfies the triangle inequality,

$$\|u+v\| \le \|u\| + \|v\|, \quad \forall\, u,v \in H. \tag{C.2.3}$$

To see this, square each side,

$$\begin{aligned}
\|u+v\|^2 &= (u+v, u+v) \\
&= \|u\|^2 + \|v\|^2 + 2\operatorname{Re}(u,v), \\
(\|u\| + \|v\|)^2 &= \|u\|^2 + \|v\|^2 + 2\|u\| \cdot \|v\|.
\end{aligned} \tag{C.2.4}$$

The desired inequality (C.2.3) then results from the following, known as Cauchy's inequality:

Lemma C.2.1. *For $u,v \in H$,*

$$\operatorname{Re}(u,v) \le \|u\| \cdot \|v\|. \tag{C.2.5}$$

Proof. We start with

$$0 \le \|u-v\|^2 = (u-v, u-v) = \|u\|^2 + \|v\|^2 - 2\operatorname{Re}(u,v), \tag{C.2.6}$$

obtaining

$$2\operatorname{Re}(u,v) \le \|u\|^2 + \|v\|^2. \tag{C.2.7}$$

Then taking $u \mapsto tu,\ v \mapsto t^{-1}v$ yields

$$2\operatorname{Re}(u,v) \le t^2\|u\|^2 + t^{-2}\|v\|^2, \quad \forall t \in (0,\infty). \tag{C.2.8}$$

Provided $u \neq 0$ and $v \neq 0$, we can take

$$t^2 = \frac{\|v\|}{\|u\|}, \tag{C.2.9}$$

and obtain the desired estimate (C.2.5). Note that replacing u by $e^{i\theta}u$ for appropriate $\theta \in \mathbb{R}$ yields

$$|(u,v)| \le \|u\| \cdot \|v\|. \tag{C.2.10}$$

$\square$

Typically, if $f, g \in H$, we say

$$f \perp g \Longleftrightarrow (f, g) = 0, \tag{C.2.11}$$

and note that

$$\begin{aligned} &\|f+g\|^2 = \|f\|^2 + \|g\|^2 + 2\,\mathrm{Re}(f,g) \\ &\Rightarrow \|f+g\|^2 = \|f\|^2 + \|g\|^2, \quad \text{if } f \perp g. \end{aligned} \tag{C.2.12}$$

The main result we want to discuss next is the concept of an orthonormal basis of H. To formulate this, let $\{u_k : k \in \mathbb{N}\}$ be an orthonormal set, so

$$(u_j, u_k) = \delta_{jk}. \tag{C.2.13}$$

Set

$$\mathcal{L} = \mathrm{Span}\{u_k : k \in \mathbb{N}\}, \tag{C.2.14}$$

i.e., $\mathcal{L}$ is the set of finite linear combinations of these elements. We say $\{u_k\}$ is an orthonormal basis of H if $\mathcal{L}$ is dense in H. Suppose this holds. Given $v \in H$, we set

$$S_N v = \sum_{k=1}^{N} \hat{v}(k) u_k, \quad \hat{v}(k) = (v, u_k). \tag{C.2.15}$$

We can write

$$v = S_N v + R_N v, \tag{C.2.16}$$

and note that $(S_N v, u_k) = (v, u_k)$ for $k \le N$, hence $R_N v \perp u_k$ for $k \le N$, so

$$R_N v \perp S_N v, \quad \text{i.e.,} \quad (S_N v, R_N v) = 0. \tag{C.2.17}$$

It follows that, for each N,

$$\|v\|^2 = \|S_N v\|^2 + \|R_N v\|^2. \tag{C.2.18}$$

In particular, $\|S_N v\| \le \|v\|$, $\forall\, v \in H$. With this result in hand, we establish the following.

Proposition C.2.2. *If $\{u_k\}$ is an orthonormal basis of H, then*

$$S_N v \longrightarrow v, \ \ \textit{in } H\textit{-norm, as } N \to \infty, \ \forall\, v \in H. \tag{C.2.19}$$

Proof. Clearly if $w \in \mathcal{L}$, then $S_N w = w$ for all sufficiently large N, so

$$w \in \mathcal{L} \Longrightarrow S_N w \to w, \quad \text{in } H\text{-norm}. \tag{C.2.20}$$

Now, given $v \in H$, take $\varepsilon > 0$ and pick $w \in \mathcal{L}$ such that $\|w - v\| < \varepsilon$. Then

$$v - S_N v = v - w + w - S_N w + S_N w - S_N v, \tag{C.2.21}$$

so

$$\begin{aligned} \|v - S_N v\| &\le \|v - w\| + \|w - S_N w\| + \|S_N(w - v)\| \\ &\le 2\|v - w\| + \|w - S_N w\|, \end{aligned} \tag{C.2.22}$$

hence

$$\limsup_{N\to\infty} \|v - S_N v\| \le 2\varepsilon, \quad \forall\, \varepsilon > 0, \tag{C.2.23}$$

which gives the desired result (C.2.19). □

We pursue this line a little further. Now let $\{u_k : k \in \mathbb{N}\}$ be an orthonormal set in H, and define $\mathcal{L}$ as in (C.2.14), but do not assume $\mathcal{L}$ is dense in H. Say

$$\overline{\mathcal{L}} = H_0, \tag{C.2.24}$$

a closed linear subspace of H. Again define $S_N v$ as in (C.2.15). We still have (C.2.16)–(C.2.18). Let us write the conclusion $\|S_N v\| \le \|v\|$ as

$$\sum_{k=1}^{N} |\hat{v}(k)|^2 = \|S_N v\|^2 \le \|v\|^2. \tag{C.2.25}$$

It follows that

$$\sum_{k=1}^{\infty} |\hat{v}(k)|^2 < \infty, \quad \text{hence} \quad \sum_{k=M}^{M+N} |\hat{v}(k)|^2 \to 0, \ \text{as } M \to \infty. \tag{C.2.26}$$

so

$$\|S_{M+N} v - S_M v\|^2 \longrightarrow 0, \quad \text{as} \quad M \to \infty, \tag{C.2.27}$$

in which case, for each $v \in H$, $(S_N v)$ is Cauchy in H. Thus this sequence converges:

$$S_N v \longrightarrow Pv \ \text{ in } H\text{-norm, as } \ N \to \infty, \ \forall\, v \in H, \tag{C.2.28}$$

this defining Pv. One checks that $P : H \to H$ is linear, that

$$P : H \longrightarrow H_0, \quad \|Pv\| \le \|v\|, \quad \forall\, v \in H. \tag{C.2.29}$$

It follows from Proposition C.2.2 that

$$v \in H_0 \Longrightarrow Pv = v. \tag{C.2.30}$$

Furthermore, with

$$v = Pv + Rv, \tag{C.2.31}$$

we see that $(Pv, u_k) = (v, u_k)$ for all k, hence $(Rv, u_k) = 0$ for all k, so

$$Rv \in H_1 = H_0^{\perp}, \quad \forall\, v \in H, \tag{C.2.32}$$

where

$$H_0^{\perp} = \{w \in H : w \perp v, \ \forall\, v \in H_0\}. \tag{C.2.33}$$

This gives an orthogonal decomposition,

$$v = v_0 + v_1, \quad v_0 \in H_0, \ v_1 \in H_1, \ \forall\, v \in H. \tag{C.2.34}$$

Such a decomposition is unique. Indeed, if also $v = v_0' + v_1'$, $v_j' \in H_j$, then $0 = (v_0 - v_0') + (v_1 - v_1')$, hence $v_j - v_j' \in H_0 \cap H_0^\perp = 0$. We say P is the orthogonal projection of H onto H_0. We record the result.

Proposition C.2.3. *If $\{u_k : k \in \mathbb{N}\}$ is an orthonormal set in H, $\mathcal{L} = \operatorname{Span}\{u_k\}$, and $H_0 = \overline{\mathcal{L}}$, and if $S_N v$ is defined as in* (C.2.15)*, then, for all $v \in H$, $S_N v \to Pv$ in H-norm, as $N \to \infty$, where P is the orthogonal projection of H onto H_0.*

We next discuss the existence of an orthonormal basis for each closed linear subspace H_0 of a separable Hilbert space H. Such a space H_0 is also separable, so there is a countable dense subset $\{x_k : k \in \mathbb{N}\}$ of H_0. Use this to form a subset $\{y_k : k \in \mathbb{N}\}$ obtained from $\{x_k\}$ by throwing away each x_j that is in the span of $\{x_\ell : \ell < j\}$. Then we have the spaces

$$V_k = \operatorname{Span}\{y_j : j \le k\}, \quad \dim V_k = k, \tag{C.2.35}$$

and we can use the Gram-Schmidt process to construct an orthonormal set $\{u_k : k \in \mathbb{N}\}$ such that

$$V_k = \operatorname{Span}\{u_1, \dots, u_k\}. \tag{C.2.36}$$

Then $\mathcal{L} = \operatorname{Span}\{u_k : k \in \mathbb{N}\}$ contains $\{x_k\}$, so $\overline{\mathcal{L}} = H_0$, and we are in the setting of Proposition C.2.3. This yields the following.

Corollary C.2.4. *If H is a separable Hilbert space and H_0 a closed linear subspace, there is a (uniquely defined) orthogonal projection*

$$P : H \longrightarrow H_0, \tag{C.2.37}$$

yielding the orthogonal decomposition

$$H = H_0 \oplus H_1, \quad H_1 = H_0^\perp. \tag{C.2.38}$$

Using this, we can establish the following result, known as the Riesz representation theorem for Hilbert space.

Proposition C.2.5. *Let H be a separable Hilbert space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$), and let*

$$\varphi : H \longrightarrow \mathbb{F} \tag{C.2.39}$$

be a continuous linear map. Then there is a unique $w \in H$ such that

$$\varphi(v) = (v, w), \quad \forall\, v \in H. \tag{C.2.40}$$

Proof. If $\varphi = 0$, $w = 0$. Assume $\varphi \ne 0$. Let $H_0 = \operatorname{Ker} \varphi = \{v \in H : \varphi(v) = 0\}$, which is a closed linear subspace of H, and use (C.2.38). Then

$$\varphi : H_1 \longrightarrow \mathbb{F} \tag{C.2.41}$$

is injective and nonzero, hence an isomorphism (so $\dim H_1 = 1$). Take a unit vector $w_1 \in H_1$. Then, for $w \in H_1$,

$$(w_1, w) = \varphi(w_1) \Longleftrightarrow w = \overline{\varphi(w_1)} w_1. \tag{C.2.42}$$

For such w, (C.2.40) holds for $v \in H_0$ and for $v \in H_1$, hence for all $v \in H$. Uniqueness is readily checked. $\square$

C.3. Linear operators

If V and W are Banach spaces, $\mathcal{L}(V,W)$ denotes the space of linear transformations $T : V \to W$ that are continuous. Continuity is equivalent to $T^{-1}(\{w \in W : \|w\| < 1\})$ being a neighborhood of 0 in V, hence to the existence of a constant $C < \infty$ such that

$$\|Tv\| \le C\|v\|, \quad \forall\, v \in V. \tag{C.3.1}$$

For this reason, we also call an element of $\mathcal{L}(V,W)$ a bounded linear operator. We define the norm of T to be

$$\|T\| = \sup\{\|Tv\| : \|v\| \le 1\}. \tag{C.3.2}$$

Then (C.3.1) holds with $C = \|T\|$, and $\|T\|$ is the smallest constant for which (C.3.1) holds.

