Lectures on Lie Groups

Michael Taylor

Contents

Introduction

These notes are based on lectures I have given on Lie groups, in Math 273, at UNC. Prerequisites include the basic first-year graduate courses in analysis, algebra, geometry, and topology, and an introductory course in manifold theory.

The first seventeen sections deal with the general theory of Lie groups. We discuss integration on a Lie group, the Lie algebra, and general results on representations. We present some classical results on compact Lie groups, such as the Peter-Weyl theorem, on the completeness of the matrix entries of irreducible unitary representations of a compact Lie group G in $L^2(G)$.

Sections 18–34 concentrate 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 $\mathrm{U}(n)$.

Sections 35–38 extend some of the general results of §§19–21 to the setting of general compact Lie groups, particularly discussing roots of their Lie algebras and weights of their representations. Sections 39–44 provide a more detailed extension of such results from the setting of $\mathrm{U}(n)$ to the setting of the orthogonal groups $\mathrm{SO}(n)$ and certain two-fold covers, denoted $\mathrm{Spin}(n)$, which are introduced in §42, via use of Clifford algebras, which are introduced in §41.

These notes end with several appendices, presenting some background material and also some material complementary to that in the main body of the notes.

After reading these notes, 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, these notes should serve as preparation for study of the monograph [T1].

1. Definition and basic 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

$$\mathrm{Gl}(n, \mathbb{R}) = \{A \in \mathrm{M}(n, \mathbb{R}) : A^{-1} \text{ exists}\}, \tag{1.1}$$

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

Proposition 1.1. *The set* $\mathrm{Gl}(n, \mathbb{R})$ *is open in* $\mathrm{M}(n, \mathbb{R})$.

Proof. 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

$$\|C\| = \sup \{\|Cv\| : v \in \mathbb{R}^n, \|v\| \le 1\}, \tag{1.2}$$

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

$$\|C\| < 1 \Longrightarrow (I + C)^{-1} = \sum_{k \ge 0} (-C)^k, \tag{1.3}$$

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

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. One shows that $\kappa(A) = A^{-1}$ gives a smooth map on $\mathrm{Gl}(n, \mathbb{R})$, with

$$D\kappa(A)X = -A^{-1}XA^{-1}. \tag{1.4}$$

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

$$\begin{aligned} (A + X)^{-1} &= (A(I + A^{-1}X))^{-1} = (I + A^{-1}X)^{-1}A^{-1} \\ &= A^{-1} + \sum_{k \ge 1} (-1)^k (A^{-1}X)^k A^{-1}, \end{aligned} \tag{1.5}$$

which yields (1.4).

Similar considerations apply to

$$\mathrm{Gl}(n, \mathbb{C}) = \{A \in \mathrm{M}(n, \mathbb{C}) : A^{-1} \text{ exists}\}, \tag{1.6}$$

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

$$\mathrm{Sl}(n,\mathbb{F}) = \{A \in \mathrm{M}(n,\mathbb{F}) : \det A = 1\} \subset \mathrm{Gl}(n,\mathbb{F}), \quad \mathbb{F} = \mathbb{R} \text{ or } \mathbb{C}. \tag{1.7}$$

Other examples are

$$\begin{aligned} \mathrm{O}(n) &= \{A \in \mathrm{M}(n,\mathbb{R}) : A^*A = I\}, \\ \mathrm{U}(n) &= \{A \in \mathrm{M}(n,\mathbb{C}) : A^*A = I\}, \end{aligned} \tag{1.9}$$

where

$$A = (a_{jk}) \Longrightarrow A^* = (\overline{a}_{kj}). \tag{1.9}$$

Also we have

$$\begin{aligned} \mathrm{SO}(n) &= \{A \in \mathrm{O}(n) : \det A = 1\}, \\ \mathrm{SU}(n) &= \{A \in \mathrm{U}(n) : \det A = 1\}. \end{aligned} \tag{1.10}$$

The proof that (1.7)–(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.

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*

$$S = \{x \in V : F(x) = p\}. \tag{1.11}$$

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 $x \in S$,

$$T_xS = \mathit{ker}\, DF(x). \tag{1.12}$$

Proof. Given $q \in S$, apply the inverse function theorem to

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

where $P_q : W \to \ker DF(q)$ is the orthogonal projection.

We leave the verification of this condition for the examples given above as an exercise. In all cases, $V = \mathrm{M}(n,\mathbb{F})$. For $\mathrm{Sl}(n,\mathbb{F})$, you take $W = \mathbb{F}$, while for $O(n)$ you take W to be the space of real, symmetric $n \times n$ matrices.

Regarding $\mathrm{O}(n)$, we note that, given $A \in \mathrm{M}(n,\mathbb{R})$, defining $A : \mathbb{R}^n \to \mathbb{R}^n$,

$$A \in \mathrm{O}(n) \Longleftrightarrow (Au, Av) = (u, v), \quad \forall\, u, v \in \mathbb{R}^n, \tag{1.13}$$

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

$$(u,v)=\sum_j u_j v_j, \tag{1.14}$$

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

$$A\in \mathrm{U}(n)\Longleftrightarrow (Au,Av)=(u,v),\quad \forall\ u,v\in\mathbb{C}^n, \tag{1.15}$$

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

$$(u,v)=\sum_j u_j\overline{v}_j. \tag{1.16}$$

Note that

$$\langle u,v\rangle=\mathrm{Re}(u,v) \tag{1.17}$$

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

$$\mathrm{U}(n)\hookrightarrow \mathrm{O}(2n). \tag{1.18}$$

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 §3.

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

$$(A,v)x=Ax+v,\quad A\in\mathrm{O}(n),\ v,x\in\mathbb{R}^n. \tag{1.19}$$

The group law is seen to be

$$(A,v)\cdot(B,w)=(AB,Aw+v). \tag{1.20}$$

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

$$(A,v)\mapsto\begin{pmatrix}A & v\\ 0 & 1\end{pmatrix}, \tag{1.21}$$

as one verifies that

$$\begin{pmatrix}A & v\\ 0 & 1\end{pmatrix}\begin{pmatrix}B & w\\ 0 & 1\end{pmatrix}=\begin{pmatrix}AB & Aw+v\\ 0 & 1\end{pmatrix}. \tag{1.22}$$

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.

2. The matrix exponential and other functions of matrices

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

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k. \tag{2.1}$$

We also denote this by $\mathrm{Exp}(tA)$. Making use of the operator norm (1.2) and noting that $\|A^k\| \le \|A\|^k$, we see that (2.1) is absolutely convergent for all A and all t. The power series (2.1) can be differentiated term by term, and we obtain

$$\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A. \tag{2.2}$$

Using this we can establish the identity

$$e^{(s+t)A} = e^{sA} e^{tA}. \tag{2.3}$$

To get this, we can first compute

$$\frac{d}{dt} \bigl[e^{(s+t)A} e^{-tA} \bigr] = e^{(s+t)A} A e^{-tA} - e^{(s+t)A} A e^{-tA} = 0, \tag{2.4}$$

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

$$e^{(s+t)A} e^{-tA} = e^{sA}. \tag{2.5}$$

Setting $s = 0$ gives

$$e^{tA} e^{-tA} = I. \tag{2.6}$$

Thus e^{-tA} is the multiplicative inverse of e^{tA}. Using this, we can multiply both sides of (2.5) on the right by e^{tA} and obtain (2.3).

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

$$AB = BA \Longrightarrow e^{sA+tB} = e^{sA} e^{tB}, \tag{2.7}$$

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

We note a few easy identities:

$$e^{tX^{-1}AX} = X^{-1} e^{tA} X, \quad e^{tA^*} = \bigl(e^{tA} \bigr)^*, \tag{2.8}$$

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

$$\det e^{tA} = e^{t \operatorname{Tr} A}. \tag{2.9}$$

If A is diagonal this is checked by the remarks above; it then follows for A diagonalizable, by (2.8). It can be shown that the set of diagonalizable matrices is dense in $\mathrm{M}(n,\mathbb{C})$, and then (2.9) holds for all A, by continuity. Alternatively, it is quite easy to show that there exists an *open* subset of $\mathrm{M}(n,\mathbb{C})$ consisting of diagonalizable matrices. Since both sides of (2.9) are holomorphic on $\mathrm{M}(n,\mathbb{C})$, this suffices.

We remark on the behavior of the exponential map on the tangent space at the identity to the groups described in (1.7)–(1.10). Making use of the criterion (1.12), one can calculate the following:

$$\begin{aligned} T_I \,\mathrm{Sl}(n,\mathbb{F}) &= \{A \in \mathrm{M}(n,\mathbb{F}) : \operatorname{Tr} A = 0\}, \\ T_I \,\mathrm{O}(n) &= \{A \in \mathrm{M}(n,\mathbb{R}) : A^* = -A\} = T_I \,\mathrm{SO}(n), \\ T_I \,\mathrm{U}(n) &= \{A \in \mathrm{M}(n,\mathbb{C}) : A^* = -A\}, \\ T_I \,\mathrm{SU}(n) &= \{A \in \mathrm{M}(n,\mathbb{C}) : A^* = -A,\ \operatorname{Tr} A = 0\}. \end{aligned} \tag{2.10}$$

Then, making use of (2.8)–(2.9), one readily verifies the following.

Proposition 2.1. *For each Lie group listed above,*

$$\mathrm{Exp} : T_I G \longrightarrow G. \tag{2.11}$$

We will discuss how this result fits in a more general framework in §12.

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

$$\frac{d}{dt} e^{A+tB}\Big|_{t=0}. \tag{2.12}$$

For this it is useful to look at

$$U(s,t) = e^{s(A+tB)}, \tag{2.13}$$

which satisfies

$$\frac{\partial U}{\partial s} = (A+tB)U(s,t), \quad U(0,t) = I. \tag{2.14}$$

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

$$\frac{\partial}{\partial s} U_t(s,t) = (A+tB)U_t(s,t) + BU(s,t), \quad U_t(0,t) = 0, \tag{2.15}$$

and in particular

$$\frac{\partial}{\partial s}U_t(s,0) = AU_t(s,0) + BU(s,0), \quad U_t(0,0) = 0. \tag{2.16}$$

This is an inhomogeneous linear ODE, whose solution is

$$\begin{aligned} U_t(s,0) &= \int_0^s e^{(s-\sigma)A} BU(\sigma,0)\, d\sigma \\ &= \int_0^s e^{(s-\sigma)A} Be^{\sigma A}\, d\sigma. \end{aligned} \tag{2.17}$$

We get (2.12) by setting $s = 1$:

$$\frac{d}{dt} e^{A+tB}\big|_{t=0} = \int_0^1 e^{(1-\sigma)A} Be^{\sigma A}\, d\sigma, \tag{2.18}$$

so

$$D\ \mathrm{Exp}(A)B = e^A \int_0^1 e^{-\sigma A} Be^{\sigma A}\, d\sigma. \tag{2.19}$$

The method (2.1) of defining the matrix exponential extends to other cases. Suppose $F(z)$ is a holomorphic function with a power series expansion

$$F(z) = \sum_{k=0}^{\infty} a_k z^k. \tag{2.20}$$

If (2.20) 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

$$F(A) = \sum_{k=0}^{\infty} a_k A^k, \tag{2.21}$$

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 $\|A\| < R$,

$$F(A)G(A) = H(A). \tag{2.22}$$

We will see more examples of (2.21) in subsequent sections.

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

$$\log(I + tA) = tA - \frac{t^2}{2}A^2 + \frac{t^3}{3}A^3 - \cdots = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} t^k A^k. \tag{2.23}$$

We aim to prove that

$$e^{\log(I+tA)} = I + tA. \tag{2.24}$$

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

$$\begin{aligned} X(t) = \log(I+tA) \Rightarrow X'(t) &= A(I - tA + t^2A^2 - \cdots) \\ &= A(I+tA)^{-1}, \end{aligned} \tag{2.25}$$

as follows from (2.23) by differentiating term by term. For such $X(t)$, we see that $X(t)$ and $X(s)$ always commute, so it follows from (2.19) (or otherwise) that

$$\frac{d}{dt} e^{X(t)} = X'(t) e^{X(t)}. \tag{2.26}$$

Consequently, if we set

$$V(t) = (I+tA)^{-1} e^{\log(I+tA)}, \tag{2.27}$$

we have $V(0) = I$ and

$$V'(t) = -A(I+tA)^{-2} e^{\log(I+tA)} + A(I+tA)^{-2} e^{\log(I+tA)} = 0, \tag{2.28}$$

so (2.24) is established.