It is clear that $\mathcal{L}(V,W)$ is a linear space. If $T_j \in \mathcal{L}(V,W)$, then

$$\|T_1 + T_2\| \le \|T_1\| + \|T_2\|. \tag{C.3.3}$$

Completeness is also easy to establish, so $\mathcal{L}(V,W)$ is a Banach space. If X is a third Banach space and $S \in \mathcal{L}(W,X)$, then $ST \in \mathcal{L}(V,X)$, and

$$\|ST\| \le \|S\| \cdot \|T\|. \tag{C.3.4}$$

The space $\mathcal{L}(V)$ is a Banach algebra for each Banach space V. Generally a Banach algebra is defined to be a Banach space B, with the structure of an algebra, so that, for each $S, T \in B$, the inequality (C.3.4) holds. Another example of a Banach algebra is the space $C(X)$, with the sup norm, for compact X. Still another, introduced in Chapter 2, is $L^1(G)$, with the convolution product

$$u * v(x) = \int_G u(g)v(g^{-1}x)\, dg, \tag{C.3.5}$$

when G is a Lie group, equipped with Haar measure.

If V is a Banach space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$), we set

$$V' = \mathcal{L}(V,\mathbb{F}), \tag{C.3.6}$$

and call V' the dual of V. We also use the notation

$$\langle v, \omega\rangle = \omega(v), \quad v \in V,\ \omega \in V'. \tag{C.3.7}$$

If $T \in \mathcal{L}(V,W)$, we have

$$T' \in \mathcal{L}(W', V'), \tag{C.3.8}$$

defined by $T'\omega = \omega T : V \to \mathbb{F}$, i.e.,

$$\langle v, T'\omega\rangle = \langle Tv, \omega\rangle, \quad v \in V,\ \omega \in W'. \tag{C.3.9}$$

A special case of (C.3.4) is that $\|T'\omega\| \le \|\omega\| \cdot \|T\|$, hence $\|T'\| \le \|T\|$. In fact, $\|T'\| = \|T\|$. The proof uses a result called the Hahn-Banach theorem, and can be found in Appendix A of [**39**].

An element $T \in \mathcal{L}(V, W)$ is said to be an isometry if $\|Tv\| = \|v\|$ for all $v \in V$. An isometry is always injective. If T is also surjective, we say it is an isomorphic isomorphism. For an example, we note from (C.1.11) that if p and q are related by (C.1.10) (we say $q = p'$ and $p = q'$), then the pairing

$$\langle h, f\rangle = \int_X f(x)h(x)\,d\mu(x) \tag{C.3.10}$$

satisfies

$$f \in L^p(X,\mu) \Rightarrow \|f\|_{L^p} = \sup\,\{|\langle h, f\rangle| : \|h\|_{L^q} \le 1\}, \tag{C.3.11}$$

and hence yields an isometry

$$\iota : L^p(X,\mu) \longrightarrow L^q(X,\mu)'. \tag{C.3.12}$$

It is an important fact that, if (X,μ) is σ-finite,

$$\begin{gathered} 1 \le q < \infty \Rightarrow \iota \text{ is an isometric isomorphism, so} \\ L^q(X,\mu)' \approx L^p(X,\mu). \end{gathered} \tag{C.3.13}$$

This is proved in standard measure theory texts, such as [**41**]. For $p = 2$, it is a simple consequence of the Hilbert space theory presented in the previous section.

In case V and W are Hilbert spaces and $T \in \mathcal{L}(V, W)$, then we also have an adjoint $T^* \in \mathcal{L}(W, V)$, given by

$$(Tv, w) = (v, T^*w), \quad v \in V,\ w \in W, \tag{C.3.14}$$

using the inner products on W and V, respectively. In this case, it it is elementary that

$$\|T\| = \sup\,\{|(Tv, w)| : \|v\|,\ \|w\| \le 1\} = \|T^*\|. \tag{C.3.15}$$

One also has $T^{**} = T$.

When H is a Hilbert space, the Banach algebra $\mathcal{L}(H)$ is a C^*-algebra. Generally a C^*-algebra B is a Banach algebra, equipped with a conjugate linear involution $T \mapsto T^*$, satisfying $\|T^*\| = \|T\|$ and

$$\|T^*T\| = \|T\|^2. \tag{C.3.16}$$

To see that (C.3.16) holds for $T \in \mathcal{L}(H)$, note that both sides are equal to the sup over $\|v_1\|, \|v_2\| \le 1$, of the absolute value of

$$(T^*Tv_1, v_2) = (Tv_1, Tv_2), \tag{C.3.17}$$

such a supremum necessarily being obtained over the set of pairs satisfying $v_1 = v_2$. Note that $C(X)$, considered above, is also a C^*-algebra. However, for a general Banach space V, $\mathcal{L}(V)$ will not have the structure of a C^*-algebra.

It is of interest to know when a sequence $T_k \in \mathcal{L}(V, W)$ has the convergence property $T_k v \to Tv$, $\forall\, v \in V$, with $T \in \mathcal{L}(V, W)$. An example of this has appeared in §C.2. Here is a general result.

Proposition C.3.1. *Let $T_k, T \in \mathcal{L}(V, W)$, and assume there exists $C < \infty$ such that $\|T_k\| \le C$ for all k, and $\|T\| \le C$. Let $\mathcal{L} \subset V$ be a dense linear subspace, and assume*

$$T_k v \longrightarrow Tv, \quad \text{in } W\text{-norm}, \tag{C.3.18}$$

for all $v \in \mathcal{L}$. Then (C.3.18) *holds for all $v \in V$.*

Proof. Take $v \in V$. Pick $\varepsilon > 0$, and then pick $w \in \mathcal{L}$ such that $\|v - w\| < \varepsilon$. Then

$$Tv - T_k v = Tv - Tw + Tw - T_k w + T_k w - T_k v, \tag{C.3.19}$$

so

$$\begin{aligned} \|Tv - T_k v\| &\le \|T(v - w)\| + \|Tw - T_k w\| + \|T_k(v - w)\| \\ &\le 2C\|v - w\| + \|Tw - T_k w\|. \end{aligned} \tag{C.3.20}$$

Hence

$$\limsup_{k\to\infty} \|Tv - T_k v\| \le 2C\varepsilon, \quad \forall\, \varepsilon > 0, \tag{C.3.21}$$

so the limsup is 0, and we have (C.3.18) for all $v \in V$. □

The following is a simple but useful strengthening.

Proposition C.3.2. *Let $T_k \in \mathcal{L}(V, W)$ and assume there exists $C < \infty$ such that $\|T_k\| \le C$ for all k. Let $\mathcal{L} \subset V$ be a dense linear subspace, and assume*

$$T_k v \ \text{converges, in } W\text{-norm}, \quad \forall\, v \in \mathcal{L}. \tag{C.3.22}$$

Then there exists $T \in \mathcal{L}(V, W)$ such that $\|T\| \le C$ and (C.3.18) *holds for all $v \in V$.*

Proof. As above, take $v \in V$, pick $\varepsilon > 0$, and then pick $w \in \mathcal{L}$ such that $\|v - w\| < \varepsilon$. Replacing (C.3.19) by

$$T_j v - T_k v = T_j v - T_j w + T_j w - T_k w + T_k w - T_k v, \tag{C.3.23}$$

we deduce as above that

$$(T_k v) \ \text{ is Cauchy in } W, \quad \forall\, v \in V. \tag{C.3.24}$$

Denote the limit by Tv. One then checks that $T \in \mathcal{L}(V, W)$, $\|T\| \le C$, and (C.3.18) holds. □

The following application of Proposition C.3.2 will play an important role in the proof of Proposition C.4.4.

Proposition C.3.3. *Let H be a separable Hilbert space. Assume $w_k \in H$ and $\|w_k\| \le 1$. Then there exists $w \in H, \|w\| \le 1$, and a subsequence (w_{k_ν}) such that*

$$(v, w_{k_\nu}) \longrightarrow (v, w), \quad \forall\, v \in H. \tag{C.3.25}$$

Proof. Let $\{u_n : n \in \mathbb{N}\}$ be an orthonormal basis of H. Define $T_k : H \to \mathbb{F}$ by $T_k v = (v, w_k)$, so $\|T_k\| \le 1$. Since $\{T_k u_1 : k \in \mathbb{N}\}$ is a bounded set in $\mathbb{F}$, there is a subsequence $T_k^{(1)}$ such that $T_k^{(1)} u_1$ converges. In turn $T_k^{(1)}$ has a further subsequence $T_k^{(2)}$ such that $T_k^{(2)} u_2$ converges, and so on, yielding nested subsequences $T_k^{(\ell)}$ such that $T_k^{(\ell)} u_j$ converges as $k \to \infty$, for each $j \in \{1, \dots, \ell\}$. Then a diagonal argument produces a subsequence T_{k_ν} such that $T_{k_\nu} u_j$ converges as $k_\nu \to \infty$, for each $j \in \mathbb{N}$.

Now Proposition C.3.2 applies, with $W = \mathbb{F}$ and $\mathcal{L} = \operatorname{Span}\{u_n : n \in \mathbb{N}\}$. We have $T \in \mathcal{L}(H, \mathbb{F})$ such that

$$T_{k_\nu} v \longrightarrow T v, \quad \forall\, v \in H. \tag{C.3.26}$$

Also $\|T\| \le 1$. Finally, we have $w \in H$ such that $Tv = (v, w)$ for all $v \in H$, and then (C.3.25) holds. $\square$

Remark. When (C.3.25) holds, we say $w_{k_\nu} \to w$ weakly in H.

C.4. Compact operators

If V and W are Banach spaces, an operator $T \in \mathcal{L}(V, W)$ is said to be compact provided T takes each bounded subset of V to a relatively compact subset of W, that is, a set with compact closure. It suffices to assume that $T(B_1)$ is relatively compact in W, where B_1 is the closed unit ball in V. We denote the space of compact operators by $\mathcal{K}(V, W)$. If $V = W$, we use the notation $\mathcal{K}(V)$.

The following proposition summarizes some elementary facts about $\mathcal{K}(V, W)$.

Proposition C.4.1. *$\mathcal{K}(V, W)$ is a closed linear subspace of $\mathcal{L}(V, W)$. Each T in $\mathcal{L}(V, W)$ with finite-dimensional range is compact. Furthermore, if $T \in \mathcal{K}(V, W)$, $S_1 \in \mathcal{L}(V_1, V)$, and $S_2 \in \mathcal{L}(W, W_2)$, then $S_2TS_1 \in \mathcal{K}(V_1, W_2)$.*

Most of these assertions are obvious. We show that if $T_j \in \mathcal{K}(V, W)$ is norm convergent to T, then T is compact. Given a sequence (x_n) in B_1, one can pick successive subsequences on which T_1x_n converges, then T_2x_n converges, and so on, and by a diagonal argument produce a single subsequence (which we will still denote (x_n)) such that for each j, T_jx_n converges as $n \to \infty$. It is then easy to show that Tx_n converges, giving compactness of T.

Here is a useful class of compact operators.

Proposition C.4.2. *If X is a compact metric space, then the natural inclusion*

$$\iota : \mathrm{Lip}(X) \longrightarrow C(X) \tag{C.4.1}$$

is compact.

This is a special case of Ascoli's theorem. More generally, let $\omega : X \times X \to [0, \infty)$ be any continuous function, vanishing on the diagonal $\Delta = \{(x, x) : x \in X\}$. Fix $K \in \mathbb{R}^+$. Let $\mathcal{F}$ be any subset of $C(X)$ satisfying

$$|u(x)| \le K, \quad |u(x) - u(y)| \le K\omega(x, y). \tag{C.4.2}$$

This condition is called equicontinuity. Ascoli's theorem states that such a set is relatively compact in $C(X)$ whenever X is a compact metric space. See, for example, Appendix A of [**39**] for a proof of Proposition C.4.2.

One useful implication is the persistence of compactness under taking adjoints.

Proposition C.4.3. *If $T \in \mathcal{K}(V, W)$, then T' is also compact.*

Proof. Let (ω_n) be a sequence in B_1', the closed unit ball in W'. Consider (ω_n) as a sequence of continuus functions on the compact space $X =$

$\overline{T(B_1)}$, B_1 being the unit ball in V. Proposition C.4.2 applies; there exists a subsequence (ω_{n_k}) converging uniformly on X. Thus $(T'\omega_{n_k})$ is a sequence in V' converging uniformly on B_1, hence in the V'-norm. This completes the proof. □

For simplicity, we restrict attention in the next result to operators in $\mathcal{K}(H)$, where H is a separable Hilbert space. A more general result is given in Appendix A of [**39**].

Proposition C.4.4. *Let H be a separable Hilbert space. If $T \in \mathcal{K}(H)$, then the image of the closed unit ball $B \subset H$ under T is compact.*

Proof. Take $v_n \in B$. Since we are given that $T(B)$ is relatively compact in H, a subsequence Kv_n converges, to a limit $w \in H$. Our task is to show that $w \in K(B)$.

Now, by Proposition C.3.3, we have a subsequence $v_n \to v$ weakly, $v \in B$. Consequently, for all $x \in H$,

$$(Kv_n, x) = (v_n, K^*x) \longrightarrow (v, K^*x) = (Kv, x), \tag{C.4.3}$$

so $Kv_n \to Kv$ weakly. Since $Kv_n \to w$ in norm, we conclude that $w = Kv$, as desired. □

We next derive some results on the spectral theory of a compact operator on a separable Hilbert space H that is self adjoint, so $A = A^*$.

Proposition C.4.5. *If H is a separable Hilbert space and $A \in \mathcal{L}(H)$ is compact and self adjoint, then either $\|A\|$ or $-\|A\|$ is an eigenvalue of S, that is, there exists $u \neq 0$ such that*

$$Au = \lambda u, \tag{C.4.4}$$

with $\lambda = \pm\|A\|$.

Proof. By Proposition C.4.4, we know that the image under A of the closed unit ball $B \subset H$ is compact, so the norm assumes a maximum on this image. Thus there exists $u \in H$ such that

$$\|u\| = 1, \quad \|Au\| = \|A\|. \tag{C.4.5}$$

Pick any unit $w \perp u$. Self-adjointness implies $\|Ax\|^2 = (A^2x, x)$, so we have, for all real s,

$$(A^2(u + sw), u + sw) \le \|A\|^2(1 + s^2), \tag{C.4.6}$$

equality holding at $s = 0$. Since the left side is equal to

$$\|A\|^2 + 2s\,\mathrm{Re}(A^2u, u) + s^2\|Aw\|^2, \tag{C.4.7}$$

this inequality for $s \to 0$ implies $\mathrm{Re}(A^2u, w) = 0$. Replacing w by iw gives $(A^2u, w) = 0$ whenever $w \perp u$. Thus A^2u is parallel to u, that is, $A^2u = cu$ for some scalar c, and (C.4.5) implies $c = \|A\|^2$. Now, assuming $A \neq 0$, set

$$v = (A + \|A\|I)u.$$

If $v = 0$, then u satisfies (C.4.4) with $\lambda = -\|A\|$. If $v \neq 0$, then v is an eigenvector of A with eigenvalue $\lambda = \|A\|$. $\square$

The space of $u \in H$ satisfying (C.4.4) is called the λ-eigenspace of A. Clearly, if A is compact and $\lambda \neq 0$, such a λ-eigenspace must be finite dimensional. If $Au_j = \lambda_j u_j,\ A = A^*$, then

$$\lambda_1(u_1, u_2) = (Au_1, u_2) = (u_1, Au_2) = \overline{\lambda}_2(u_1, u_2). \tag{C.4.8}$$

With $\lambda_1 = \lambda_2$ and $u_1 = u_2$, this implies that each eigenvalue of $A = A^*$ is real. With $\lambda_1 \neq \lambda_2$, it then yields $(u_1, u_2) = 0$, so any distinct eigenspaces of $A = A^*$ are orthogonal.

We also note that if $H_0 \subset H$ is a closed linear subspace,

$$A : H_0 \to H_0 \Longrightarrow A : H_0^\perp \to H_0^\perp. \tag{C.4.9}$$

Indeed, we have

$$(u, Av) = (Au, v) = 0, \quad \forall\, u \in H_0,\ v \in H_0^\perp.$$

Now if A is compact and self adjoint on H, we can apply Proposition C.4.5, restrict A to the orthogonal complement of its $\pm\|A\|$-eigenspaces (where its norm must be strictly smaller, as a consequence of Proposition C.4.5), apply the proposition again, and continue. In this fashion we obtain mutually orthogonal finite-dimensional spaces $H_j \subset H$, sums of μ_j and $-\mu_j$-eigenspaces of A, with

$$\|A\| = \mu_1 > \mu_2 > \cdots, \tag{C.4.10}$$

and

$$A : E_k \to E_k, \quad E_k = \Big(\bigoplus_{j \le k} H_j\Big)^\perp, \quad \|A|_{E_k}\| < \mu_k. \tag{C.4.11}$$

We claim $\mu_j \searrow 0$ is H is infinite dimensional. Indeed, if all $\mu_j \ge a > 0$, then picking one unit vector v_j in each H_j gives an orthonormal set $\{v_j\} \subset H$ such that, for $j \neq k$,

$$\|Av_j - Av_k\|^2 = \mu_j^2 + \mu_k^2 \ge 2a^2,$$

contradicting compactness of A. We have the following result, known as the spectral theorem for compact, self-adjoint operators.

Proposition C.4.6. *If $A \in \mathcal{L}(H)$ is a compact, self-adjoint operator on a Hilbert space H, then H has an orthonormal basis u_j of eigenvectors of A. With $Au_j = \lambda_j u_j$, (λ_j) is a sequence of real numbers with only 0 as an accumulation point.*

There is also a spectral theorem for general bounded self-adjoint operators, but for the purpose of these notes the compact case suffices. The general result is teated in Chapter 8 of [**39**].

Appendix D

Positive definite zonal functions

A function u on the unit sphere $S^{n-1} \subset \mathbb{R}^n$ is said to be a zonal function if u is a function of x_n alone. Such a function can be expanded in zonal harmonics. In more detail, we have

$$L^2(S^{n-1}) = \bigoplus_{\ell \geq 0} V_\ell, \tag{D.0.1}$$

where each V_ℓ is an eigenspace of the Laplace-Beltrami operator Δ on S^{n-1}:

$$V_\ell = \{f \in C^\infty(S^{n-1}) : \Delta f = -\lambda_\ell^2 f\}, \tag{D.0.2}$$

where $\{-\lambda_\ell^2 : \ell \geq 0\} = \operatorname{Spec} \Delta$, i.e., $\lambda_\ell^2 = \ell(\ell + n - 2)$. Each space V_ℓ has a one-dimensional subspace of zonal functions, spanned by

$$\mathfrak{z}_\ell(x) = A_{\ell n} C_\ell^\alpha(x_n), \quad \alpha = \frac{n-2}{2}, \tag{D.0.3}$$

where C_ℓ^α is a Gegenbauer polynomial, given by the generating function

$$(1 - 2tr + r^2)^{-\alpha} = \sum_{\ell=0}^{\infty} C_\ell^\alpha(t) r^\ell. \tag{D.0.4}$$

(We assume $n \geq 3$.) See Chapter 8, §§8.3–8.4. Taking $t = 1$ gives $(1-r)^{-2\alpha}$ for the left side, and we see that

$$C_\ell^\alpha(1) > 0, \tag{D.0.5}$$

for $\alpha > 0,\ \ell \geq 0$. We set $A_{\ell n} = 1/C_\ell^\alpha(1)$ (which is > 0) in (D.0.3), so

$$\mathfrak{z}_\ell(p_0) = 1, \tag{D.0.6}$$

where $p_0 = (0, \ldots, 0, 1)$ is the "north pole" of S^{n-1}.

We are interested in zonal functions of the form

$$u(x) = \sum_{\ell \geq 0} c_\ell \mathfrak{z}_\ell(x), \quad c_\ell \geq 0. \tag{D.0.7}$$

If $u \in C(S^{n-1})$ has the form (1.7), we say

$$u \in \mathcal{PZ}(S^{n-1}). \tag{D.0.8}$$

At first glance, one might think the normalization $\mathfrak{z}(p_0) > 0$ is pretty arbitrary, and that the class of functions given by (D.0.7) is not particularly distinguished. However, as we will see, this is far from the case.

In fact $\mathcal{PZ}(S^{n-1})$ is naturally equivalent to the space of K-bi-invariant elements of $C(G)$ that are *positive definite*, as functions on the compact group G, where

$$G = SO(n), \quad K = SO(n-1). \tag{D.0.9}$$

This natural class of functions on G is independent of any choice of normalized eigenfunctions. Further general properties of positive-definite functions imply that

$$u, v \in \mathcal{PZ}(S^{n-1}) \Longrightarrow uv \in \mathcal{PZ}(S^{n-1}). \tag{D.0.10}$$

Furthermore, the class $\mathcal{Z}(S^{n-1})$ of zonal functions on S^{n-1} has a natural convolution product, and

$$u, v \in \mathcal{PZ}(S^{n-1}) \Longrightarrow u * v \in \mathcal{PZ}(S^{n-1}). \tag{D.0.11}$$

In §D.1 we gather some facts about positive-definite functions. We work more generally on a compact Lie group G, and replace S^{n-1} by $M = G/K$, where K is a closed subgroup of G. In §§D.2–D.3 we define $\mathcal{PZ}(M)$ and identify it with the class of K-bi-invariant functions on G that are positive definite. We establish (D.0.10), in the more general setting of $\mathcal{PZ}(M)$. For (D.0.11), we require M to be a rank-one symmetric space.

D.1. Positive definite functions on G

Let G be a compact Lie group. A function $u \in C(G)$ is said to be positive definite provided

$$L(u) : L^2(G) \longrightarrow L^2(G), \tag{D.1.1}$$

defined as left convolution by u:

$$L(u)f(x) = u * f(x) = \int_G u(y)f(y^{-1}x)\,dy, \quad f \in L^2(G), \tag{D.1.2}$$

is a positive, self-adjoint operator. An equivalent condition is that, for each strongly continuous unitary representation π of G on a Hilbert space H, $\pi(u) : H \to H$ is a positive, self-adjoint operator, where

$$\pi(u)f = \int_G u(y)\pi(y)f\,dy, \quad f \in H. \tag{D.1.3}$$

Note that in (D.1.2), $L(y)f(x) = f(y^{-1}x)$ is the left regular representation of G on $L^2(G)$. Note also that

$$\begin{aligned} L(u)f(x) &= \int_G u(xy^{-1})f(y)\,dy \\ &= \int_G K_u(x,y)f(y)\,dy, \end{aligned} \tag{D.1.4}$$

where

$$K_u(x,y) = u(xy^{-1}). \tag{D.1.5}$$

Hence an alternative characterization is that $u \in C(G)$ is positive definite if and only if, for each finite set $\{g_1, \dots, g_N\} \subset G$, the $N \times N$ matrix

$$A = (a_{jk}) = \big(u(g_j g_k^{-1})\big) \text{ is positive and self-adjoint in } M(N, \mathbb{C}). \tag{D.1.6}$$

We denote the set of positive definite elements of $C(G)$ by $\mathcal{P}(G)$. The following is a consequence of the characterization (D.1.6).