Note from (2.19) that

$$D\,\mathrm{Exp}(0)B = B, \tag{2.29}$$

i.e., $D\,\mathrm{Exp}(0)$ is the identity operator on $\mathrm{M}(n,\mathbb{R})$. 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

$$\mathrm{Exp} : \mathcal{O} \longrightarrow \Omega, \quad \text{diffeomorphically,} \tag{2.30}$$

hence there is a smooth inverse from Ω to $\mathcal{O}$. The results (2.23)–(2.24) provide an explicit formula for this inverse.

3. Quaternions and the group $\mathrm{Sp}(n)$

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:

$$\xi = a + bi + cj + dk, \tag{3.1}$$

with $a, b, c, d \in \mathbb{R}$. We call a the real part of ξ ($a = \mathrm{Re}\,\xi$) 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

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik, \tag{3.2}$$

and

$$i^2 = j^2 = k^2 = -1. \tag{3.3}$$

Otherwise stated, if we write

$$\xi = a + u, \quad a \in \mathbb{R}, \quad u \in \mathbb{R}^3, \tag{3.4}$$

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

$$\xi\eta = (a + u)(b + v) = (ab - u \cdot v) + av + bu + u \times v. \tag{3.5}$$

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.

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

$$\overline{\xi} = a - bi - cj - dk = a - u. \tag{3.6}$$

A calculation gives

$$\xi\overline{\eta} = (ab + u \cdot v) - av + bu - u \times v. \tag{3.7}$$

In particular,

$$\mathrm{Re}(\xi\overline{\eta}) = (\xi, \eta), \tag{3.8}$$

the right side denoting the Euclidean inner product on $\mathbb{R}^4$. Setting $\eta = \xi$ in (3.7) gives

$$\xi\overline{\xi} = |\xi|^2, \tag{3.9}$$

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

$$\xi^{-1} = |\xi|^{-2}\overline{\xi}. \tag{3.10}$$

A routine calculation gives

$$\overline{\xi\eta} = \overline{\eta}\,\overline{\xi}. \tag{3.11}$$

Hence

$$|\xi\eta|^2 = (\xi\eta)(\overline{\xi\eta}) = \xi\eta\overline{\eta}\overline{\xi} = |\eta|^2\xi\overline{\xi} = |\xi|^2|\eta|^2, \tag{3.12}$$

or

$$|\xi\eta| = |\xi|\,|\eta|. \tag{3.13}$$

Note that $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$ sits in $\mathbb{H}$ as a commutative subring, for which the properties (3.9) and (3.13) are familiar.

We consider the set of unit quaternions:

$$\mathrm{Sp}(1) = \{\xi \in \mathbb{H} : |\xi| = 1\}. \tag{3.14}$$

Using (3.10) and (3.13) it is clear that $\mathrm{Sp}(1)$ is a group under multiplication. It sits in $\mathbb{R}^4$ as the unit sphere S^3. We compare $\mathrm{Sp}(1)$ with the group $\mathrm{SU}(2)$, consisting of 2×2 complex matrices of the form

$$U = \begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix}, \quad \xi, \eta \in \mathbb{C}, \quad |\xi|^2 + |\eta|^2 = 1. \tag{3.15}$$

The group $\mathrm{SU}(2)$ is also diffeomorphic to S^3. Furthermore we have:

Proposition 3.1. *The groups* $\mathrm{SU}(2)$ *and* $\mathrm{Sp}(1)$ *are isomorphic under the correspondence*

$$U \mapsto \xi + j\eta, \tag{3.16}$$

for U as in (3.15).

Proof. The correspondence (3.16) is clearly bijective. To see that it is a homomorphism of groups, we calculate:

$$\begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix} \begin{pmatrix} \xi' & -\overline{\eta}' \\ \eta' & \overline{\xi}' \end{pmatrix} = \begin{pmatrix} \xi\xi' - \overline{\eta}\eta' & -\xi\overline{\eta}' - \overline{\eta}\overline{\xi}' \\ \eta\xi' + \overline{\xi}\eta' & -\eta\overline{\eta}' + \xi\overline{\xi}' \end{pmatrix}, \tag{3.17}$$

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

$$(3.18)\qquad \begin{aligned}(\xi + j\eta)(\xi' + j\eta') &= \xi\xi' + \xi j\eta' + j\eta\xi' + j\eta j\eta' \\ &= \xi\xi' - \overline{\eta}\eta' + j(\eta\xi' + \overline{\xi}\eta').\end{aligned}$$

Comparison of (3.17) and (3.18) verifies that (3.16) yields a homomorphism of groups.

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

$$(3.19)\qquad A = (a_{jk}) \in \mathrm{M}(n, \mathbb{H}), \quad a_{jk} \in \mathbb{H}.$$

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

$$(3.20)\qquad (A\xi)_j = \sum_k a_{jk}\xi_k.$$

Note that

$$(3.21)\qquad A : \mathbb{H}^n \longrightarrow \mathbb{H}^n$$

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

$$(3.22)\qquad (\xi b)_j = \xi_j b, \quad \xi \in \mathbb{H}^n,\ b \in \mathbb{H}.$$

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

$$(3.23)\qquad (AB)_{jk} = \sum_\ell a_{j\ell}b_{\ell k}.$$

We define a conjugation on $\mathrm{M}(n, \mathbb{H})$; with A given by (3.19),

$$(3.24)\qquad A^* = (\overline{a}_{kj}).$$

A calculation using (3.11) gives

$$(3.25)\qquad (AB)^* = B^*A^*.$$

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

$$(3.26)\qquad \mathrm{Sp}(n) = \{A \in \mathrm{M}(n, \mathbb{H}) : A^*A = I\}.$$

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 $A \in \mathrm{M}(n, \mathbb{H})$,

$$(3.27)\qquad A^*A = I \Longleftrightarrow AA^* = I.$$

In particular,

$$A \in \mathrm{Sp}(n) \Longleftrightarrow A^* \in \mathrm{Sp}(n) \Longleftrightarrow A^{-1} \in \mathrm{Sp}(n). \tag{3.28}$$

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

$$(AB)^* AB = B^* A^* AB = B^* B = I. \tag{3.29}$$

Hence $\mathrm{Sp}(n)$, defined by (3.26), is a group. We claim that (3.26) 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, 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

$$\langle \xi, \eta \rangle = \sum_j \overline{\eta}_j \xi_j. \tag{3.30}$$

From (3.8) we have

$$\mathrm{Re}\langle \xi, \eta \rangle = (\xi, \eta), \tag{3.31}$$

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

$$\begin{aligned} \langle A\xi, \eta \rangle &= \sum_{j,k} \overline{\eta}_j a_{jk} \xi_k \\ &= \sum_{j,k} \overline{\overline{a}_{jk} \eta_j}\, \xi_k \\ &= \langle \xi, A^* \eta \rangle. \end{aligned} \tag{3.32}$$

Hence

$$\langle A\xi, A\eta \rangle = \langle \xi, A^* A\eta \rangle. \tag{3.33}$$

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 (3.30). Given (3.31), we have

$$\mathrm{Sp}(n) \hookrightarrow \mathrm{O}(4n). \tag{3.34}$$

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

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,

$$(A,B) = \operatorname{Tr} AB^* = \operatorname{Tr} B^*A, \tag{4.1}$$

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

$$L_g, R_g : \mathrm{M}(n,\mathbb{C}) \longrightarrow \mathrm{M}(n,\mathbb{C}), \quad L_g X = gX, \quad R_g X = Xg. \tag{4.2}$$

Clearly each such map is a linear isometry on $\mathrm{M}(n,\mathbb{C})$, 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,

$$dV(x) = \sqrt{\det(h_{jk})}\, dx_1 \cdots dx_N. \tag{4.3}$$

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:

$$\int_G f(g)\, dg = \frac{1}{V(G)} \int_G f\, dV. \tag{4.4}$$

We have left invariance

$$\int_G f(hg)\, dg = \int_G f(g)\, dg \tag{4.5}$$

and right invariance

$$\int_G f(gh)\, dg = \int_G f(g)\, dg, \tag{4.6}$$

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 (4.2), we have

$$DL_{g^{-1}},\ DR_{g^{-1}} : T_gG \longrightarrow T_eG \approx \mathfrak{g}. \tag{4.7}$$

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

$$\begin{aligned}\langle U, V\rangle_\ell &= \langle DL_{g^{-1}}U, DL_{g^{-1}}V\rangle_{\mathfrak{g}},\\ \langle U, V\rangle_r &= \langle DR_{g^{-1}}U, DR_{g^{-1}}V\rangle_{\mathfrak{g}}.\end{aligned} \tag{4.8}$$

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 (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 $h \in G$,

$$\int_G f(hg)\, dV_\ell(g) = \int_G f(g)\, dV_\ell(g), \quad \int_G f(gh)\, dV_r(g) = \int_G f(g)\, dV_r(g). \tag{4.9}$$

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

$$I_\ell(f) = \int_G f(g)\, dV_\ell(g), \tag{4.10}$$

so right translation by h yields

$$I_\ell^h(f) = \int_G f(gh)\, dV_\ell(g). \tag{4.11}$$

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

$$I_\ell^k(f) = \alpha(h) I_\ell(f), \tag{4.12}$$

for a map

$$\alpha : G \longrightarrow (0, \infty). \tag{4.13}$$

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

$$\alpha(h_1h_2) = \alpha(h_1)\alpha(h_2), \quad \forall\, h_j \in G, \tag{4.14}$$

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 (4.4), so

$$\int_G 1\, dg = 1. \tag{4.15}$$

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. 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

$$\omega_\ell(e) = \omega_e, \quad L_g^* \omega_\ell = \omega_\ell, \quad \forall\, g \in G, \tag{4.16}$$

and a unique N-form ω_r on G such that

$$\omega_r(e) = \omega_e, \quad R_g^* \omega_r = \omega_r, \quad \forall\, g \in G. \tag{4.17}$$

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 (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

$$R_h^* \omega_\ell = \alpha(h) \omega_\ell, \tag{4.18}$$

for all $h \in G$, with α as in (4.12). From this point of view, (4.14) follows from the identity

$$R_{h_1h_2}^* = R_{h_2}^* R_{h_1}^*. \tag{4.19}$$

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

$$\int_G f(g^{-1})\,dV_\ell = I(f) \tag{4.20}$$

is right-invariant, i.e., equal to $\int_G f(g)\,dV_r(g)$ for some right-invariant Haar measure dV_r. If G is compact and (4.15) holds, then $I(1) = 1$, and we have

$$\int_G f(g^{-1})\,dg = \int_G f(g)\,dg. \tag{4.21}$$

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

$$dV_\ell(X) = \varphi(X)\,dX, \tag{4.22}$$

where

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}, \quad dX = dx_{11} \cdots dx_{nn}, \tag{4.23}$$

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

$$L_g : \mathrm{Gl}(n, \mathbb{R}) \longrightarrow \mathrm{Gl}(n, \mathbb{R}), \quad L_g X = gX, \tag{4.24}$$

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

$$\begin{aligned} \int u(X)\varphi(X)\,dX &= \int u(L_g X)\varphi(L_g X)\,|\det DL_g(X)|\,dX \\ &= \int u(gX)\varphi(gX)\,|\det g|^n\,dX. \end{aligned} \tag{4.25}$$

The left invariance of dV_ℓ demands that this equal

$$\int u(gX)\varphi(X)\,dX. \tag{4.26}$$

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

$$\varphi(gX) = |\det g|^{-n}\varphi(X), \quad \forall\, g, X \in \mathrm{Gl}(n, \mathbb{R}). \tag{4.27}$$

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

$$\varphi(g) = |\det g|^{-n}, \tag{4.28}$$

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

$$dV_\ell(X) = |\det X|^{-n}\,dX. \tag{4.29}$$

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

5. Representations of a group

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

$$\pi : G \longrightarrow \mathrm{End}(V) \tag{5.1}$$

such that

$$\pi(e) = I, \quad \pi(gg') = \pi(g)\pi(g'), \quad \forall\, g, g' \in G. \tag{5.2}$$

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

$$(\pi(g)v, \pi(g)w) = (v, w), \quad \forall\, g \in G,\ v, w \in V, \tag{5.3}$$

we say π is an orthogonal representation. If V is complex with a Hermitian inner product and (5.3) holds, we say π is a unitary representation. Representations of a *compact* Lie group are unitarizable, as follows.