Proposition D.1.1. *We have*

$$u, v \in \mathcal{P}(G) \Longrightarrow uv \in \mathcal{P}(G). \tag{D.1.7}$$

Proof. Take $\{g_1, \dots, g_N\} \subset G$. Parallel to (D.1.6), given $v \in \mathcal{P}(G)$, we have

$$B = \big(v(g_j g_k^{-1})\big) \text{ positive, self-adjoint.} \tag{D.1.8}$$

To establish (D.1.7), it suffices to deduce from (D.1.6) and (D.1.8) that

$$C = (c_{jk}), \quad c_{jk} = a_{jk} b_{jk} \tag{D.1.9}$$

is positive. In fact, (D.1.6) and (D.1.8) imply

$$A \otimes B \geq 0 \ \text{ in } \ M(N^2, \mathbb{C}), \tag{D.1.10}$$

hence

$$\sum a_{jk} b_{j'k'} \eta_{jj'} \overline{\eta}_{kk'} \geq 0, \quad \forall\, \eta_{jj'} \in \mathbb{C}. \tag{D.1.11}$$

Take $\eta_{jj'} = \xi_j \delta_{jj'}$. We get

$$\sum a_{jk} b_{j'k'} \eta_{jj'} \overline{\eta}_{kk'} = \sum a_{jk} b_{jk} \xi_j \overline{\xi}_k, \tag{D.1.12}$$

so (2.11) implies

$$\sum a_{jk} b_{jk} \xi_j \overline{\xi}_k \geq 0, \quad \forall\, \xi_j \in \mathbb{C}, \tag{D.1.13}$$

as desired. □

We also have convolution $u*v$, defined as in (D.1.2), and if π is a strongly continuous unitary representation of G on H,

$$\pi(u * v) = \pi(u)\pi(v). \tag{D.1.14}$$

Now, if $\pi(u)$ and $\pi(v)$ *commute*, we can deduce that $\pi(u*v)$ is positive, self-adjoint, provided $\pi(u)$ and $\pi(v)$ are, but we cannot draw such a conclusion if these factors do not commute. More on this later.

The following ties in with the normalization (D.0.6).

Proposition D.1.2. *Let $e \in G$ denote the identity element. Then*

$$u \in \mathcal{P}(G) \Longrightarrow u(e) \geq |u(x)|, \quad \forall\, x \in G. \tag{D.1.15}$$

Proof. Let $\{\pi_\alpha\}$ be a complete set of irreducible unitary representations of G on V_α, of dimension d_α. We have the following Fourier inversion formula for $u \in L^2(G)$:

$$u(x) = \sum_\alpha d_\alpha \operatorname{Tr}\big(\pi_\alpha(u)\pi_\alpha(x)^*\big), \tag{D.1.16}$$

with convergence in L^2-norm. To check absolute and uniform convergence, note that

$$d_\alpha |\operatorname{Tr}(\pi_\alpha(u)\pi_\alpha(x)^*)| \leq d_\alpha \|\pi_\alpha(u)\|_{TR}. \tag{D.1.17}$$

If $u \in C(G)$ is positive definite, then each $\pi_\alpha(u) \geq 0$, and

$$\|\pi_\alpha(u)\|_{TR} = \operatorname{Tr} \pi_\alpha(u). \tag{D.1.18}$$

In such a case,

$$|u(x)| \leq \sum_\alpha d_\alpha \operatorname{Tr} \pi_\alpha(u) = u(e), \tag{D.1.19}$$

as asserted in (D.1.15). □

D.2. K-bi-invariant functions

Let G be a compact Lie group and K a closed subgroup, and set $M = G/K$, endowed with a G-invariant Riemannian metric. Set

$$p_0 = [e] \in M = G/K. \tag{D.2.1}$$

A function in $C(M)$ corresponds to a function $u \in C(G)$ satisfying $u(xk) = u(x)$, for each $x \in G,\ k \in K$. A function in $C(M)$ invariant under the action of K corresponds to a function $u \in C(G)$ satisfying

$$u(k_1 x k_2) = u(x), \quad \forall\, k_j \in K,\ x \in G. \tag{D.2.2}$$

We say u is K-bi-invariant, or $u \in C(K \setminus G/K)$.

We denote by $\mathcal{Z}(M)$ the space of continuous functions on M invariant under the action of K. We have a natural isomorphism

$$\mathcal{Z}(M) \approx C(K \setminus G/K), \tag{D.2.3}$$

and these spaces are also naturally isomorphic to the space of K-bi-invariant functions in $C(G)$. Note that

$$u, v \in \mathcal{Z}(M) \Longrightarrow uv \in \mathcal{Z}(M). \tag{D.2.4}$$

We say

$$u \in \mathcal{PZ}(M) \tag{D.2.5}$$

provided $u \in \mathcal{Z}(M)$ and its counterpart in $C(G)$ is positive definite. By Proposition D.1.1,

$$u, v \in \mathcal{PZ}(M) \Longrightarrow uv \in \mathcal{PZ}(M). \tag{D.2.6}$$

Recall the convolution product of u and v in $C(G)$:

$$u * v(x) = \int\limits_G u(y) v(y^{-1}x)\, dy. \tag{D.2.7}$$

If u and v are right K-invariant, $u*v$ might not be right K-invariant, except in special cases. However, if u and v are K-bi-invariant, so is $u * v$. This gives rise to a convolution product on $\mathcal{Z}(M)$:

$$u, v \in \mathcal{Z}(M) \Longrightarrow u * v \in \mathcal{Z}(M). \tag{D.2.8}$$

If G is non-commutative, the convolution product on $C(G)$ is non-commutative. As we will see, sometimes the convolution product on $\mathcal{Z}(M)$ is commutative (in particular, this happens when $M = S^{n-1}$).

For general $u \in C(G)$, if π is a unitary representation of G,

$$v(x) = u(k_1 x k_2) \Longrightarrow \pi(v) = \pi(k_1^{-1})\pi(u)\pi(k_2^{-1}). \tag{D.2.9}$$

Hence, if u is K-bi-invariant, then

$$\pi(u) = \pi(k_1)\pi(u)\pi(k_2), \quad \forall\, k_j \in K. \tag{D.2.10}$$

Equivalently, given $u \in C(G)$,

$$\text{(D.2.11)}\qquad \begin{aligned} u \in C(K \setminus G/K) \Longrightarrow \pi(u) &= \pi(u)\pi(k) \\ &= \pi(k)\pi(u), \quad \forall\, k \in K. \end{aligned}$$

Consequently, if $u \in C(K \setminus G/K)$, then, for all $k \in K$,

$$\text{(D.2.12)}\qquad \begin{aligned} &\pi(k) = \text{id. on } \mathcal{R}\pi(u), \text{ and} \\ &\pi(k) = \text{id. on } \mathcal{R}\pi(u)^* = \bigl(\mathcal{N}\pi(u)\bigr)^{\perp}, \end{aligned}$$

where $\mathcal{R}\pi(u)$ is the range of $\pi(u)$ and $\mathcal{N}\pi(u)$ is the null space of $\pi(u)$. If $\pi(u)$ is self-adjoint, the two conditions in (D.2.12) are equivalent to each other.

D.3. Specialization to $M = S^{n-1}$

We now assume G and K are given by (D.0.9), so $M = S^{n-1}$, with its standard metric. G acts on M by rotations, hence as a unitary group on $L^2(M)$, and then it acts on each space V_ℓ in (D.0.1). Call this representation π_ℓ. Since S^{n-1} is a rank-one symmetric space, dim $K \setminus G/K = 1$ and each V_ℓ has a one-dimensional subspace of zonal functions,

$$\mathcal{Z}_\ell = \operatorname{Span}(\mathfrak{z}_\ell), \tag{D.3.1}$$

with $\mathfrak{z}_\ell$ as in (D.0.3) (normalized by (D.0.6). (Hence each π_ℓ is irreducible.) The space $\mathcal{Z}_\ell$ is the subspace of V_ℓ on which $\pi_\ell(k)$ acts as the identity for all $k \in K$. Hence, by (D.2.12), if $u \in \mathcal{Z}(M)$,

$$\pi_\ell(u) : V_\ell \longrightarrow \mathcal{Z}_\ell, \tag{D.3.2}$$

and ditto for $\pi_\ell(u)^*$. Generally, if π is a unitary representation of G and $u \in C(G)$,

$$\pi(u)^* = \pi(u^*), \quad u^*(x) = \overline{u(x^{-1})}. \tag{D.3.3}$$

In this setting, we have

$$\mathfrak{z}_\ell^* = \mathfrak{z}_\ell, \tag{D.3.4}$$

so $\pi_\ell(\mathfrak{z}_\ell)$ is a scalar muitiple of an orthogonal projection:

$$\pi_\ell(\mathfrak{z}_\ell) = \gamma_\ell Z_\ell, \tag{D.3.5}$$

with $\gamma_\ell \in \mathbb{R}$ and

$$Z_\ell = \text{orthogonal projection of } V_\ell \text{ onto Span } (\mathfrak{z}_\ell). \tag{D.3.6}$$

The Weyl orthogonality relations imply that if π_α is an irreducible representation of G,

$$\pi_\alpha \ \text{ not } \ \equiv \pi_\ell \Longrightarrow \pi_\alpha(\mathfrak{z}_\ell) = 0. \tag{D.3.7}$$

In view of the inversion formula (D.1.16), this implies $\pi_\ell(\mathfrak{z}_\ell) \neq 0$, so $\gamma_\ell \neq 0$ in (D.3.5). Whether $\gamma_\ell > 0$ or $\gamma_\ell < 0$, either $\mathfrak{z}_\ell$ or $-\mathfrak{z}_\ell$ is mapped by π_ℓ to a positive operator, so it must be positive definite. In fact, since $\mathfrak{z}(p_0) > 0$, it follows from Proposition D.1.2 that it must be $\mathfrak{z}_\ell$ that is positive definite, so $\gamma_\ell > 0$ and

$$\mathfrak{z}_\ell \text{ is positive definite.} \tag{D.3.8}$$

In fact, via (D.1.19),

$$\mathfrak{z}(p_0) = 1 \Longrightarrow \gamma_\ell = \frac{1}{d_\ell}, \quad d_\ell = \dim V_\ell. \tag{D.3.9}$$

We also deduce from Proposition D.1.2 that

$$|\mathfrak{z}_\ell(x)| \leq 1, \quad \forall\, x \in S^{n-1}. \tag{D.3.10}$$

Note that (D.3.5) extends:

$$u \in \mathcal{Z}(S^{n-1}) \Longrightarrow \pi_\ell(u) = \Gamma_\ell(u) Z_\ell, \quad \Gamma_\ell : \mathcal{Z}(S^{n-1}) \to \mathbb{C}. \tag{D.3.11}$$

This observation enables us to prove the following.

Proposition D.3.1. *The convolution product on $\mathcal{Z}(S^{n-1})$ is commutative:*

$$u, v \in \mathcal{Z}(S^{n-1}) \Longrightarrow u * v = v * u \quad in \quad \mathcal{Z}(S^{n-1}). \tag{D.3.12}$$

Consequently, with $\mathcal{PZ}(M)$ defined as in §D.2,

$$u, v \in \mathcal{PZ}(S^{n-1}) \Longrightarrow u * v \in \mathcal{PZ}(S^{n-1}). \tag{D.3.13}$$

Proof. To get (D.3.12), it suffices to show that

$$u, v \in \mathcal{Z}(S^{n-1}) \Longrightarrow \pi_\alpha(u)\pi_\alpha(v) = \pi_\alpha(v)\pi_\alpha(u), \tag{D.3.14}$$

for each irreducible unitary representation π_α of $= SO(n)$. In fact, either π_α is equivalent to π_ℓ for some ℓ (which always holds if $n = 3$), in which case (4.11) holds, or else, by the Weyl orthogonality relations, $\pi_\alpha(u) = 0$. Since each Z_ℓ is a rank-one projection, (D.3.14) follows. As noted previously, (D.3.13) is a consequence of such commutativity. $\square$

We now show that $\mathcal{PZ}(M)$, defined as in §D.2, coincides with the space characterized by (D.0.7)–(D.0.8) when $M = S^{n-1}$.

Proposition D.3.2. *With $\mathcal{PZ}(M)$ defined as in §D.2, a function $u \in \mathcal{Z}(S^{n-1})$ belongs to $\mathcal{PZ}(S^{n-1})$ if and only if* (D.0.7) *holds.*

Proof. By (D.3.9)–(D.3.10), if (D.0.7) holds and the sum is bounded, then $\sum c_\ell < \infty$, and the sum converges absolutely and uniformly. That such a sum belongs to $\mathcal{PZ}(S^{n-1})$ then follows readily from (D.3.8).

For the converse, if $u \in \mathcal{Z}(S^{n-1})$, we can write

$$u = \sum b_\ell \mathfrak{z}_\ell, \tag{D.3.15}$$

with $b_\ell \in \mathbb{C}$, and convergence in L^2-norm. We need to show that if $u \in \mathcal{PZ}(S^{n-1})$, then each b_ℓ is ≥ 0. In fact, for u as in (D.3.15) and each $\ell \geq 0$, we have

$$\pi_\ell(u) = b_\ell \gamma_\ell Z_\ell. \tag{D.3.16}$$

Since $\gamma_\ell > 0$, we have $b_\ell \geq 0$ when $u \in \mathcal{PZ}(S^{n-1})$. $\square$

We hence have (D.0.10)–(D.0.11).

Remark. While we have desired to deduce (D.0.10) from "general principles," we can also deduce it (in the setting of $\mathcal{PZ}(S^{n-1})$ defined by (D.0.7)–(D.0.8)) from special function identities. In fact, given $\alpha > 0$ and $\ell_j \geq 0$,

$$C^{\alpha}_{\ell_1}(t)C^{\alpha}_{\ell_2}(t) = \sum_{\ell \in \mathcal{S}(\ell_1,\ell_2)} \sigma^{\alpha}_{\ell_1,\ell_2}(\ell) C^{\alpha}_{\ell}(t), \tag{D.3.17}$$

where

$$\mathcal{S}(\ell_1,\ell_2) = \{\ell \in \mathbb{Z} : |\ell_1 - \ell_2| \leq \ell \leq \ell_1 + \ell_2 \text{ and } \ell = \ell_1 + \ell_2 \text{ mod } 2\}, \tag{D.3.18}$$

and

$$\sigma^{\alpha}_{\ell_1,\ell_2}(\ell) > 0 \quad \text{whenever} \quad \ell \in \mathcal{S}(\ell_1,\ell_2). \tag{D.3.19}$$

A formula for these coefficients is given in [**46**], p. 491. For another application of this, see [**32**], §5.

Appendix E

Complementary results

In this appendix we discuss a variety of results on Lie groups and Lie algebras, not so much needed as background for the main text, but related to this material in interesting ways. Section E.1 discusses a class of Lie algebras known as two-step nilpotent Lie algebras, the leading example of which is the Heisenberg algebra. This is germane to quantum theory, and its representation theory plays a significant role in [**38**]. Here, we take a look at automorphism groups of such a Lie algebra $\mathfrak{n}$, and see that every compact Lie group arises as a group of automorphisms of such a Lie algebra. We also discuss some generalizations of this.

Section E.2 defines induced representations, leading from an irreducible representation ρ of a closed subgroup H of a compact G to the induced representation π of G, typically not irreducible. We establish the Frobenius reciprocity theorem, which states that, if λ is an irreducible representation of G, then the number of times π contains λ is equal to the number of times $\lambda\big|_H$ contains ρ.

Section E.3 studies the geometrical properties of a Lie group G, equipped with a bi-invariant Riemannian metric. It is shown that the constant speed geodesics through the identity element coincide with the one parameter subgroups $\gamma_X(t) = \mathrm{Exp}(tX)$, $X \in \mathfrak{g}$. We use this to produce formulas for the Riemann curvature tensor. We bring in a theorem from differential geometry to deduce that if G is such a group, and if the center of $\mathfrak{g}$ is 0 (and G is connected), then G is compact.

Section E.4 gives a brief discussion of how G_2, the smallest of the exceptional compact Lie groups, is related to E_8, the largest and most mysterious.

Section E.5 analyzes integrals of the form

$$\int_{\mathrm{U}(n)} u(g) \otimes v(g)\, dg, \tag{E.0.1}$$

where $u, v \in C(S^1)$ and $u(g), v(g) \in \mathrm{End}(\mathbb{C}^n)$ are defined by the spectral theorem. The trace of such an integral arose in [**8**], and the analysis there made use of work of Dyson [**17**]. The analysis of (E.0.1) has the complicating feature that, unlike its trace, the integrand is not a central function.

E.1. Two-step nilpotent Lie algebras

Every 2-step nilpotent Lie algebra $\mathfrak{n}$ has the form

$$\mathfrak{n} = V \oplus \mathfrak{z}, \tag{E.1.1}$$

as a vector space direct sum, where $\mathfrak{z}$ is central and the Lie bracket on V is uniquely determined by an anti-symmetric bilinear map

$$A : V \times V \longrightarrow \mathfrak{z}. \tag{E.1.2}$$

Namely,

$$[(X_1, Z_1), (X_2, Z_2)] = (0, A(X_1, X_2)), \quad X_j \in V,\ Z_j \in \mathfrak{z}. \tag{E.1.3}$$

A structure equivalent to (E.1.2) is $A : \Lambda^2 V \to \mathfrak{z}$; another equivalent structure is

$$A' : \mathfrak{z}' \longrightarrow (\Lambda^2 V)' \approx \Lambda^2 V'. \tag{E.1.4}$$

Inner products on $\mathfrak{z}$ and on V produce isomorphisms $\mathfrak{z}' \approx \mathfrak{z}$ and $\Lambda^2 V' \approx \mathrm{Sk}(V)$, the space of skew-adjoint linear operators on V, and hence the structure (E.1.4) is equivalent to

$$j : \mathfrak{z} \longrightarrow \mathrm{Sk}(V), \tag{E.1.5}$$

related to A by

$$\langle j(Z)X, Y\rangle = \langle A(X, Y), Z\rangle, \tag{E.1.6}$$

for $X, Y \in V,\ Z \in \mathfrak{z}$, where the left side of (E.1.6) uses the inner product on V and the right side uses the inner product on $\mathfrak{z}$.

This precisely captures all 2-step nilpotent Lie algebras. To guarantee that the center is precisely $\mathfrak{z}$, we add the non-degeneracy hypothesis

$$A(X, Y) = 0 \ \ \forall\, Y \in V \Longrightarrow X = 0, \tag{E.1.7}$$

or equivalently, if we have inner products on V and $\mathfrak{z}$ and use (E.1.5) to define the Lie algebra structure,

$$\bigcap_{Z \in \mathfrak{z}} \ker j(Z) = 0 \subset V. \tag{E.1.8}$$

Example 1. The Heisenberg Lie algebra $\mathcal{H}^n$ has the form (E.1.1)–(E.1.3) with

$$V = T^*\mathbb{R}^n \approx \mathbb{R}^{2n}, \quad \mathfrak{z} = \mathbb{R}, \tag{E.1.9}$$

and A the symplectic form on V (specified below). For $X_j = (x_j, y_j)^t \in V$, we have

$$A(X_1, X_2) = x_1 \cdot y_2 - x_2 \cdot y_1, \quad \langle X_1, X_2\rangle = x_1 \cdot x_2 + y_1 \cdot y_2, \tag{E.1.10}$$

and hence

$$j(Z) = ZJ, \quad J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}, \quad Z \in \mathbb{R}. \tag{E.1.11}$$

Generalizing Example 1, one says $\mathfrak{n}$ is of Heisenberg type if it is defined by the structures (E.1.1) and (E.1.5), with

$$j(Z)^2 = -|Z|^2 I, \quad \forall\, Z \in \mathfrak{z}. \tag{E.1.12}$$

This is equivalent to requiring

$$j(Z)j(W) + j(W)j(Z) = -2\langle Z, W\rangle\, I, \quad \forall\, Z, W \in \mathfrak{z}. \tag{E.1.13}$$

In other words, j extends to a unital representation of the Clifford algebra $C\ell(\mathfrak{z})$ on V. For example, we can take a representation of $C\ell(\mathfrak{z})$ on a direct sum of spaces of spinors. Note that

$$j(Z) = \sigma_D(Z) \tag{E.1.14}$$

is the symbol of a Dirac operator on $\mathfrak{z}$.

We bring in some notation. Given vector spaces V and $\mathfrak{z}$, and given $A \in \mathcal{L}(\Lambda^2 V, \mathfrak{z})$, we denote by $\mathfrak{n}_A$ the two-step nilpotent Lie algebra given by (E.1.1)–(E.1.3). The set of Lie algebras so produced is hence parametrized by $\mathcal{L}(\Lambda^2 V, \mathfrak{z})$. The condition (E.1.7) that the center of $\mathfrak{n}_A$ be exactly $\mathfrak{z}$ is that A belong to

$$\mathcal{N}^0(V, \mathfrak{z}) = \{A \in \mathcal{L}(\Lambda^2 V, \mathfrak{z}) : A(X, Y) = 0\ \forall\, Y \in V \Rightarrow X = 0\}. \tag{E.1.15}$$

We examine when A_1 and $A_2 \in \mathcal{N}^0(V, \mathfrak{z})$ yield isomorphic Lie algebras:

$$T : \mathfrak{n}_{A_1} \xrightarrow{\approx} \mathfrak{n}_{A_2}. \tag{E.1.16}$$

Since T must preserve the common center $\mathfrak{z}$, we see that T must have the form

$$T(X, Z) = (QX, RX + SZ), \quad Q \in \mathrm{Gl}(V),\ S \in \mathrm{Gl}(\mathfrak{z}),\ R \in \mathcal{L}(V, \mathfrak{z}). \tag{E.1.17}$$

The condition that such T be a Lie algebra isomorphism is

$$A_2(QX, QY) = SA_1(X, Y), \quad \forall\, X, Y \in V, \tag{E.1.18}$$

or equivalently $A_2(X, Y) = SA_1(Q^{-1}X, Q^{-1}Y)$. Thus, given $A_1 \in \mathcal{N}^0(V, \mathfrak{z})$, the set of elements $A_2 \in \mathcal{N}^0(V, \mathfrak{z})$ for which $\mathfrak{n}_{A_2} \approx \mathfrak{n}_{A_1}$ consists of the orbit of A_1 under the natural action on $\mathcal{L}(\Lambda^2 V, \mathfrak{z})$ of $\mathrm{Gl}(V) \times \mathrm{Gl}(\mathfrak{z})$.

We next make some remarks on identifying groups of automorphisms of a 2-step nilpotent Lie algebra $\mathfrak{n}$, constructed via (E.1.1) and either (E.1.2) or (E.1.5). Suppose a group G has representations π on V and ρ on $\mathfrak{z}$. If

$$A(\pi(g)X, \pi(g)Y) = \rho(g)A(X, Y), \quad \forall\, X, Y \in V, \quad g \in G, \tag{E.1.19}$$

it is clear that

$$(X, Z) \mapsto (\pi(g)X, \rho(g)Z) \tag{E.1.20}$$

yields a Lie algebra automorphism of $\mathfrak{n} = V \oplus \mathfrak{z}$. (This specializes (E.1.18) to $A_1 = A_2 = A$.) If we assume V and $\mathfrak{z}$ have inner products, via which we pass from (E.1.2) to (E.1.5), and the operators $\pi(g)$ and $\rho(g)$ preserve these inner products, then the hypothesis

$$j(\rho(g)Z) = \pi(g)j(Z)\pi(g)^{-1}, \quad \forall\, g \in G,\ Z \in \mathfrak{z} \tag{E.1.21}$$

readily yields (E.1.19), displaying the action of G as a group of automorphisms of $\mathfrak{n}$ in (E.1.20). If we do not assume $\pi(g)$ and $\rho(g)$ preserve these inner products, replace (E.1.21) by

$$j(\rho(g)^t Z) = \pi(g)^t j(Z)\pi(g), \quad \forall g \in G,\ Z \in \mathfrak{z}. \tag{E.1.22}$$

Example 2. Dilations. If we take $G = \mathbb{R}$ and $\pi(t)X = e^t X$, $\rho(t)Z = e^{2t}Z$ in (E.1.20), it is clear that (E.1.19) holds. Thus each two-step nilpotent Lie algebra $\mathfrak{n} = V \oplus \mathfrak{z}$ has the group of dilations

$$\delta(t)(X, Z) = (e^t X, e^{2t} Z) \tag{E.1.23}$$

as a group of automorphisms.