Proposition 5.1. *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 (5.3) holds.*

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

$$(u, v) = \int_G ((\pi(g)u, \pi(g)v))\, dg. \tag{5.4}$$

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

$$\begin{aligned} (\pi(h)u, \pi(h)v) &= \int_G ((\pi(g)\pi(h)u, \pi(g)\pi(h)v))\, dg \\ &= \int_G ((\pi(gh)u, \pi(gh)v))\, dg \\ &= (u, v), \end{aligned} \tag{5.5}$$

by right invariance of Haar measure on G.

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 5.2. *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.

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

Lemma 5.3. *Suppose π and λ are finite-dimensional, irreducible unitary representations of G on V and W. Assume $A : V \to W$ satisfies*

$$A\pi(g) = \lambda(g)A, \quad \forall g \in G. \tag{5.6}$$

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 5.3 follows from the next lemma.

Lemma 5.4. *Suppose π is a finite-dimensional, irreducible unitary representation of G on V. Assume $B : V \to V$ satisfies*

$$B\pi(g) = \pi(g)B, \quad \forall\ g \in G. \tag{5.7}$$

Then B is a scalar multiple of the identity.

Proof. Set $B = B_1 + iB_2,\ B_j^* = B_j$. It follows from (5.7) and unitarity that

$$B_j\pi(g) = \pi(g)B_j, \quad \forall\ g \in G. \tag{5.8}$$

Now each B_j is diagonalizable, and

$$B_jv = av \Rightarrow B_j\pi(g)v = \pi(g)B_jv = a\pi(g)v, \quad \forall\ g \in G, \tag{5.9}$$

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

Finally, we set $B = A^*A$ to prove the last assertion in Lemma 5.3. In fact, if λ and π are unitary, (5.6) 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 $a \in \mathbb{C}$.

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.

We end with the following result, for which there will be an important analogue in the next section.

Proposition 5.5. *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*

$$Pv = \int_G \pi(g)v\,dg. \tag{5.10}$$

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

The proof consists of four easy pieces:

$$\pi(g)Pv = Pv, \quad \forall\ g \in G, \tag{5.11}$$

$$P^* = \int \pi(g^{-1})\,dg = P, \tag{5.12}$$

$$P^2 = \iint \pi(g)\pi(h)\,dg\,dh = \iint \pi(gh)\,dg\,dh = P, \tag{5.13}$$

$$\pi(g)v = v\ \forall\, g \Longrightarrow Pv = v. \tag{5.14}$$

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

6. 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

$$P(A) = \int_G \lambda(g) A \pi(g)^{-1}\, dg. \tag{6.1}$$

It is readily verified that

$$\lambda(g) P(A) \pi(g)^{-1} = P(A), \quad \forall\, g \in G. \tag{6.2}$$

In other words, $P(A)$ intertwines π and λ. Now Schur's lemma, established in §5, gives the following:

$$\begin{aligned} \pi \text{ not} \approx \lambda &\Longrightarrow P(A) = 0, \quad \forall A, \\ \pi = \lambda &\Longrightarrow P(A) = c_\pi(A) I, \end{aligned} \tag{6.3}$$

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

$$\int_G \pi(g) A \pi(g)^*\, dg = d_\pi^{-1} (\mathrm{Tr}\, A) I. \tag{6.4}$$

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

$$\begin{aligned} \sum_{k,\ell} \int_G \pi(g)_{jk} A_{k\ell} \overline{\pi(g)}_{m\ell}\, dg &= d_\pi^{-1}\, \delta_{jm}\ \mathrm{Tr}\, A \\ &= d_\pi^{-1}\, \delta_{jm} \sum_{k,\ell} \delta_{k\ell} A_{k\ell}, \end{aligned} \tag{6.5}$$

hence

$$\int_G \pi(g)_{jk} \overline{\pi(g)}_{m\ell}\, dg = d_\pi^{-1}\, \delta_{jm}\, \delta_{k\ell}. \tag{6.6}$$

These are Weyl orthogonality relations. They are complemented by

$$\int_G \pi(g)_{jk} \overline{\lambda(g)}_{m\ell}\, dg = 0, \quad \pi \text{ not} \approx \lambda, \tag{6.7}$$

which follows from the first part of (6.3).

7. The Peter-Weyl theorem

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 §6 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 7.1. *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 sketch a proof of this 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 $C(G)$ 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:

$$\mathcal{E} \ \text{ separates points of } \ G, \tag{7.1}$$

$$u \in \mathcal{E} \Longrightarrow \overline{u} \in \mathcal{E}, \tag{7.2}$$

$$u, v \in \mathcal{E} \Longrightarrow uv \in \mathcal{E}. \tag{7.3}$$

Of these conditions, (7.1) follows directly from the hypothesis $G \subset \mathrm{U}(N)$. As for (7.2), 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 $V_\alpha \otimes V_\beta$ can be decomposed into irreducibles, by Proposition 5.2, and this gives (7.3).

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

Proposition 7.2. *If* G *is a compact Lie group, then there is an injective representation*

$$\rho : G \longrightarrow \mathrm{U}(N). \tag{7.4}$$

We say G *has a faithful unitary representation.*

Actually, we will prove this result in §11, 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

$$u = \sum_{\alpha \in \mathcal{I}} d_\alpha^{1/2} \sum_{j,k} \hat{u}_{jk}(\alpha)\, \pi^\alpha_{jk}(g), \tag{7.5}$$

where

$$\hat{u}_{jk}(\alpha) = d_\alpha^{1/2} \int_G u(g)\, \overline{\pi^\alpha_{jk}(g)}\, dg, \tag{7.6}$$

the convergence in (7.5) holding in L^2-norm. Let us also set

$$\mathcal{P}_\alpha u = d_\alpha^{1/2} \sum_{j,k} \hat{u}_{jk}(\alpha)\pi^\alpha_{jk}(g), \tag{7.7}$$

the orthogonal projection of u onto the space

$$\mathcal{V}_\alpha = \operatorname{span}\{\pi^\alpha_{jk} : 1 \le j,k \le d_\alpha\}. \tag{7.8}$$

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

$$u = \sum_{\alpha\in\mathcal{I}} \mathcal{P}_\alpha u, \tag{7.9}$$

convergence in L^2-norm.

Another way to write (7.7) is as

$$\mathcal{P}_\alpha u(g) = d_\alpha^{1/2} \operatorname{Tr}\bigl(\hat{u}(\alpha)^t \pi^\alpha(g)\bigr), \tag{7.10}$$

where

$$\hat{u}(\alpha) = d_\alpha^{1/2} \int_G u(g)\overline{\pi}^\alpha(g)\, dg. \tag{7.11}$$

Here we can define $\overline{\pi}^\alpha$ as the representation of G on V'_α given by

$$\overline{\pi}^\alpha(g) = \pi^\alpha(g^{-1})^t : V'_\alpha \longrightarrow V'_\alpha, \tag{7.12}$$

so $\hat{u}(\alpha) : V'_\alpha \to V'_\alpha$. If $(\pi^\alpha_{jk}(g))$ is the matrix representation of $\pi^\alpha(g)$ with respect to an orthonormal basis of V_α, then $(\overline{\pi^\alpha_{jk}(g)})$ 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

$$C : V_\alpha \longrightarrow V'_\alpha, \quad (u,v) = \langle u, Cv\rangle, \tag{7.13}$$

and a straightforward calculation gives

$$\overline{\pi}^\alpha(g) = C\pi^\alpha(g)C^{-1}. \tag{7.14}$$

We also note that for a unitary representation π^α we have (as asserted shortly below (7.3))

$$\pi^\alpha \text{ irreducible} \Longrightarrow \overline{\pi}^\alpha \text{ irreducible}. \tag{7.15}$$

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 $\pi^\alpha(g)$.

8. 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

$$\operatorname{Tr} \rho(g) = \chi_\rho(g), \tag{8.1}$$

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 $\alpha \neq \beta$,

$$\int_G \chi_\alpha(g)\overline{\chi_\beta(g)}\, dg = \sum_{j,k} \int_G \pi^\alpha_{jj}(g)\overline{\pi^\beta_{kk}(g)}\, dg = 0, \tag{8.2}$$

as a consequence of (6.7). On the other hand, using (6.6) we have

$$\begin{aligned} \int_G \chi_\alpha(g)\overline{\chi_\alpha(g)}\, dg &= \sum_{j,k} \int_G \pi^\alpha_{jj}(g)\overline{\pi^\alpha_{kk}(g)}\, dg \\ &= \sum_{j,k} d_\alpha^{-1}\delta_{jk} \\ &= 1. \end{aligned} \tag{8.3}$$

Hence $\{\chi_\alpha : \alpha \in \mathcal{I}\}$ is an orthonormal set in $L^2(G)$. We have more:

Proposition 8.1. *The set $\{\chi_\alpha : \alpha \in \mathcal{I}\}$ is an orthonormal basis of*

$$L^2_{\mathcal{C}}(G) = \{u \in L^2(G) : u \text{ is central}\}. \tag{8.4}$$

Proof. Take $u \in L^2_{\mathcal{C}}(G)$. By (7.9)–(7.10) we can write

$$u = \sum_{\alpha\in\mathcal{I}} \mathcal{P}_\alpha u, \quad \mathcal{P}_\alpha u = d_\alpha^{1/2} \operatorname{Tr}\big(\hat{u}(\alpha)^t\, \pi^\alpha(g)\big) \in \mathcal{V}_\alpha, \tag{8.5}$$

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 $h \in G$,

$$u_h(g) = u(h^{-1}gh) \Longrightarrow \hat{u}_h(\alpha) = \overline{\pi}^\alpha(h)\hat{u}(\alpha)\overline{\pi}^\alpha(h)^{-1}. \tag{8.6}$$

Hence

$$u \text{ central} \Longrightarrow \hat{u}(\alpha)\overline{\pi}^{\alpha}(h) = \overline{\pi}^{\alpha}(h)\hat{u}(\alpha), \quad \forall\, h \in G. \tag{8.7}$$

Thus, by Schur's lemma, $\hat{u}(\alpha)$ is a scalar multiple of the identity. Taking traces in (7.11) gives $\operatorname{Tr}\hat{u}(\alpha) = d_\alpha^{1/2}\int_G u(g)\overline{\chi_\alpha(g)}\,dg$, which gives

$$\begin{aligned} u \text{ central} &\Longrightarrow \hat{u}(\alpha) = c^\alpha(u)I, \quad c^\alpha(u) = d_\alpha^{-1/2}(u,\chi_\alpha)_{L^2(G)} \\ &\Longrightarrow \mathcal{P}_\alpha u(g) = (u,\chi_\alpha)_{L^2(G)}\,\chi_\alpha(g), \end{aligned} \tag{8.8}$$

finishing the proof.

We now establish a generalization of Proposition 5.5.