Example 3. Let $\mathfrak{n} = \mathcal{H}^n = T^*\mathbb{R}^n \oplus \mathbb{R}$, as in Example 1, and let $G = \mathrm{Sp}(n, \mathbb{R})$, the group of linear operators on $T^*\mathbb{R}^n$ preserving the symplectic form, hence yielding a representation π of G on $V = T^*\mathbb{R}^n$. Let ρ be the trivial representation of G on $\mathbb{R}$. Then (E.1.19) obviously holds, so $\mathrm{Sp}(n, \mathbb{R})$ acts as a group of automorphisms of $\mathcal{H}^n$.

Example 4. Let G be a compact semisimple Lie group with Lie algebra $\mathfrak{g}$, and let π be a representation of G on V, via operators preserving its inner product. Let $\mathfrak{z} = \mathfrak{g}$, as a linear space, with inner product given by the negative of the Killing form. Then take

$$j = d\pi : \mathfrak{z} \longrightarrow \mathrm{Sk}(V), \tag{E.1.24}$$

to define a 2-step nilpotent Lie algebra $\mathfrak{n}$. We have (E.1.21) with

$$\rho(g)Z = (\mathrm{Ad}\, g)Z, \tag{E.1.25}$$

so G acts as a group of automorphisms of $\mathfrak{n}$.

Remark. Throughout the constructions above, we need not insist that the inner products on $\mathfrak{z}$ and V be positive-definite. They could be non-degenerate inner products with other signatures. Thus we can extend the scope of Example 4 to include noncompact semisimple Lie groups. Also we can replace the hypothesis that G be semisimple by the more general hypothesis that $\mathfrak{g}$ possess a non-degenerate Ad-invariant inner product, so G can be a real reductive group. We do need G to act on V, preserving a non-degenerate inner product. For example, we could take $V = \mathfrak{g}$, $\pi(g) = \operatorname{Ad} g$, or V could be some G-invariant subspace of $\otimes^k \mathfrak{g}$, as long as it inherits a non-degenerate inner product. Other examples:

$$G = \mathrm{SO}(p, q), \quad V = \mathbb{R}^{p,q}. \tag{E.1.26}$$

For the nilpotent Lie algebras considered in Examples 3 and 4, we have both the group of automorphisms constructed there (action of $\mathrm{Sp}(n, \mathbb{R})$ and of G, respectively) and the groups of dilations constructed in Example 2. These are mutually commuting groups of automorphisms of $\mathfrak{n}$. Some of the nilpotent Lie algebras of Example 4 have a much larger group of automorphisms, such as described in the next example.

Example 5. Let $G = \mathrm{SO}(n)$, and let $\rho = \mathrm{Ad}$, as in (E.1.25). We take $V = \mathbb{R}^n$ and let π be the standard representation of $\mathrm{SO}(n)$ on $\mathbb{R}^n$. Then $\mathfrak{g} \approx \mathrm{Skew}(\mathbb{R}^n)$, and we have

$$\pi(g)X = gX, \quad \rho(g)Z = gZg^t, \quad X \in \mathbb{R}^n,\ Z \in \mathrm{Skew}(\mathbb{R}^n), \tag{E.1.27}$$

where we use the fact that $g^{-1} = g^t$ for $g \in \mathrm{SO}(n)$. We set $\mathfrak{n} = V \oplus \mathfrak{g}$, with Lie bracket defined by j as in (E.1.23), which in this setting is tautological:

$$j(Z) = Z, \quad Z \in \mathfrak{g} \approx \mathrm{Skew}(\mathbb{R}^n). \tag{E.1.28}$$

Example 4 specializes to yield $\mathrm{SO}(n)$ acting on $\mathfrak{n}$ as a group of automorphisms. We claim this enlarges to

$$\mathrm{Gl}(n, \mathbb{R}) \longrightarrow \operatorname{Aut} \mathfrak{n}, \tag{E.1.29}$$

given by (E.1.20), where $\pi(g)$ and $\rho(g)$ are again defined by (E.1.27), for $g \in \mathrm{Gl}(n, \mathbb{R})$. To verify (E.1.22), note that

$$j(\rho(g)^t Z) = g^t Z g \ \text{ and } \ \pi(g)^t j(Z) \pi(g) = g^t Z g, \tag{E.1.30}$$

for all $g \in \mathrm{Gl}(n, \mathbb{R})$, $Z \in \mathrm{Skew}(\mathbb{R}^n)$, in the current setting. Note that the automorphism group (E.1.29) contains both the $\mathrm{SO}(n)$ action and the group $\delta(t)$ of dilations from Example 2.

E.2. The Frobenius reciprocity theorem

Let G be a compact group, H a closed subgroup. Let ρ be an irreducible representation of H. We define the induced representation

$$\pi = \mathrm{Ind}_H^G(\rho) \tag{E.2.1}$$

as follows. The representation π is given by the left G-action on

$$^H W_\rho = \{u \in L^2(G, V_\rho) : u(gh) = \rho(h)^{-1}u(g),\ \forall\, h \in H\}. \tag{E.2.2}$$

The following result is the Frobenius reciprocity theorem.

Theorem E.2.1. *Let ρ and π be as above, and let λ be an irreducible representation of G. Set*

$$\begin{aligned} \mu(\pi, \lambda) &= \textit{Number of times } \pi \textit{ contains } \lambda, \\ \nu(\lambda, \rho) &= \textit{Number of times } \lambda\big|_H \textit{ contains } \rho. \end{aligned} \tag{E.2.3}$$

Then

$$\mu(\pi, \lambda) = \nu(\lambda, \rho). \tag{E.2.4}$$

Proof. We first note that, by the orthogonality relations,

$$\nu(\lambda, \rho) = \int\limits_H \overline{\chi_\lambda(h)}\chi_\rho(h)\, dh. \tag{E.2.5}$$

We aim to show that $\mu(\pi, \lambda)$ is equal to the same integral.

We next note that, if we set

$$^H H_\rho = \{u \in L^2(G) : {}^H R \text{ acts like copies of } \overline{\rho}\}, \tag{E.2.6}$$

where $^H R$ denotes the right regular representation of H on $L^2(G)$, then $^H H_\rho$ is isomorphic to a sum of d_ρ copies of $^H W_\rho$, and

$$\begin{aligned} &\text{the left } G\text{-action on } {}^H H_\rho \text{ is isomorphic to} \\ &\text{a sum of } d_\rho \text{ copies of } \pi. \end{aligned} \tag{E.2.7}$$

Now the orthogonal projection of $L^2(G)$ on $^H H_\rho$ is given by

$$P_1 v(x) = d_\rho \int\limits_H v(xh)\chi_\rho(h)\, dh, \tag{E.2.8}$$

and the orthogonal projection of $L^2(G)$ onto the space where G acts on the left like copies of λ is given by

$$\begin{aligned} P_2 w(x) &= d_\lambda \int\limits_G w(g^{-1}x)\overline{\chi_\lambda(g)}\, dg \\ &= d_\lambda \int\limits_G w(g)\overline{\chi_\lambda(g^{-1}x)}\, dg. \end{aligned} \tag{E.2.9}$$

Note that these projections commute, and, by (E.2.7),

$$d_\rho d_\lambda \mu(\pi, \lambda) = \operatorname{Tr} P_2 P_1. \tag{E.2.10}$$

Now, by (E.2.8)–(E.2.9), we have

$$\begin{aligned} P_2 P_1 v(x) &= d_\rho d_\lambda \int\limits_G \int\limits_H v(gh) \chi_\rho(h) \overline{\chi_\lambda(g^{-1}x)}\, dh\, dg \\ &= d_\rho d_\lambda \int\limits_G \int\limits_H v(g) \chi_\rho(h) \overline{\chi_\lambda(hg^{-1}x)}\, dh\, dg, \end{aligned} \tag{E.2.11}$$

so $\operatorname{Tr} P_2 P_1$ is clearly $d_\rho d_\lambda$ times the right side of (E.2.5), and the theorem is proved. $\square$

E.3. Differential geometric properties of compact Lie groups

Let G be a Lie group. Assume G has a bi-invariant Riemannian metric (as it would have if it were compact). We will obtain some interesting results on the interplay between the geometry of G, and the algebraic behavior of G and its Lie algebra $\mathfrak{g}$. Differential geometry background can be found in Appendix C of [**39**], or Chapter 6 of [**40**].

We begin with the following observation.

Lemma E.3.1. *The map $\psi : G \to G$ given by $\psi(x) = x^{-1}$ is an isometry of G, fixing the identity element e.*

Proof. This map takes a left-invariant metric to a right-invariant metric and vice-versa, hence it takes a bi-invariant metric to a bi-invariant metric. We have

$$D\psi(e) = -I \quad \text{on} \quad \mathfrak{g} = T_e G, \tag{E.3.1}$$

so such a metric tensor is preserved. □

More generally, for each $g \in G$, we have

$$\psi_g : G \longrightarrow G, \quad \psi_g(x) = gx^{-1}g, \tag{E.3.2}$$

an isometry of G, fixing g, and satisfying $D\psi_g(g) = -I$ on $T_g G$.

Using this, we establish the following key result.

Proposition E.3.2. *If γ is a unit speed geodesic on G satisfying $\gamma(0) = e$, then*

$$\gamma(s + t) = \gamma(s)\gamma(t). \tag{E.3.3}$$

Proof. Fix $t \in \mathbb{R}$ and consider $\sigma(s) = \gamma(t + s)$. This is the unit speed geodesic satisfying $\sigma(0) = \gamma(t),\ \sigma'(0) = \gamma'(t)$. It follows that $\tilde{\sigma}(s) = \psi_{\gamma(t)}(\sigma(s))$ is the unit speed geodesic satisfying $\tilde{\sigma}(0) = \gamma(t),\ \tilde{\sigma}'(0) = -\gamma'(t)$. This forces $\tilde{\sigma}(s) = \gamma(t - s)$, i.e.,

$$\gamma(t - s) = \psi_{\gamma(t)}(\gamma(t + s)) = \gamma(t)\gamma(t + s)^{-1}\gamma(t). \tag{E.3.4}$$

Taking $t = 0$ gives

$$\gamma(-s) = \gamma(s)^{-1}, \tag{E.3.5}$$

and then taking $s \mapsto -s$ gives

$$\gamma(s + t) = \gamma(t)\gamma(s - t)\gamma(t). \tag{E.3.6}$$

Taking $s = t$ gives $\gamma(2t) = \gamma(t)^2$, and then we obtain by induction that

$$\gamma((n + 1)t) = \gamma(t)\gamma((n - 1)t)\gamma(t) = \gamma(t)^{n+1}, \tag{E.3.7}$$

for each $n \in \mathbb{N}$. A limiting argument then gives (E.3.3) when s and t have the same sign. In such a case, (E.3.4) gives

$$\gamma(t-s) = \gamma(t)\gamma(s)^{-1}\gamma(t)^{-1}\gamma(t) = \gamma(t)\gamma(-s), \tag{E.3.8}$$

so we have (E.3.3) in general. □

Proposition E.3.2 implies that if $X \in \mathfrak{g}$, then $\gamma_X(t) = \operatorname{Exp} tX$ is a constant speed geodesic through e. The geodesic equation gives $\nabla_X X = 0$. Given also $Y \in \mathfrak{g}$, we have $\nabla_Y Y = 0$ and $\nabla_{X+Y}(X+Y) = 0$, hence

$$\nabla_X Y + \nabla_Y X = 0. \tag{E.3.9}$$

Since

$$\nabla_X Y - \nabla_Y X = [X, Y], \tag{E.3.10}$$

we obtain

$$\nabla_X Y = \frac{1}{2}[X, Y]. \tag{E.3.11}$$

The Riemann tensor R is defined by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{E.3.12}$$

It then follows that, if $X, Y, Z \in \mathfrak{g}$,

$$R(X,Y)Z = -\frac{1}{4}[[X,Y],Z]. \tag{E.3.13}$$

The Ricci tensor is then given by

$$\operatorname{Ric}(X,Y) = \sum_j \langle R(X, E_j)Y, E_j \rangle, \tag{E.3.14}$$

where $\{E_j\}$ is an orthonormal basis of $\mathfrak{g}$. It follows that if $X, Y \in \mathfrak{g}$,

$$\operatorname{Ric}(X,Y) = \frac{1}{4} \sum_j \langle [X, E_j], [Y, E_j] \rangle, \tag{E.3.15}$$

since the inner product on $\mathfrak{g}$ is Ad-invariant. In particular,

$$\operatorname{Ric}(X,X) = \frac{1}{4} \sum_j \|[X, E_j]\|^2. \tag{E.3.16}$$

This readily yields the following.