Proposition 8.2. *Let G be a compact Lie group and ρ a unitary representation of G on a finite-dimensional vector space V. Set*

$$P_\alpha = d_\alpha \int_G \overline{\chi_\alpha(g)}\,\rho(g)\,dg. \tag{8.9}$$

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

Proof. As shown in §5, one has an orthogonal direct sum decomposition

$$V = V_1 \oplus \cdots \oplus V_K, \tag{8.10}$$

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

$$\begin{aligned} u \in V_j,\ \beta_j \neq \alpha &\Longrightarrow P_\alpha u = 0, \\ u \in V_j,\ \beta_j = \alpha &\Longrightarrow P_\alpha u = u. \end{aligned} \tag{8.11}$$

The first part of (8.11) follows from the identity

$$\int_G \overline{\chi_\alpha(g)}\,\pi^\beta_{k\ell}(g)\,dg = 0 \ \Longleftarrow \beta \neq \alpha, \tag{8.12}$$

a consequence of (6.7). The second part of (8.11) follows from the identity

$$d_\alpha \int_G \overline{\chi_\alpha(g)}\,\pi^\alpha_{k\ell}(g)\,dg = \delta_{k\ell}, \tag{8.13}$$

a consequence of (6.6).

The number of factors V_j in (8.10) 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.,

$$\mu(\pi^\alpha,\rho) = d_\alpha^{-1}\operatorname{Tr}P_\alpha = \int_G \chi_\rho(g)\overline{\chi_\alpha(g)}\,dg. \tag{8.14}$$

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

$$\rho \approx \lambda \Longleftrightarrow \chi_\rho = \chi_\lambda. \tag{8.15}$$

9. 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

$$\int_G f(g)\,dg = \frac{1}{o(G)} \sum_{g\in G} f(g). \tag{9.1}$$

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

$$L(g)u(x) = u(g^{-1}x), \quad g, x \in G,\ u \in L^2(G), \tag{9.2}$$

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 §7 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 (7.8), we have

$$L^2(G) = \bigoplus_{\alpha\in\mathcal{I}} \mathcal{V}_\alpha, \quad \dim \mathcal{V}_\alpha = d_\alpha^2, \tag{9.3}$$

and hence

$$\sum_{\alpha\in\mathcal{I}} d_\alpha^2 = o(G). \tag{9.4}$$

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

$$\dim L^2_{\mathcal{C}}(G) = o(\mathcal{C}), \tag{9.5}$$

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

$$o(\mathcal{I}) = o(\mathcal{C}), \tag{9.6}$$

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

Let us illustrate some of these results on 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,

$$\text{sgn} : S_n \longrightarrow \{\pm 1\}. \tag{9.7}$$

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

$$D_n(x) = \prod_{1 \le j < k \le n} (x_j - x_k). \tag{9.8}$$

Then, for $\sigma \in S_n$,

$$\prod_{1 \le j < k \le n} (x_{\sigma(j)} - x_{\sigma(k)}) = \text{sgn}(\sigma)\, D_n(x). \tag{9.9}$$

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

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

$$\rho_n(\sigma) e_j = e_{\sigma(j)}, \tag{9.10}$$

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

$$\rho_n(\sigma)(e_1 + \cdots + e_n) = e_1 + \cdots + e_n, \quad \forall\, \sigma \in S_n. \tag{9.11}$$

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

$$V_{n-1} = \{u \in \mathbb{C}^n : u_1 + \cdots + u_n = 0\}. \tag{9.12}$$

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

Lemma 9.1. *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

$$e_1 - e_2 \in W, \tag{9.13}$$

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

$$(v_1, v_2, \dots, v_n), \quad v_1 \ne v_2. \tag{9.14}$$

Then

$$(v_1, v_2, v_3, \dots, v_n) - (v_2, v_1, v_3, \dots, v_n) = (v_1 - v_2, v_2 - v_1, 0, \dots, 0) \in W \tag{9.15}$$

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_S^n, 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

$$1,\ \mathrm{sgn},\ \pi_S^n \tag{9.16}$$

exhausts the set of irreducible representations of S_3. In fact, in view of (9.4), the dimension check

$$o(S_3) = 6 = 1^2 + 1^2 + 2^2 \tag{9.17}$$

verifies this.

The group S_4 has the irreducible representations (9.16) and a couple more. One is given by

$$\pi_Q^4(\sigma) = \mathrm{sgn}(\sigma)\, \pi_S^4(\sigma), \tag{9.18}$$

acting on V_3. Since the representations π_S^4 and π_Q^4 are three-dimensional representations, we have

$$\det \pi_Q^4(\sigma) = \mathrm{sgn}(\sigma)\, \det \pi_S^4(\sigma), \tag{9.19}$$

so they cannot be equivalent. (By contrast, $\mathrm{sgn} \cdot \pi_S^3$ is equivalent to π_S^3.)

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

$$\beta : S_4 \longrightarrow S_3. \tag{9.20}$$

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 (9.20). Then the representation

$$\pi_S^3 \circ \beta \tag{9.21}$$

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

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

$$\det \pi_S^4(\sigma) = \mathrm{sgn}(\sigma), \quad \det \pi_Q^4(\sigma) = 1. \tag{9.22}$$

Hence π_Q^4 acts as a group of *rotations* on $V_3^R \approx \mathbb{R}^3$. In fact, we claim π_Q^4 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 9.1, 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 $-\mathcal{T}$. 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

$$\pi_S^4(\sigma) = \mathrm{sgn}(\sigma)\, \pi_Q^4(\sigma), \tag{9.23}$$

which is another way of putting (9.18).

We note another perspective on (9.20). Namely π_Q^4 acts on the 3-point set consisting of opposite pairs of faces of the cube $\mathcal{Q}$.

10. 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

$$u * v(x) = \int_G u(g)v(g^{-1}x)\,dg. \tag{10.1}$$

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:

$$u * (v * w) = (u * v) * w. \tag{10.2}$$

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

$$\pi(u) = \int_G u(g)\,\pi(g)\,dg. \tag{10.3}$$

The following relates (10.3) to the convolution product.

Proposition 10.1. *We have*

$$\pi(u * v) = \pi(u)\pi(v). \tag{10.4}$$

Proof. The definitions give

$$\begin{aligned} \pi(u * v) &= \int (u * v)(g)\,\pi(g)\,dg \\ &= \iint u(h)v(h^{-1}g)\pi(g)\,dg\,dh \\ &= \iint u(h)\pi(h)\,v(h^{-1}g)\pi(h^{-1}g)\,dg\,dh \\ &= \int u(h)\pi(h)\,dy\;\pi(v) \\ &= \pi(u)\pi(v), \end{aligned} \tag{10.5}$$

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 10.2. *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

$$
\begin{aligned}
v * u(x) &= \int v(g)u(g^{-1}x)\,dg \\
&= \int v(g)u(xg^{-1})\,dg \quad \text{(if } u \text{ is central)} \\
&= \int v(h^{-1}x)u(h)\,dh \\
&= u * v(x).
\end{aligned} \tag{10.6}
$$

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

$$
C_g u(x) = u(g^{-1}xg), \tag{10.7}
$$

and note that

$$
\begin{aligned}
\pi(C_h u) &= \int \pi(g)u(h^{-1}gh)\,dg \\
&= \int \pi(hxh^{-1})u(x)\,dx \\
&= \pi(h)\pi(u)\pi(h)^{-1},
\end{aligned} \tag{10.8}
$$

so

$$
u \ \text{ central} \Longrightarrow \pi(u)\pi(h) = \pi(h)\pi(u), \quad \forall\ h \in G. \tag{10.9}
$$

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

$$
\begin{aligned}
&u \ \text{ central} \Longrightarrow \pi^\alpha(u) = \sigma^\alpha(u)I, \\
&\sigma^\alpha(u) = \frac{1}{d_\alpha}\int \chi_\alpha(g)u(g)\,dg,
\end{aligned} \tag{10.10}
$$

where $\chi_\alpha(g) = \operatorname{Tr}\pi^\alpha(g)$. Compare (8.8), noting that $\hat{u}(\alpha) = d_\alpha^{1/2}\overline{\pi}^\alpha(u)$.

Generalizing (10.7)–(10.8), we note that

$$
G_{g,h}u(x) = u(g^{-1}xh) \Rightarrow \pi(C_{g,h}u) = \pi(g)\pi(u)\pi(h)^{-1}. \tag{10.11}
$$

The following is a useful formula for the projection $\mathcal{P}_\alpha$ defined in (7.7).

Proposition 10.3. *For $u \in L^2(G)$,*

$$\mathcal{P}_\alpha u = d_\alpha\, \chi_\alpha * u = d_\alpha\, u * \chi_\alpha. \tag{10.12}$$

Proof. By Proposition 10.2, the last two functions in (10.12) are equal. By (7.10) we have

$$\begin{aligned} \mathcal{P}_\alpha u(g) &= d_\alpha^{1/2}\ \mathrm{Tr}\big(\hat{u}(\alpha)^t \pi^\alpha(g)\big) \\ &= d_\alpha\ \mathrm{Tr}\big(\overline{\pi}^\alpha(u)^t \pi^\alpha(g)\big). \end{aligned} \tag{10.13}$$

Meanwhile

$$\begin{aligned} u * \chi_\alpha(g) &= \mathrm{Tr} \int_G u(h)\pi^\alpha(h^{-1}g)\, dh \\ &= \mathrm{Tr}\Big[\int_G u(h)\pi^\alpha(h^{-1})\, dh\ \pi^\alpha(g)\Big]. \end{aligned} \tag{10.14}$$

Finally,

$$\overline{\pi}^\alpha(h) = \pi^\alpha(h^{-1})^t \Rightarrow \overline{\pi}^\alpha(u)^t = \int_G u(h)\pi^\alpha(h^{-1})\, dh. \tag{10.15}$$

Then (10.13)–(10.15) yield (10.12).

We define the following involution on functions on G:

$$u^*(g) = \overline{u(g^{-1})}, \tag{10.16}$$

and note that

$$(u * v)^* = v^* * u^*, \tag{10.17}$$

and if π is a unitary representation of G,

$$\pi(u^*) = \pi(u)^*, \tag{10.18}$$

as is readily checked.

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

$$K_f : L^2(G) \longrightarrow L^2(G), \quad K_f u(x) = f * u(x) = \int f(g)u(g^{-1}x)\, dg. \tag{10.19}$$

The estimate

$$\|K_f u\|_{L^2} \le \|f\|_{L^1} \|u\|_{L^2} \tag{10.20}$$

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

$$
\begin{aligned}
(K_f u, v) &= \iint f(g)u(g^{-1}x)\overline{v(x)}\,dg\,dx \\
&= \iint f(xy^{-1})u(y)\overline{v(x)}\,dy\,dx \\
&= (u, K_{f^*}v),
\end{aligned}
\tag{10.21}
$$

or

$$
K_f^* = K_{f^*}. \tag{10.22}
$$

Also it follows from (10.2) that

$$
K_{f*g} = K_f K_g. \tag{10.23}
$$

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

$$
L(g)u(x) = u(g^{-1}x), \quad R(g)u(x) = u(xg). \tag{10.24}
$$

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

$$
L(f)u(x) = \int f(g)u(g^{-1}x)\,dg = f * u(x) = K_f u(x), \tag{10.25}
$$

and

$$
R(f)u(x) = \int f(g)u(xg)\,dg = \int f(hx^{-1})u(h)\,dh = u * \check{f}(x), \tag{10.26}
$$

where

$$
\check{f}(g) = f(g^{-1}). \tag{10.27}
$$

Then (10.23) becomes $L(f*g) = L(f)L(g)$, a result parallel to (10.4). Similarly one has $R(f*g) = R(f)R(g)$.

Here is another useful result.

Proposition 10.4. *For all $f \in L^1(G)$, we have*

$$
R(g)K_f = K_f R(g), \quad \forall\ g \in G. \tag{10.28}
$$

In addition,

$$
f \quad \textit{central} \Longrightarrow L(g)K_f = K_f L(g), \quad \forall\ g \in G. \tag{10.29}
$$

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

$$\forall\, u \in L^2(G), \quad L(g)u \text{ and } R(g)u \ \text{ are continuous from } G \text{ to } L^2(G). \tag{10.30}$$

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 (9.1). 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

$$u = \sum_{g \in G} u(g) g. \tag{10.31}$$

Then convolution is given by

$$\begin{aligned} u * v &= \frac{1}{o(G)} \sum_{g,h \in G} u(g)v(h)\, gh \\ &= \frac{1}{o(G)} \sum_{g,x \in G} u(g)v(g^{-1}x)\, x, \end{aligned} \tag{10.32}$$

which is consistent with (10.1).

11. Approximate identities and the Peter-Weyl theorem in general

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 §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 avarage 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 $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\psi(g) = \varphi(d(g, e)) \ \text{ is a central function on } \ G. \tag{11.1}$$

Also, with ψ^* defined as in (10.12), 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

$$\psi_\nu(g) = \varphi(\nu d(g, e)), \quad \nu \geq 1. \tag{11.2}$$

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

$$\Psi_\nu(g) = A_\nu^{-1} \psi_\nu(g), \quad A_\nu = \int \psi_\nu(g)\, dg, \tag{11.3}$$

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

$$\Phi_\nu(g) = \Psi_\nu * \Psi_\nu(g). \tag{11.4}$$

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

$$C_\nu u = \Psi_\nu * u, \quad K_\nu u = \Phi_\nu * u. \tag{11.5}$$

Proposition 11.1. *The operators C_ν and K_ν are approximate identities. That is, as $\nu \to \infty$,*

$$u \in C(G) \Longrightarrow C_\nu u \to u \ \text{ and } \ K_\nu u \to u \ \text{ uniformly.} \tag{11.6}$$

Also

$$u \in L^2(G) \Longrightarrow C_\nu u \to u \ \text{ and } \ K_\nu u \to u \ \text{ in } \ L^2\text{-norm.} \tag{11.7}$$

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

$$\|C_\nu u - u\|_{\sup} \le \omega(1/\nu), \quad \|K_\nu u - u\|_{\sup} \le \omega(2/\nu). \tag{11.8}$$

This gives (11.6). The result (11.7) for C_ν follows from

$$\|C_\nu u\|_{L^2} \le \|u\|_{L^2}, \quad \|u\|_{L^2} \le \|u\|_{\sup}, \quad C(G) \ \text{ dense in } \ L^2(G). \tag{11.9}$$

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 $\nu \ge N$,

$$\begin{aligned} \|C_\nu u - u\|_{L^2} &= \|C_\nu u - C_\nu v + C_\nu v - v + v - u\|_{L^2} \\ &\le \|C_\nu(u - v)\|_{L^2} + \|C_\nu v - v\|_{\sup} + \|v - u\|_{L^2} \\ &< 3\varepsilon, \end{aligned} \tag{11.10}$$

giving (11.7) 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 (10.22)–(10.23) that

$$C_\nu = C_\nu^*, \quad K_\nu = C_\nu^2; \tag{11.11}$$

hence

$$(K_\nu u, u) = \|C_\nu u\|_{L^2}^2 \ge 0, \quad \forall\, u \in L^2(G), \tag{11.12}$$

i.e., K_ν is a positive semi-definite self-adjoint operator on $L^2(G)$.

Proposition 11.2. *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,

$$C_\nu : L^2(G) \longrightarrow \mathrm{Lip}(G). \tag{11.13}$$

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 11.3. *Let $K : L^2(G) \to L^2(G)$ be compact and self-adjoint, and assume*

$$KR(g) = R(g)K, \quad \forall\, g \in G, \tag{11.14}$$

where $R(g)$ denotes the right-regular representation of G on $L^2(G)$, given by (10.24). 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 (11.14), $R(g) : E_\lambda \to E_\lambda$. 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,

$$R(g)u_k = U^{-1}\pi^\alpha(g)e_k = U^{-1}\sum_j \pi^\alpha_{jk}(g)e_j, \tag{11.15}$$

or

$$u_k(xg) = \sum_j u_j(x)\,\pi^\alpha_{jk}(g), \quad \forall\, x, g \in G. \tag{11.16}$$

Taking $x = e$ gives

$$u_k(g) = \sum_j u_j(e)\,\pi^\alpha_{jk}(g), \tag{11.17}$$

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 11.4. *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 (11.14), 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 11.3. The task that remains is to construct such an injective operator K.

To this end, set

$$K = \sum_{\nu \ge 0} 2^{-\nu} K_\nu, \tag{11.18}$$

with K_ν as in (11.4)–(11.5). Then K is a norm limit of compact operators, hence compact, and also clearly positive, self-adjoint. By Proposition 10.4 each K_ν has the property (11.14), 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 11.1 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 $\{\pi^\alpha_{jk} : 1 \leq j, k \leq d_\alpha\}$, and $\mathcal{P}_\alpha$ the orthogonal projection of $L^2(G)$ on $\mathcal{V}_\alpha$. We went 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 §7 under the hypothesis $G \subset U(N)$). The following will be useful for this.

Proposition 11.5. *For any $u \in L^1(G)$,*

$$K_u : \mathcal{V}_\alpha \longrightarrow \mathcal{V}_\alpha. \tag{11.19}$$

Furthermore, for any $v \in L^2(G)$,

$$\mathcal{P}_\alpha(u * v) = u * (\mathcal{P}_\alpha v). \tag{11.20}$$

Proof. Recall from Proposition 10.3 that

$$\mathcal{P}_\alpha v = d_\alpha\, \chi_\alpha * v. \tag{11.21}$$

Hence, by Proposition 10.2,

$$\begin{aligned} \mathcal{P}_\alpha(u * v) &= d_\alpha\, \chi_\alpha * u * v \\ &= d_\alpha\, u * \chi_\alpha * v, \end{aligned} \tag{11.22}$$

and we have (11.19)–(11.20).

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

$$\Pi_N\, f = \sum_{|\alpha| \leq N} \mathcal{P}_\alpha f. \tag{11.23}$$

Then the content of Proposition 11.4 is that, as $N \to \infty$,

$$f \in L^2(G) \Longrightarrow \Pi_N f \to f \quad \text{in} \quad L^2\text{-norm}. \tag{11.24}$$

Here is a result on uniform convergence.

Proposition 11.6. *Assume $f \in C(G)$ has the form*

$$f = u * v, \quad u, v \in L^2(G). \tag{11.25}$$

Then, as $N \to \infty$,

$$\Pi_N f \to f \quad \text{uniformly on } G. \tag{11.26}$$

Proof. It follows from (11.20) that

$$\Pi_N f = u * (\Pi_N v). \tag{11.27}$$

Now convolution yields a continuous bilinear map

$$L^2(G) \times L^2(G) \longrightarrow C(G), \tag{11.28}$$

so using $\Pi_N v \to v$ in L^2-norm in (11.27) yields (11.26).

From Proposition 11.1 it follows that the set of functions of the form (11.25) is dense in $C(G)$, so we have:

Corollary 11.7. *The linear span 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 11.8. *Every compact Lie group has a faithful finite-dimensional unitary representation.*

To see this, let

$$K_\alpha = \{g \in G : \pi^\alpha(g) = I\}. \tag{11.29}$$

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 11.7, for any $g \in G,\ g \ne e$, there exists $\alpha \in \mathcal{I}$ such that $\pi^\alpha(g) \ne I$. 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 $\pi^\alpha(h) \ne I$ 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

$$\bigcap_{\alpha\in\mathcal{S}} K_\alpha \subset \mathcal{O}. \tag{11.30}$$

Thus the proof of Proposition 11.8 is completed by the following assertion.

Proposition 11.9. *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 §12 we will prove

$$D\,\mathrm{Sq}(e) = 2I. \tag{11.31}$$

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

$$\mathrm{Sq}(x) = 2x + R(x), \quad |R(x)| \le C|x|^2. \tag{11.32}$$

We use the Euclidean norm $|x|$. It follows that, in this coordinate system,

$$|x| < \frac{1}{2C} \Longrightarrow |\,\mathrm{Sq}(x)| > \frac{3}{2}|x|. \tag{11.33}$$

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 $\mathcal{O}$.

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:

$$\pi(f)u = \int_G f(g)\pi(g)u\,dg, \quad \pi(f) : V \to V. \tag{11.34}$$

Here we will use left-invariant Haar measure. Note that, for any $h \in G$,

$$\pi(h)\pi(f)u = \int f(h^{-1}g)\pi(g)u\,dg. \tag{11.35}$$

From this it is easy to see that, for all $u \in V$,

$$f \in C_0^\infty(G) \Rightarrow \pi(h)\pi(f)u \ \text{ is a smooth } V\text{-valued function of } h, \tag{11.36}$$

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 11.1) we have

$$\pi(f_\nu)u \longrightarrow u \ \text{ in } \ V, \quad \forall\, u \in V. \tag{11.37}$$

We hence have:

Proposition 11.10. *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.*

We close this section with the following corollary to the Peter-Weyl theorem, which will prove useful later.

Proposition 11.11. *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*

$$\pi(g) = \pi_1(g_1) \otimes \pi_2(g_2), \tag{11.38}$$

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 (11.38) is clear, and irreducibility can be established as follows. We have $\chi_\pi(g_1, g_2) = \chi_{\pi_1}(g_1)\chi_{\pi_2}(g_2)$, and hence

$$\iint_{G_1 \times G_2} |\chi_\pi(g_1, g_2)|^2 \, dg_1 \, dg_2 = 1. \tag{11.39}$$

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 $L^2(G_1 \times G_2)$.

12. Lie algebras

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

$$DL_g(h)\,X(h) = X(gh), \tag{12.1}$$

where

$$L_g : G \to G, \quad L_g x = gx, \quad DL_g : T_hG \to T_{gh}G. \tag{12.2}$$

In fact, such X is uniquely specified by

$$X(g) = DL_g(e)\,X_0. \tag{12.3}$$

We denote by $\mathfrak{g}$ the set of left-invariant vector fields on G, so $\mathfrak{g} \approx T_gG$ as a linear space.

A vector field $X \in \mathfrak{g}$ generates a flow $\mathcal{F}_X^t$ on G; cf. Appendix A. 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

$$\mathcal{F}_X^0 g = g, \quad \frac{d}{dt}\mathcal{F}_X^t g = X(\mathcal{F}_X^t g). \tag{12.4}$$

The following property helps reveal the nature of this flow.

Proposition 12.1. *Given $X \in \mathfrak{g}$, $g, h \in G$,*

$$g\,\mathcal{F}_X^t h = \mathcal{F}_X^t(gh). \tag{12.4}$$

Proof. Denote the left side of (12.4) by $x(t)$ and the right side by $y(t)$. Then $x(0) = y(0) = gh$. The result (12.4) 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 (12.1). Uniqueness for ODE then yields $x(t) \equiv y(t)$.

Let us set

$$\gamma_X(t) = \mathcal{F}_X^t e, \tag{12.5}$$

for $X \in \mathfrak{g}$. Then taking $h = e$ in (12.4) gives

$$\mathcal{F}_X^t g = g\,\gamma_X(t). \tag{12.6}$$

The following is a key group property.

Proposition 12.2. *For $X \in \mathfrak{g}$,*

$$\gamma_X(s+t) = \gamma_X(s)\gamma_X(t). \tag{12.7}$$

Proof. This follows from (12.4) 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,

$$\begin{aligned}\gamma_X(s+t) &= \mathcal{F}_X^t(\mathcal{F}_X^s e) = \mathcal{F}_X^t(\mathcal{F}_X^s e \cdot e) \\ &= \mathcal{F}_X^s e \cdot \mathcal{F}_X^t e = \gamma_X(s)\gamma_X(t),\end{aligned} \tag{12.8}$$

the third identity using (12.4).

We see that $\gamma_X(t)$ is a one-parameter subgroup of G. Note that (12.7) implies $\gamma_X(t)$ is well defined for all $t \in \mathbb{R}$, hence, by (12.6), $\mathcal{F}_X^t$ is well defined for all $t \in \mathbb{R}$, when $X \in \mathfrak{g}$. We can characterize $\gamma_X(t)$ as the unique smooth one-parameter subgroup of G satisfying $\gamma_X'(0) = X(e)$. In fact, if $\gamma(t)$ is another such one-parameter subgroup and we set $\mathcal{F}^t g = g\gamma(t)$, we see that

$$\begin{aligned}\frac{d}{dt}\mathcal{F}^t g &= \frac{d}{ds} g\gamma(t+s)\big|_{s=0} \\ &= \frac{d}{ds} g\gamma(t)\gamma(s)\big|_{s=0} \\ &= DL_{g\gamma(t)}(e)\, X(e) \\ &= X(\mathcal{F}^t g),\end{aligned} \tag{12.9}$$

so by uniqueness for ODE, $\mathcal{F}^t \equiv \mathcal{F}_X^t$.