Proposition E.3.3. *Given $X \in \mathfrak{g}$,*

$$\operatorname{Ric}(X,X) > 0, \quad \textit{unless } X \in \mathfrak{z}, \textit{ the center of } \mathfrak{g}. \tag{E.3.17}$$

Proposition E.3.3 has the following important consequence.

Proposition E.3.4. *If G is a connected Lie group with a bi-invariant Riemannian metric, and if the center of its Lie algebra is 0, then G is compact. Hence, if G is a compact Lie group and its Lie algebra has trivial center, then its universal covering group $\widetilde{G}$ is compact.*

Example. The universal covering group of $\mathrm{SO}(n)$ is compact, for $n \geq 3$.

Proposition E.3.4 follows from Proposition E.3.3 together with the following classical result, known as Meyer's theorem.

Proposition E.3.5. *If M is a complete, connected Riemannian manifold of dimension n and*

$$\mathrm{Ric}(X, X) \geq (n-1)\kappa \|X\|^2 \tag{E.3.18}$$

for some $\kappa > 0$, then M is compact, of diameter $\leq \pi/\sqrt{\kappa}$.

We sketch a proof of this, refering to Chapter 1 of [**10**] or §19 of [**29**] for details. Let γ_s be a 1-parameter family of curves such that $\gamma_s(a) \equiv p$ and $\gamma_s(b) \equiv q$. Assume γ_0 is a constant speed geodesic. Define the energy

$$E(s) = \frac{1}{2} \int_a^b \langle \gamma_s'(t), \gamma_s'(t) \rangle \, dt. \tag{E.3.19}$$

We set $T = \gamma_0'(t),\ V = \partial_s \gamma_s(t)|_{s=0}$. One has

$$\begin{aligned} E'(0) &= \frac{1}{2} \int_a^b V \langle T, T \rangle \, dt \\ &= \int_a^b \langle \nabla_V T, T \rangle \, dt. \end{aligned} \tag{E.3.20}$$

Using $T\langle V, T\rangle = \langle \nabla_T V, T \rangle + \langle V, \nabla_T T \rangle$, we get

$$E'(0) = -\int_a^b \langle V, \nabla_T T \rangle \, dt. \tag{E.3.21}$$

If γ_0 is a geodesic, this is 0 for all variations γ_s, fixed at the endpoints $t = a$ and $t = b$. This yields the geodesic equation $\nabla_T T = 0$. Going further, one has the second variational formula

$$E''(0) = \int_a^b \Big[\langle R(V, T)V, T \rangle + \langle \nabla_T V, \nabla_T V \rangle \Big] \, dt. \tag{E.3.22}$$

With this background, we turn to the proof of Proposition E.3.5. Let γ_0 be a unit speed geodesic on M, say $\gamma_0 : [0, \ell] \to M$, so γ_0 has length

ℓ. Say $\gamma_0(0) = p$, $\gamma_0(\ell) = q$, and $\gamma_0'(t) = T$. Let $\{T, e_1, \dots, e_{n-1}\}$ be an orthonormal basis of T_pM, and extend e_j along γ_0 by $\nabla_T e_j = 0$. Set

$$V_j(t) = \sin \frac{\pi t}{\ell}\, e_j(t). \tag{E.3.23}$$

For each j, construct a family of curves γ_s so that $\partial_s \gamma_s(t) = V_j(t)$ at $s = 0$. Then use (E.3.22) with $V = V_j$ and denote the result by $E_j''(0)$. We have

$$\begin{aligned}\sum_j E_j''(0) &= \int_0^\ell \Big\{ (n-1)\frac{\pi^2}{\ell^2} \cos^2 \frac{\pi t}{\ell} - \mathrm{Ric}(T,T) \sin^2 \frac{\pi t}{\ell} \Big\}\, dt \\ &\le (n-1) \int_0^\ell \Big\{ \frac{\pi^2}{\ell^2} \cos^2 \frac{\pi t}{\ell} - \kappa \sin^2 \frac{\pi t}{\ell} \Big\}\, dt \\ &= (n-1)\frac{\ell}{2}\Big(\frac{\pi^2}{\ell^2} - \kappa\Big).\end{aligned} \tag{E.3.24}$$

This is < 0 if $\ell > \pi/\sqrt{\kappa}$. In such a case, γ_0 cannot be length minimizing from $p = \gamma_0(0)$ to $q = \gamma_0(\ell)$. Consequently, given $p, q \in M$, these points can be joined by a curve of length $\le \pi/\sqrt{\kappa}$. This proves the diameter estimate of Proposition P.4, hence the compactness.

E.4. From G_2 to E_8

The complexification of Der($\mathbb{O}$), analyzed in §11.4, is the first of 5 exceptional complex simple Lie algebras, introduced by Killing and Cartan, denoted $\mathfrak{G}_2, \mathfrak{F}_4, \mathfrak{E}_6, \mathfrak{E}_7$, and $\mathfrak{E}_8$. We describe a uniform construction of $\mathfrak{G}_2$ and $\mathfrak{E}_8$, due to Freudenthal. In each case, the complex Lie algebra has a $\mathbb{Z}/(3)$ grading:

$$\mathfrak{g} = \mathfrak{g}_{-1} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_1, \quad \mathbb{Z}/(3) = \{-1, 0, 1\}. \tag{E.4.1}$$

We will have $[\mathfrak{g}_j, \mathfrak{g}_k] \subset \mathfrak{g}_{j+k}$, with $j + k$ computed mod 3. In each case, the complex Lie algebra $\mathfrak{g}_0$ has a representation ρ on a complex vector space, with contragredient representation ρ' on V'. We set

$$\mathfrak{g}_1 = V, \quad \mathfrak{g}_{-1} = V', \tag{E.4.2}$$

and define the actions $[\mathfrak{g}_0, \mathfrak{g}_j] \to \mathfrak{g}_j$ via these representations. In the cases $\mathfrak{g} = \mathfrak{G}_2$ or $\mathfrak{E}_8$, we take respectively

$$\mathfrak{g}_0 = s\ell(3, \mathbb{C}), \quad \mathfrak{g}_0 = s\ell(9, \mathbb{C}), \tag{E.4.3}$$

and, respectively,

$$V = \mathbb{C}^3 \ \text{ and } \ V = \Lambda^3\mathbb{C}^9. \tag{E.4.4}$$

There is a natural representation ρ of $\mathfrak{g}_0$ on V in each case. In the first case, we have $\Lambda^3 V = \Lambda^3\mathbb{C}^3 \approx \mathbb{C}$, via an invariant complex volume element, and in the second case $\Lambda^3 V \to \Lambda^9\mathbb{C}^9 \approx \mathbb{C}$. Thus we have natural bilinear maps

$$V \times V \longrightarrow V', \quad V' \times V' \longrightarrow V, \tag{E.4.5}$$

which are anti-symmetric. These define Lie brackets

$$[\mathfrak{g}_1, \mathfrak{g}_1] \to \mathfrak{g}_{-1}, \quad [\mathfrak{g}_{-1}, \mathfrak{g}_{-1}] \to \mathfrak{g}_1. \tag{E.4.6}$$

It remains to specify

$$[\mathfrak{g}_1, \mathfrak{g}_{-1}] \to \mathfrak{g}_0. \tag{E.4.7}$$

This is done as follows. Given $v \in V, v' \in V'$, we define $[v, v'] \in \mathfrak{g}_0$ by

$$-B(\lambda, [v, v']) = \langle \rho(\lambda)v, v' \rangle, \quad \lambda \in \mathfrak{g}_0, \tag{E.4.8}$$

where B is the Killing form on the simple Lie algebra $\mathfrak{g}_0$.

In this fashion, the Lie algebras are constructed. For $\mathfrak{G}_2$, the construction outlined here is consistent with the analysis of the complexification of Der($\mathbb{O}$) done in §11.4. For $\mathfrak{E}_8$, one needs to verify that the "products" defined above satisfy the Jacobi identity. For details on this, and the analysis of the root system for $\mathfrak{E}_8$, see [**1**], [**35**].

E.5. Dyson integrals and generalizations

Let $u, v : S^1 \to \mathbb{C}$ be continuous. Given $g \in \mathrm{U}(n)$, we define $u(g), v(g) \in \mathrm{End}(\mathbb{C}^n)$ by the spectral representation. As advertised in §4.15, it is of interest to obtain formulas for

$$\int_{\mathrm{U}(n)} X_u(g) X_v(g)\, dg, \tag{E.5.1}$$

where

$$X_u(g) = \mathrm{Tr}\, u(g). \tag{E.5.2}$$

Such integrals arise in random matrix theory. An evaluation of (E.5.1) was made by Bump, Diaconis, and Keller in [**8**]. The work made use of identities of Dyson [**17**], also presented in [**28**]. Here we present an alternative approach. Going further, we seek formulas for

$$\int_{\mathrm{U}(n)} u(g) \otimes v(g)\, dg, \tag{E.5.3}$$

for which (E.5.1) is the trace. One complicating feature of (E.5.3) is that the integrand is not a central function.

To attack (E.5.3), we use Fourier series:

$$u(g) = \sum_{j=-\infty}^{\infty} \hat{u}(j) g^j, \quad \hat{u}(j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(\theta) e^{-ij\theta}\, d\theta. \tag{E.5.4}$$

We see that (E.5.3) is equal ro

$$\sum_{j,k} \hat{u}(j)\hat{v}(k) A_{jk}, \quad A_{jk} = \int_{\mathrm{U}(n)} g^j \otimes g^k\, dg. \tag{E.5.5}$$

Performing the measure-preserving transformation $g \mapsto e^{i\psi} g$ on $\mathrm{U}(n)$, we see that

$$A_{jk} = e^{i(j+k)\psi} A_{jk}, \tag{E.5.6}$$

for all $\psi \in \mathbb{R}$, and hence $A_{jk} = 0$ for $j \neq -k$. Hence

$$\int_{\mathrm{U}(n)} u(g) \otimes v(g)\, dg = \sum_j \hat{u}(j)\hat{v}(-j) A_j, \tag{E.5.7}$$

with

$$A_j = \int_{\mathrm{U}(n)} g^j \otimes g^{-j}\, dg. \tag{E.5.8}$$

Note also that performimg the measure-preserving transformation $g \mapsto g^{-1}$ on $\mathrm{U}(n)$ yields

$$A_j = A_{-j}. \tag{E.5.9}$$

From (E.5.7) we deduce that

$$\int_{\mathrm{U}(n)} X_u(g) X_v(g)\, dg = \sum_j \hat{u}(j)\hat{v}(-j) a_{jn}, \tag{E.5.10}$$

where

$$a_{jn} = \operatorname{Tr} A_j = \int_{\mathrm{U}(n)} |\operatorname{Tr} g^j|^2\, dg. \tag{E.5.11}$$

Clearly

$$a_{0n} = n^2. \tag{E.5.12}$$

Meanwhile, as shown in §4.15, for $j \geq 1$, $a_{jn} = j \wedge n$. By (E.5.9) we have $a_{-j,n} = a_{jn}$. Hence

$$a_{jn} = |j| \wedge n, \quad \text{for} \quad j \neq 0. \tag{E.5.13}$$

This provides an analysis of (E.5.1).