We pause to prove (11.31), i.e.,

$$\mathrm{Sq}(x) = x \cdot x \Longrightarrow D\,\mathrm{Sq}(e)\, X = 2X, \quad \forall\, X \in T_e G. \tag{12.10}$$

To see this, since we know Sq is smooth, it suffices to note that

$$\mathrm{Sq}(\gamma_X(t)) = \gamma_X(2t) \Longrightarrow \frac{d}{dt}\,\mathrm{Sq}(\gamma_X(t))\big|_{t=0} = 2X. \tag{12.11}$$

Thus the last detail in the proof of Proposition 11.9 is taken care of.

We now define the *exponential map*:

$$\mathrm{Exp} : \mathfrak{g} \to G, \quad \mathrm{Exp}\, X = \gamma_X(1) = \mathcal{F}_X^1 e. \tag{12.12}$$

In view of the results above, we have $\gamma_{sX}(t) = \gamma_X(st)$, and hence

$$\mathrm{Exp}\, tX = \gamma_X(t). \tag{12.13}$$

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 $X \in T_eG \approx \mathfrak{g}$,

$$\mathrm{Exp}\, tX = e^{tX}, \tag{12.14}$$

the right side denoting the matrix exponential defined in §2.

Note that (12.13) implies

$$D\,\mathrm{Exp}(0) : T_eG \to T_eG \ \text{ is the identity map.} \tag{12.15}$$

Hence, by the inverse function theorem, 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

$$Xf(x) = \frac{d}{dt} f(\mathcal{F}_X^t x)\big|_{t=0}. \tag{12.16}$$

See Appendix A for more details. If $X \in \mathfrak{g}$, we can write this as

$$Xf(x) = \frac{d}{dt} f(x\gamma_X(t))\big|_{t=0}. \tag{12.17}$$

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

$$XL(g)f = L(g)Xf, \quad \forall\, g \in G,\ f \in C^\infty(G), \tag{12.18}$$

where, as usual,

$$L(g)f(x) = f(g^{-1}x). \tag{12.19}$$

In fact, the map $f(x) \mapsto f(x\gamma_X(t))$ commutes with $L(g)$, so any vector field X of the form (12.17), 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 (12.18), then so does the commutator (or Lie bracket)

$$[X, Y] = XY - YX, \tag{12.20}$$

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

$$\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \quad (X, Y) \mapsto [X, Y], \tag{12.21}$$

satisfying two identities. One is

$$[X, Y] = -[Y, X]. \tag{12.22}$$

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

$$\operatorname{ad} X(Y) = [X, Y]. \tag{12.23}$$

Then the Jacobi identity is

$$\operatorname{ad}[X, Y] = [\operatorname{ad} X, \operatorname{ad} Y], \tag{12.24}$$

where

$$[\operatorname{ad} X, \operatorname{ad} Y] = (\operatorname{ad} X)(\operatorname{ad} Y) - (\operatorname{ad} Y)(\operatorname{ad} X). \tag{12.25}$$

Plugging in the definition (12.23), one can write out the Jacobi identity as

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0. \tag{12.26}$$

It is routine to show that the commutator (12.20) satisfies (12.22) and (12.26).

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 Appendix C), 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 12.3. *H 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.

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 12.4. *Let G be a connected Lie group, $g \in G$. Then there exist $X_1, \dots, X_K \in \mathfrak{g}$ such that*

$$g = (\operatorname{Exp} X_1) \cdots (\operatorname{Exp} X_K). \tag{12.27}$$

Proof. Put a left-invariant Riemannian metric on G. By (12.15) and the inverse function theorem, there exists $\delta > 0$ such that, for $h \in G$,

$$\operatorname{dist}(h, e) < \delta \Longrightarrow h = \operatorname{Exp} Y(h), \quad \text{for some} \quad Y(h) \in \mathfrak{g}. \tag{12.28}$$

Given any $g \in G$, pick a smooth path $\sigma(t)$ from e to g, and find $g_1, \dots, g_K$ on the path such that

$$g_1 = e, \dots, g_K = g, \quad \operatorname{dist}(g_{j-1}, g_j) < \delta. \tag{12.29}$$

Then

$$g = (g_1^{-1} g_2)(g_2^{-1} g_3) \cdots (g_{K-2}^{-1} g_{K-1})(g_{K-1}^{-1} g_K), \tag{12.30}$$

and $\operatorname{dist}(g_{j-1}^{-1} g_j, e) = \operatorname{dist}(g_j, g_{j-1}) < \delta$, so each $g_{j-1}^{-1} g_j = \operatorname{Exp}(Y_j)$ for some $Y_j \in \mathfrak{g}$.

13. 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 11.10, when $\dim V < \infty$ all vectors $v \in V$ are smooth. We define the map $d\pi : \mathfrak{g} \to \operatorname{End}(V)$ as follows:

$$d\pi(X)v = \frac{d}{dt}\,\pi(\operatorname{Exp} tX)v\Big|_{t=0}, \quad X \in \mathfrak{g},\ v \in V. \tag{13.1}$$

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

$$\pi(f) = \int\limits_G f(g)\pi(g)\,dg, \tag{13.2}$$

as before, except now we use *right-invariant* Haar measure on G.

Lemma 13.1. *Given $f \in C_0^\infty(G)$, $X \in \mathfrak{g}$, we have*

$$\pi(f)d\pi(X) = -\pi(Xf). \tag{13.3}$$

Proof. Plugging in the definitions yields

$$\begin{aligned} \pi(f)d\pi(X)v &= \frac{d}{dt}\int f(g)\pi(g)\pi(\operatorname{Exp} tX)v\,dg\Big|_{t=0} \\ &= \frac{d}{dt}\int f(g)\pi(g\ \operatorname{Exp} tX)v\,dg\Big|_{t=0} \\ &= \frac{d}{dt}\int f\big(g\ \operatorname{Exp}(-tX)\big)\pi(g)v\,dg\Big|_{t=0} \\ &= -\int (Xf)(g)\pi(g)v\,dg \\ &= -\pi(Xf)v. \end{aligned} \tag{13.4}$$

Here the third identity uses the right invariance of Haar measure and the fourth identity uses (12.17).

We can deduce the following important consequence.

Proposition 13.2. *Given $X \in \mathfrak{g}$,*

$$d\pi([X,Y]) = [d\pi(X), d\pi(Y)]. \tag{13.5}$$

Proof. For any $f \in C_0^\infty(G),\ v \in V$, we have

$$\begin{aligned} \pi(f)\big(d\pi(X)d\pi(Y) - d\pi(Y)d\pi(X)\big)v \\ &= \pi(YXf)v - \pi(XYf)v \\ &= -\pi([X,Y]f)v \\ &= \pi(f)d\pi([X,Y])v. \end{aligned} \tag{13.6}$$

Letting $f = f_\nu$ be an approximate identity gives the result.

The following is an important connection between Lie algebra and Lie group representations.

Proposition 13.3. *For all* $X \in \mathfrak{g}$,

$$\pi(\operatorname{Exp} tX) = e^{t\, d\pi(X)}. \tag{13.7}$$

Proof. Let $A = d\pi(X) \in \operatorname{End}(V)$ and let $\gamma(t)$ denote the left side of (13.7). We want to show that $\gamma(t) \equiv e^{tA}$. It is clear that $\gamma : \mathbb{R} \to \operatorname{Gl}(V)$ is a smooth one-parameter group, and (13.1) gives $\gamma'(0) = d\pi(X) = A$. The group property gives

$$\gamma'(t) = \frac{d}{ds}\gamma(s+t)\big|_{s=0} = A\gamma(t) = \gamma(t)A, \tag{13.8}$$

and hence

$$\frac{d}{dt}\,\gamma(t)e^{-tA} = \gamma(t)Ae^{-tA} - \gamma(t)Ae^{-tA} = 0, \tag{13.9}$$

so $\gamma(t)e^{-tA} \equiv I$.

We next relate irreducibility of π and of $d\pi$.

Proposition 13.4. *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 $X \in \mathfrak{g}$,

$$v \in V_0 \Rightarrow d\pi(X)v = \frac{d}{dt}\pi(\operatorname{Exp} tX)v\big|_{t=0} \in V_0, \tag{13.10}$$

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 $X \in \mathfrak{g}$,

$$v \in V_0 \Rightarrow \pi(\operatorname{Exp} tX)v = e^{t\, d\pi(X)}v = \sum_{k\ge 0} \frac{t^k}{k!}\, d\pi(X)^k\, v \in V_0. \tag{13.11}$$

Now if G is connected, any $g \in G$ can be written in the form (12.27), so

$$v \in V_0 \Rightarrow \pi(g)v = \pi(\operatorname{Exp} X_1) \cdots \pi(\operatorname{Exp} X_K)v \in V_0. \tag{13.12}$$

Suppose V has a Hermitian inner product and the representation π of G on V is unitary. Then, for $X \in \mathfrak{g}$,

$$e^{-t\, d\pi(X)} = \pi(\gamma_X(t))^{-1} = \pi(\gamma_X(t))^* = (e^{t\, d\pi(X)})^*, \tag{13.13}$$

and hence

$$d\pi(X)^* = -d\pi(X). \tag{13.14}$$

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 13.5. *Let $\mathfrak{g}$ be a Lie algebra 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:*

$$\begin{gathered} A \in \operatorname{End}(V), \quad \alpha(X)A = A\alpha(X) \ \ \textit{for all} \ \ X \in \mathfrak{g} \\ \Longrightarrow A \ \ \textit{is a scalar multiple of the identity.} \end{gathered} \tag{13.15}$$

The proof is as for Lemma 5.4. 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.

14. 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

$$K_g : G \to G, \quad K_g(x) = gxg^{-1}, \tag{14.1}$$

and set

$$\mathrm{Ad}(g) = DK_g(e) : T_eG \to T_eG \approx \mathfrak{g}. \tag{14.2}$$

Note that

$$K_{gh} = K_g \circ K_h \Longrightarrow \mathrm{Ad}(gh) = \mathrm{Ad}(g)\,\mathrm{Ad}(h). \tag{14.3}$$

Proposition 14.1. *For* $g \in G,\ X \in \mathfrak{g}$*,*

$$\mathrm{Exp}(t\ \mathrm{Ad}(g)X) = g\ \mathrm{Exp}(tX)\,g^{-1}. \tag{14.4}$$

Proof. Both sides of (14.4) are one-parameter subgroups of G. Call them $\gamma(t)$ and $\sigma(t)$, respectively. It follows from (12.4) that $\gamma'(0) = \mathrm{Ad}(g)X$. Meanwhile the chain rule plus (14.2) gives $\sigma'(0) = \mathrm{Ad}(g)X$. The uniqueness result established in (12.9) then implies $\gamma(t) \equiv \sigma(t)$.

Let us take $g = \mathrm{Exp}\, sY$. By (12.6) the right side of (14.4) is then equal to

$$\mathcal{F}_Y^{-s} \circ \mathcal{F}_X^t \circ \mathcal{F}_Y^s\, e = \mathcal{F}_{X(s)}^t\, e, \quad X(s) = \mathcal{F}_{Y\#}^s X, \tag{14.5}$$

the last identity using (B.1). Consequently, comparing the left side of (14.4), and noting that $\mathcal{F}_{X(s)}^t e = \mathrm{Exp}(tX(s))$, we have

$$\mathrm{Ad}(\mathrm{Exp}\, sY)X = \mathcal{F}_{Y\#}^s X. \tag{14.6}$$

Taking the s-derivative at $s = 0$ and using (B.3)–(B.5), we have

$$\frac{d}{ds}\ \mathrm{Ad}(\mathrm{Exp}\, sY)X\big|_{s=0} = [Y, X]. \tag{14.7}$$

According to (13.1), the left side of (14.7) is the Lie algebra representation derived from Ad, i.e., $d\ \mathrm{Ad}(Y)X$. We use the notation $\mathrm{ad}(Y)$ instead of $d\ \mathrm{Ad}(Y)$:

$$\mathrm{ad}(Y)X = [Y, X], \tag{14.8}$$

a notation already brought forward in (12.23).