Operator analysis of (E.5.3)

For the operator analysis of (E.5.3), we need an analysis of the operators A_j, given by (E.5.8), which we proceed to investigate. Let us define

$$\sigma : \mathbb{C}^n \otimes \mathbb{C}^n \longrightarrow \mathbb{C}^n \otimes \mathbb{C}^n \tag{E.5.14}$$

by

$$\sigma(u \otimes v) = v \otimes u. \tag{E.5.15}$$

(Here u and v are elements of $\mathbb{C}^n$, not of $C(S^1)$.) Let I denote the identity operator on $\mathbb{C}^n \otimes \mathbb{C}^n$. We will establish the following

Proposition E.5.1. *For $j \geq 1, n \geq 2$, we have*

$$A_j = \frac{n^2 - (j \wedge n)}{n(n^2 - 1)} \sigma + \frac{(j \wedge n) - 1}{n^2 - 1} I. \tag{E.5.16}$$

Proof. Take $X \in \mathrm{U}(n)$, acting on $\mathbb{C}^n \otimes \mathbb{C}^n$ by $X(u \otimes v) = Xu \otimes Xv$. Then

$$\begin{aligned} &X^{-1} A_j X(u \otimes v) \\ &= \int_{\mathrm{U}(n)} X^{-1} g^j X u \otimes X^{-1} g^{-j} X v\, dg \\ &= A_j(u \otimes v), \end{aligned} \tag{E.5.17}$$

since $g \mapsto X^{-1}gX$ is a measure preserving transformation on $\mathrm{U}(n)$. Thus A_j commutes with the action of $\mathrm{U}(n)$ on $\mathbb{C}^n \otimes \mathbb{C}^n$. It follows that

$$A_j = \alpha_{jn}\sigma + \beta_{jn} I, \tag{E.5.18}$$

for some scalars α_{jn} and β_{jn}. Taking the trace of the right side of (E.5.18) and comparing with (E.5.11)–(E.5.13), we have

$$n\alpha_{jn} + n^2\beta_{jn} = j \wedge n. \tag{E.5.19}$$

To obtain a second identity involving α_{jn} and β_{jn}, we proceed as follows. Write (E.5.18) as

$$A_j(u \otimes v) = \alpha_{jn} v \otimes u + \beta_{jn} u \otimes v. \tag{E.5.20}$$

Setting $u = v$ and taking the inner product of both sides with $u \otimes u$ yields

$$\begin{aligned}(\alpha_{jn} + \beta_{jn})|u|^4 &= \int_{\mathrm{U}(n)} (g^j u \otimes g^{-j} u, u \otimes u)\, dg \\ &= \int_{\mathrm{U}(n)} (g^j u, u)(g^{-j} u, u)\, dg \\ &= \int_{\mathrm{U}(n)} |(g^j u, u)|^2\, dg.\end{aligned} \tag{E.5.21}$$

On the other hand, taking $u \perp v$ and taking the inner product of both sides of (E.5.20) with $v \otimes u$ yields

$$\begin{aligned}\alpha_{jn}|u|^2|v|^2 &= \int_{\mathrm{U}(n)} (g^j u \otimes g^{-j} v, v \otimes u)\, dg \\ &= \int_{\mathrm{U}(n)} (g^j u, v)(v, g^j u)\, dg \\ &= \int_{\mathrm{U}(n)} |(g^j u, v)|\, dg.\end{aligned} \tag{E.5.22}$$

Now let $e_1, \dots, e_n$ denote the standard orthonormal basis of $\mathbb{C}^n$, and set $u = e_1$. Apply (E.5.21) for $\ell = 1$ and (E.5.22) with $v = e_\ell$ for $\ell > 1$, and sum, to get

$$\begin{aligned}\alpha_{jn} + \beta_{jn} + (n-1)\alpha_{jn} &= \sum_{\ell=1}^{n} \int_{\mathrm{U}(n)} |(g^j e_1, e_\ell)|^2\, dg \\ &= \int_{\mathrm{U}(n)} |g^j e_1|^2\, dg.\end{aligned} \tag{E.5.23}$$

Since $|g^j e_1| = 1$, we have our second identity:

$$n\alpha_{jn} + \beta_{jn} = 1. \tag{E.5.24}$$

Now (E.5.16) follows. $\square$

With (E.5.16) in hand (and noting that $A_0 = I$), we can finish the computation of (E.5.7). We obtain

$$\begin{aligned}\int_{\mathrm{U}(n)} u(g) \otimes v(g)\, dg = {} & \frac{1}{n^2-1}\big[w(0) - F_n w(0)\big]\Big(I - \frac{\sigma}{n}\Big) \\ & - \frac{1}{n^2-1}\big[w(0) - \hat{w}(0)\big](I - n\sigma) \\ & + \hat{w}(0) I,\end{aligned} \tag{E.5.25}$$

where

$$w(\theta) = u * \check{v}(\theta) = \frac{1}{2\pi}\int_{S^1} u(\varphi) v(\varphi - \theta)\, d\varphi, \tag{E.5.26}$$

and where $F_n w$ denotes the nth Fejér mean of the Fourier series of w:

$$F_n w(\theta) = \sum_{|j|<n} \Big(1 - \frac{|j| \wedge n}{n}\Big) \hat{w}(j) e^{ij\theta}. \tag{E.5.27}$$

Note that $\hat{w}(0)$ is equal to the product of the mean values of u and v.

Remark. Taking $v \in \mathbb{C}^n$ orthogonal to u and taking the inner product of both sides of (E.5.20) with $u \otimes v$ yields the identity

$$\int_{\mathrm{U}(n)} (g^j u, u)\overline{(g^j v, v)}\, dg = \beta_{jn} |u|^2 |v|^2, \quad \text{for } \ u \perp v. \tag{E.5.28}$$

Bibliography

[1] J. Adams, Lectures on Exceptional Lie Groups, Univ. Chicago press, 1996.

[2] D. Altschuler and C. Itzykson, Remarks on integration over Lie algebras, Ann. Inst. Henri Poincaré 54 (1991), 1–8.

[3] M. Artin, Algebra, Prentice-Hall, Englewood Cliffs, NJ, 1991.

[4] J. Baez, The octonions, Bull. AMS 39 (2002), 145–205.

[5] P. Bleher and A. Its (eds.), Random Matrix Models and Their Applications, MSRI Publications #40, Cambridge Univ. Press, 2001.

[6] R. Bryant, Submanifolds and special structures on the octonions, J. Diff. Geom. 17 (1982), 185–232.

[7] R. Bryant, On the geometry of almost-complex 6-manifolds, Asian J. Math. 10 (2006), 561–605.

[8] D. Bump, P. Diaconis, and J. Keller, Unitary correlations and the Fejer kernel, Math. Phys. Anal. Geom. 5 (2002), 101–123.

[9] R. Cahn, Semi-simple Lie Algebras and Their Representations, Dover, New York, 2006.

[10] J. Cheeger and D. Ebin, Comparison Theorems in Riemannian Geometry, North-Holland, Amsterdam, 1975.

[11] M. Curtis, Matrix Gorups, Springer, NY, 1979.

[12] P. Deift, Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach, Courant Lecture Notes #3, AMS, Providence, R.I., 1998.

[13] P. Diaconis and S. Evans, Linear functionals of eigenvalues of random matrices, Trans. AMS 353 (2001), 2615–2633.

[14] P. Diaconis and M. Shahshahani, On the eigenvalues of random matrices, Trans. AMS 353 (2001), 2615–2633.

[15] T. Dray and C. Manoque, The Geometry of the Octonions, World Scientific, NJ, 2015.

[16] J.J. Duistermaat and J. Kolk, Lie Groups, Springer-Verlag, New York, 2000.

[17] F. Dyson, Statistical theory of energy levels of complex systems, I–III, J. Math. Phys. 3 (1962), 140–156; 157–165; 166–175.

[18] N. Ercolani and K. McLaughlin, Asymptotics and integrable structures for biorthogonal polynomials associated to a random matrix model, Physica D 152–153 (2001), 232–268.

[19] W. Fulton and J. Harris, Representation Theory, a First Course, Springer-Verlag, New York, 1991.

[20] R. Goodman and N. Wallach, Representations and invariants of the classical groups, Cambridge Univ. Press, 1998.

[21] F. R. Harvey, Spinors and Calibrations, Academic Press, NY, 1990.

[22] M. Hausner and J. Schwartz, Lie Groups; Lie Algebras, Gordon and Breach, London, 1968.

[23] S. Helgason, Groups and Geometric Analysis, Academic Press, New York, 1984.

[24] R. Howe, Remarks on classical invariant theory, Trans. AMS 313 (1989), 539–570.

[25] C. Itzykson and J. Zuber, The planar approximation, II, J. Math. Phys. 21 (1980), 411–421.

[26] H. B. Lawson and M.-L. Michelson, Spin Geometry, Princeton Univ. Press, Princeton, N.J., 1989.

[27] K. McCrimmon, Jordan algebras and their applications, BAMS 84 (1978), 612–627.

[28] M. Mehta, Random Matrices, Academic Press, New York, 1991.

[29] J. Milnor, Morse Theory, Princeton Univ. Press, Princeton, NJ, 1963.

[30] I. Porteous, Clifford Algebras and Classical Groups, Cambridge Univ. Press, 1995.

[31] C. Procesi, Lie Groups – an Approach through Invariants and Representations, Springer, New York, 2007.

[32] J. Rosenblatt and M. Taylor, Group actions and singular martingales II, the recognition problem, Canad. J. Math. 56 (2004), 431–448.

[33] S. Sakai, C^*-algebras and W^*-algebras, Springer-Verlag, New York, 1971.

[34] B. Simon, Representations of finite and compact groups, AMS, Providence, R.I., 1996.

[35] T. Springer and F. Veldcamp, Octonions, Jordan Algebras, and Exceptional Groups, Springer, Berlin, 2000.

[36] N. Steenrod, The Topology of Fiber Bundles, Princeton Univ. Press, Princeton NJ, 1951.

[37] R. Strichartz, The Campbell-Hausdorff formula and DuHamel's principle, Preprint.

[38] M. Taylor, Noncommutative Harmonic Analysis, AMS, Providence, R.I., 1986.

[39] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer-Verlag, New York, 1996, 2nd ed. 2011.

[40] M. Taylor, Introduction to Analysis in Several Variables, AMS, 2020.

[41] M. Taylor, Measure Theory and Integration, AMS, Providence RI, 2006.

[42] M. Taylor, Linear Algebra, AMS, 2020.

[43] M. Taylor, Differential Geometry, Lecture Notes, available at http://mtaylor.web.unc.edu/notes

[44] M. Taylor, Introduction to Differential Equations, AMS, 2011, 2nd ed. 2022.

[45] V.S. Varadarajan, Lie groups, Lie algebras, and their representations, Prentice-Hall, Englewood Cliffs, N.J., 1974.

[46] N. Vilenkin, Special functions and the theory of group representations, Transl. Math. Monogr. 22, AMS, Providence, RI, 1968.

[47] N. Wallach, Harmonic Analysis on Homogeneous Spaces, Marcel Dekker, New York, 1973.

[48] K. Yosida, Functional Analysis, Springer, N.Y., 1965.

[49] D. Zelobenko, Compact Lie Groups and their Representations, Transl. Math. Monogr., Vol. 40, AMS, Providence, R.I., 1973.

Index