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 (14.1) extends to a map linear in $x \in \mathrm{M}(n, \mathbb{F})$, and we simply have

$$\mathrm{Ad}(g)X = gXg^{-1}, \quad X \in \mathfrak{g} \subset \mathrm{M}(n, \mathbb{F}). \tag{14.9}$$

If $g = \mathrm{Exp}\, sY = \gamma_Y(s)$, we have

$$\mathrm{Ad}(\mathrm{Exp}\, sY)X = \gamma_Y(s)\, X\, \gamma_Y(-s), \quad X, Y \in \mathfrak{g} \subset \mathrm{M}(n, \mathbb{F}). \tag{14.10}$$

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

$$\frac{d}{ds}\, \mathrm{Ad}(\mathrm{Exp}\, sY)X\big|_{s=0} = YX - XY, \quad X, Y \in \mathfrak{g} \subset \mathrm{M}(n, \mathbb{F}). \tag{14.11}$$

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 (13.5), applied to the "identity" representation $G \hookrightarrow \mathrm{Gl}(n, \mathbb{F})$ of G on $\mathbb{F}^n$. Compare also the presentation in Appendix E.

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

$$R_{g^{-1}}^* \omega = R_{g^{-1}}^* L_g^* \omega = K_g^* \omega = \det \mathrm{Ad}(g)\, \omega. \tag{14.12}$$

(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

$$\mathrm{Ad}(\mathrm{Exp}\, X) = e^{\mathrm{ad}\, X}, \tag{14.13}$$

which is a special case of (13.7), to formulate the unimodularity condition in purely Lie algebra terms. In view of Proposition 12.4, when G is connected, $\det \mathrm{Ad}(g) = 1$ for all $g \in G$ if and only if $\det \mathrm{Ad}(\mathrm{Exp}\, X) = 1$ for all $X \in \mathfrak{g}$. Now, for a general linear map A on a finite-dimensional vector space,

$$\det e^A = e^{\mathrm{Tr}\, A}, \tag{14.14}$$

so we have:

Proposition 14.2. *A connected Lie group G is unimodular if and only if* $\operatorname{Tr}\operatorname{ad}(X) = 0$ *for all* $X \in \mathfrak{g}$.

We give an example of a Lie group that is not unimodular, namely the 2-dimensional group $\mathrm{Aff}(1)$, 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

$$(a, b)(a', b') = (aa', b + ab'). \tag{14.15}$$

This group is isomorphic to the group of matrices

$$\left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a > 0,\ b \in \mathbb{R} \right\}. \tag{14.16}$$

Compare (1.19)–(1.22). The Lie algebra of $\mathrm{Aff}(1)$ is isomorphic to the matrix Lie subalgebra of $\mathrm{M}(2, \mathbb{R})$ spanned by

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{14.17}$$

We have

$$[X, Y] = Y, \quad \operatorname{Tr}\operatorname{ad} X = 1. \tag{14.18}$$

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 (2.19) for the derivative of the matrix exponential in terms of Ad. As shown in (2.19), if we consider

$$\mathrm{Exp} : \mathrm{M}(n, \mathbb{R}) \longrightarrow \mathrm{Gl}(n, \mathbb{R}), \quad \mathrm{Exp}\, X = e^X, \tag{14.19}$$

then

$$D\,\mathrm{Exp}(X)Y = e^X \int_0^1 e^{-\sigma X} Y e^{\sigma X}\, d\sigma. \tag{14.20}$$

Now, by (14.9) and (14.13),

$$e^{-\sigma X} Y e^{\sigma X} = \mathrm{Ad}(e^{-\sigma X})Y = e^{-\sigma\, \mathrm{ad}\, X} Y, \tag{14.21}$$

so we can rewrite (14.20) as

$$D\,\mathrm{Exp}(X)Y = e^X\, \Xi(\mathrm{ad}\, X)Y, \tag{14.22}$$

where

$$\Xi(z) = \int_0^1 e^{-\sigma z}\, d\sigma = \frac{1-e^{-z}}{z} \tag{14.23}$$

is an entire holomorphic function and $\Xi(A)$ is defined as in (2.20)–(2.21) for a linear transformation A on a finite-dimensional vector space; in this case $V = \mathfrak{g}$ and $A = \mathrm{ad}\, X$.

We now point out analogues of (14.19)–(14.23) valid for a general Lie group G, arising by defining e^{tX} for $x \in \mathfrak{g}$ as an operator on functions:

$$e^{tX} u(x) = u(\mathcal{F}_X^t x) = u(x \cdot \mathrm{Exp}\, tX). \tag{14.24}$$

Note that

$$\begin{aligned} \frac{d}{dt} e^{tX} u(x) &= Du(\mathcal{F}_X^t x)\, X(\mathcal{F}_X^t x) \\ &= Xu(\mathcal{F}_X^t x) \\ &= X e^{tX} u(x). \end{aligned} \tag{14.25}$$

Now the formulas yielding (14.20)–(14.22) have a straightforward extension to this setting:

$$\frac{d}{dt} e^{X+tY} u(x)\big|_{t=0} = e^X\, \Xi(\mathrm{ad}\, X) Y\, u(x). \tag{14.26}$$

These observations will be useful in §15.

15. The Campbell-Hausdorff formula

The Campbell-Hausdorff formula has the form

$$\mathrm{Exp}(X)\ \mathrm{Exp}(Y) = \mathrm{Exp}\big(\mathcal{C}(X,Y)\big), \tag{15.1}$$

where G is any Lie group, with Lie algebra $\mathfrak{g}$, and $\mathrm{Exp} : \mathfrak{g} \to G$ is the exponential map defined by (12.12); 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 [HS], which was also independently discovered by [Str].

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

$$e^{X(s)} = e^X e^{sY}, \tag{15.2}$$

near $s = 0$. Note that this function satisfies the ODE

$$\frac{d}{ds} e^{X(s)} = e^{X(s)} Y. \tag{15.3}$$

We can produce an ODE for $X(s)$ by using the following formula, derived in (2.19):

$$\frac{d}{ds} e^{X(s)} = e^{X(s)} \int_0^1 e^{-\tau X(s)} X'(s) e^{\tau X(s)}\, d\tau. \tag{15.4}$$

As shown in (14.22), we can rewrite this as

$$\frac{d}{ds} e^{X(s)} = e^{X(s)} \Xi\big(\mathrm{ad}\ X(s)\big) X'(s), \tag{15.5}$$

with

$$\Xi(z) = \int_0^1 e^{-\tau z}\, d\tau = \frac{1 - e^{-z}}{z}. \tag{15.6}$$

Comparing (15.3) and (15.5), we obtain

$$\Xi\big(\mathrm{ad}\ X(s)\big) X'(s) = Y, \quad X(0) = X. \tag{15.7}$$

We can obtain a more convenient ODE for $X(s)$ as follows. Note that

$$e^{\mathrm{ad}\ X(s)} = \ \mathrm{Ad}\ e^{X(s)} = \ \mathrm{Ad}\ e^X \cdot \mathrm{Ad}\ e^{sY} = e^{\mathrm{ad}\ X}\ e^{s\ \mathrm{ad}\ Y}. \tag{15.8}$$

Now let $\Psi(\zeta)$ be holomorphic near $\zeta = 1$ and satisfy

$$\Psi(e^a) = \frac{1}{\Xi(a)} = \frac{a}{1 - e^{-a}}, \tag{15.9}$$

explicitly,

$$\Psi(\zeta) = \frac{\zeta \ \log \zeta}{\zeta - 1}. \tag{15.10}$$

It follows that

$$\Psi\big(e^{\mathrm{ad}\, X} e^{s\ \mathrm{ad}\, Y}\big)\Xi\big(\mathrm{ad}\ X(s)\big) = I, \tag{15.11}$$

so we can transform (15.7) to

$$X'(s) = \Psi\big(e^{\mathrm{ad}\, X} e^{s\ \mathrm{ad}\, Y}\big)Y, \quad X(0) = X. \tag{15.12}$$

Integrating gives the Campbell-Hausdorff formula for $X(s)$ in (15.2):

$$X(s) = X + \int_0^s \Psi\big(e^{\mathrm{ad}\, X} e^{t\ \mathrm{ad}\, Y}\big)Y\ dt. \tag{15.13}$$

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

$$e^X e^Y = e^{\mathcal{C}(X,Y)}, \quad \mathcal{C}(X,Y) = X + \int_0^1 \Psi\big(e^{\mathrm{ad}\, X} e^{t\ \mathrm{ad}\, Y}\big)Y\ dt. \tag{15.14}$$

The formula (15.14) gives a power series in ad X and ad Y which is norm-summable provided

$$\|\mathrm{ad}\ X\| \le x, \quad \|\mathrm{ad}\ Y\| \le y, \tag{15.15}$$

with $e^{x+y} - 1 < 1$, i.e.,

$$x + y < \log 2. \tag{15.16}$$

We can extend the analysis above to the case where X and Y are vector fields on a manifold M, asking for a vector field $X(s)$ such that

$$\mathcal{F}^1_{X(s)} = \mathcal{F}^1_X \mathcal{F}^s_Y, \tag{15.17}$$

where $\mathcal{F}^t_X$ is the flow generated by X, evaluated at time t. If there is such a family $X(s)$, depending smoothly on s, material in Appendix D, in place of (2.19), leads to a formula parallel to (15.4), and hence to (15.7), in this context. However, we cannot always solve (15.7), because ad $X(s)$ tends not to act as a bounded operator on a Banach space of vector fields, and in fact one cannot always solve (15.17) for $X(s)$ is this case. However, if there is a *finite dimensional* Lie algebra $\mathfrak{g}$ of vector fields containing X and Y, then the analysis (15.8)–(15.16) extends. We have

$$\mathcal{F}_X^t \mathcal{F}_Y^t = \mathcal{F}_{\mathcal{C}(t,X,Y)}^t, \tag{15.18}$$

with

$$\mathcal{C}(t,X,Y) = X + \int_0^1 \Psi\big(e^{\mathrm{ad}\, tX} e^{\mathrm{ad}\, stY}\big) Y\, ds, \tag{15.19}$$

provided $\|\mathrm{ad}\, tX\| + \|\mathrm{ad}\, tY\| < \log 2$, the operator norm $\|\mathrm{ad}\, X\|$ being computed using any convenient norm on $\mathfrak{g}$. In particular, if $M = G$ is a Lie group with Lie algebra $\mathfrak{g}$, and $X, Y \in \mathfrak{g}$, this analysis applies, to yield the Campbell-Hausdorff formula for general Lie groups.

Another approach to extending the Campbell-Hausdorff formula to general Lie groups is given by studying e^{tX} as operators for $X \in \mathfrak{g}$, given by (14.24). In this approach, we take $X, Y \in \mathfrak{g}$, near 0, and look for $\mathcal{C}(X,Y) \in \mathfrak{g}$ such that

$$e^X e^Y u(x) = e^{\mathcal{C}(X,Y)} u(x). \tag{15.20}$$

The construction of $\mathcal{C}(X,Y)$ uses the same formulas as in (15.2)–(15.14). Again we have

$$\mathcal{C}(X,Y) = X + \int_0^1 \Psi(e^{\mathrm{ad}\, X} e^{t\ \mathrm{ad}\, Y}) Y\, dt. \tag{15.21}$$

Note that the left and right sides of (15.20) are equal respectively to

$$u\big(x(\mathrm{Exp}\, X)(\mathrm{Exp}\, Y)\big) \quad \text{and} \quad u(x \cdot \mathrm{Exp}\, \mathcal{C}(X,Y)), \tag{15.22}$$

so from (15.20) we again deduce

$$(\mathrm{Exp}\, X)(\mathrm{Exp}\, Y) = \mathrm{Exp}\, \mathcal{C}(X,Y). \tag{15.23}$$

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

$$\mathrm{Exp} : U \longrightarrow G \tag{15.24}$$

is a diffeomorphism of U onto a neighborhood $\mathcal{O}$ of $e \in G$. Then, for each $p \in G$, define

$$\psi_p : U \longrightarrow G, \quad \psi_p(X) = p\ \mathrm{Exp}(X). \tag{15.25}$$

Proposition 15.1. *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,

$$\psi_q^{-1} \circ \psi_p(X) = Y \Longrightarrow \operatorname{Exp}(Y) = q^{-1}p \operatorname{Exp}(X), \tag{15.26}$$

and hence, if $Z_{pq} = \operatorname{Exp}^{-1}(q^{-1}p)$, we have

$$\psi_q^{-1} \circ \psi_p(X) = \mathcal{C}(Z_{pq}, X). \tag{15.27}$$

The analyticity in X then follows from the explicit formula (15.19).

The formula (15.19) 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

$$\begin{aligned} (p \operatorname{Exp} X)(q \operatorname{Exp} Y) &= pq(q^{-1} \operatorname{Exp} X\, q) \operatorname{Exp} Y \\ &= pq \operatorname{Exp}(\operatorname{Ad}(q^{-1})X) \operatorname{Exp} Y \\ &= pq \operatorname{Exp}(\mathcal{C}(\operatorname{Ad}(q^{-1})X, Y)), \end{aligned} \tag{15.28}$$

the second identity by (14.4). 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

$$\begin{aligned} \Psi(1+z) = (1+z)\frac{\log(1+z)}{z} &= (1+z)\Bigl(1 - \frac{z}{2} + \frac{z^2}{3} - \cdots\Bigr) \\ &= 1 + \sum_{k \ge 1} \frac{(-1)^{k-1}}{k(k+1)}\, z^k \\ &= 1 + \frac{z}{2} - \frac{z^2}{6} + \cdots . \end{aligned} \tag{15.29}$$

If we set $\operatorname{ad} X = \xi$ and $\operatorname{ad} Y = \eta$, we have

$$\begin{aligned} \Psi(e^{\xi}e^{t\eta}) &= \Psi\bigl((I + \xi + \tfrac{1}{2}\xi^2 + \cdots)(I + t\eta + \tfrac{1}{2}t^2\eta^2 + \cdots)\bigr) \\ &= \Psi(I + \xi + t\eta + \tfrac{1}{2}\xi^2 + t\xi\eta + \tfrac{1}{2}t^2\eta^2 + \cdots) \\ &= I + \tfrac{1}{2}\xi + \tfrac{1}{2}t\eta + \tfrac{1}{12}\xi^2 + \tfrac{1}{3}t\xi\eta - \tfrac{1}{6}t\eta\xi + \tfrac{1}{12}t^2\eta^2 + \cdots). \end{aligned} \tag{15.30}$$

Noting that $\eta(Y) = [Y,Y] = 0$, we see that

$$\int_0^1 \Psi(e^{\operatorname{ad} X}e^{t \operatorname{ad} Y})Y\, dt = Y + \frac{1}{2}\xi(Y) + \frac{1}{12}\xi^2(Y) - \frac{1}{12}\eta\xi(Y) + \cdots , \tag{15.31}$$

and hence

$$\mathcal{C}(X,Y)=X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+\cdots. \tag{15.32}$$

We complement (15.32) with a complete power series expansion, as follows. We have

$$\Psi(e^{\xi}e^{t\eta})=I+\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k(k+1)}\big(e^{\xi}e^{t\eta}-I\big)^k, \tag{15.33}$$

and

$$(e^{\xi}e^{t\eta}-I)^k=\Big(\sum_{(\ell,m)\in\mathcal{S}_1}\frac{\xi^\ell}{\ell!}\,\frac{t^m\eta^m}{m!}\Big)^k, \tag{15.34}$$

where we set $\mathcal{S}_1=\{(\ell,m):\ell,m\ge 0,\ \ell+m>0\}$. More generally, set

$$\mathcal{S}_k=\{(\ell_1,\dots,\ell_k,m_1,\dots,m_k):\ell_j\ge 0,m_j\ge 0,\ell_j+m_j>0\}. \tag{15.35}$$

Then we can expand the right side of (15.34), to obtain

$$(e^{\xi}e^{t\eta}-I)^k=\sum_{\mathcal{S}_k}t^{m_1+\cdots+m_k}\frac{\xi^{\ell_1}}{\ell_1!}\frac{\eta^{m_1}}{m_1!}\cdots\frac{\xi^{\ell_k}}{\ell_k!}\frac{\eta^{m_k}}{m_k!}. \tag{15.36}$$

Plugging this into (15.33) and then into (15.21), we obtain

$$\begin{aligned}\mathcal{C}(X,Y)=X+Y+\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k(k+1)}\sum_{\mathcal{S}_k}\frac{1}{m_1+\cdots+m_k+1}\frac{(\operatorname{ad}X)^{\ell_1}}{\ell_1!}\frac{(\operatorname{ad}Y)^{m_1}}{m_1!}\cdots\\ \times\frac{(\operatorname{ad}X)^{\ell_k}}{\ell_k!}\frac{(\operatorname{ad}Y)^{m_k}}{m_k!}\,Y.\end{aligned} \tag{15.37}$$

16. 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 13.2. Let G and H be Lie groups, and suppose

$$\rho : G \longrightarrow H \tag{16.1}$$

is a smooth group homomorphism. Denote the Lie algebras by $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and set

$$\sigma = D\rho(e) : \mathfrak{g} \longrightarrow \mathfrak{h}. \tag{16.2}$$

Thus, if $X \in \mathfrak{g}$, generating the one-parameter group $\gamma_X(t)$, we have

$$\rho \circ \gamma_X(t) = \gamma_{\sigma(X)}(t). \tag{16.3}$$

Proposition 16.1. *The linear map σ in (16.2) is a Lie algebra homomorphism, i.e.,*

$$X, Y \in \mathfrak{g} \Longrightarrow \sigma([X,Y]) = [\sigma(X), \sigma(Y)]. \tag{16.4}$$

Proof. We make use of results on the adjoint representation established in §14. With K_g given by (14.1), we have

$$\rho \circ K_g(x) = \rho(gxg^{-1}) = K_{\rho(g)}(\rho(x)). \tag{16.5}$$

Regard each side of (16.5) 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

$$D\rho(e) \circ DK_g(e) = DK_{\rho(g)}(e) \circ D\rho(e), \tag{16.6}$$

or

$$\sigma \circ \operatorname{Ad}(g) = \operatorname{Ad}(\rho(g)) \circ \sigma, \tag{16.7}$$

as maps from $\mathfrak{g}$ to $\mathfrak{h}$. Taking $g = \gamma_X(t)$ and using (16.3), we have

$$\sigma \circ \operatorname{Ad}(\gamma_X(t)) = \operatorname{Ad}(\gamma_{\sigma(X)}(t)) \circ \sigma. \tag{16.8}$$

Taking d/dt at $t = 0$ and using (14.7) then gives

$$\sigma \circ \operatorname{ad} X = \operatorname{ad} \sigma(X) \circ \sigma, \tag{16.9}$$

which is equivalent to (16.4).

We aim for a converse to Proposition 16.1. 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 $\operatorname{Exp} : U \to G$ is a diffeomorphism onto a neighborhood $\mathcal{O}$ of $e \in G$, and assume the Cambell-Hausdorff formula (15.19) holds for $t \in [0,1],\ X, Y \in U$. Let us define

$$\rho : \mathcal{O} \longrightarrow H \tag{16.10}$$

by

$$\rho(g) = \operatorname{Exp}(\sigma \circ \operatorname{Exp}^{-1}(g)). \tag{16.11}$$

Lemma 16.2. *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

$$g_1, g_2 \in \mathcal{O}_1 \Longrightarrow \rho(g_1 g_2) = \rho(g_1)\rho(g_2). \tag{16.12}$$

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 (15.19) implies that, if σ is a Lie algebra homomorphism, then, for X_j sufficiently close to 0,

$$\sigma \mathcal{C}(X_1, X_2) = \mathcal{C}(\sigma(X_1), \sigma(X_2)). \tag{16.13}$$

Hence

$$\begin{aligned} \rho(g_1 g_2) &= \operatorname{Exp} \mathcal{C}(\sigma(X_1), \sigma(X_2)) \\ &= \operatorname{Exp} \sigma(X_1) \ \operatorname{Exp} \sigma(X_2) \\ &= \rho(g_1)\rho(g_2), \end{aligned} \tag{16.14}$$

as asserted.

A map $\rho : \mathcal{O} \to H$ as in Lemma 16.2 is called a local homomorphism of G to H. We have the following result.

Proposition 16.3. *Let $\mathcal{O}_1 \subset \mathcal{O}$ and $\rho : \mathcal{O} \to H$ be as in Lemma 16.2. 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 $\mathcal{O}_1 = B_{\delta/2}(e)$. 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 $|t_{j+1} - t_j| < \delta/2$. Thus

$$g_0 = e, \quad g_N = g, \quad g_{j+1} = x_j g_j, \quad x_j \in \mathcal{O}_1, \quad g = x_{N-1} \cdots x_2 x_1. \tag{16.15}$$

See Figure 16.1. We set

$$\rho_\gamma(g) = \rho(x_{N-1}) \cdots \rho(x_2)\rho(x_1). \tag{16.16}$$

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 (16.16). So say we add one point, $t_{j+1/2} \in (t_j, t_{j+1})$.

Then the factor $\rho(x_j)$ in (16.16) gets replaced by $\rho(z_j)\rho(y_j)$, 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 (16.12), $\rho(z_j)\rho(y_j) = \rho(x_j)$, and indeed (16.16) 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 (16.15) and take $g'_j = \sigma(t_j)$, with

$$g'_0 = e, \quad g'_N = g, \quad g'_{j+1} = x'_j g'_j, \quad g = x'_{N-1} \cdots x'_2 x'_1. \tag{16.17}$$

See Figure 16.2. Here $\operatorname{dist}(x_j, e) < \delta/8$ and $\operatorname{dist}(x'_j, e) < \delta/8$. We also have

$$g'_j = z_j g_j, \quad \operatorname{dist}(z_j, e) < \frac{\delta}{8}. \tag{16.18}$$

In order to show that $\rho_\gamma(g) = \rho_\sigma(g)$, it will suffice to show that, for each k,

$$\rho_\sigma(g'_k) = \rho(z_k)\rho_\gamma(g_k), \tag{16.19}$$

and we do this by induction.

Clearly (16.19) holds for $k = 0$. Suppose it holds for $k = j - 1$. That is, we assume

$$\rho_\sigma(g'_{j-1}) = \rho(z_{j-1})\rho_\gamma(g_{j-1}), \tag{16.20}$$

and try to show

$$\rho_\sigma(g'_j) = \rho(z_j)\rho_\gamma(g_j), \tag{16.21}$$

assuming $j \le N$. In fact,

$$g'_j = x'_j z_{j-1} x_j^{-1} g_j, \quad \text{and} \quad g'_j = z_j g_j, \quad \text{so} \quad z_j = x'_j z_{j-1} x_j^{-1}, \tag{16.22}$$

and x'_j, x_j, z_{j-1} are all sufficiently close to the identity that (16.12) gives

$$\rho(z_j) = \rho(x'_j)\rho(z_{j-1})\rho(x_j^{-1}), \tag{16.23}$$

which does lead to (16.21) from (16.20).

At this point we can define

$$\rho : G \longrightarrow H, \quad \rho(g) = \rho_\gamma(g), \tag{16.24}$$

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 (16.24), take any $h \in G$ and write

$$h = y_{M-1} \cdots y_2 y_1, \quad y_j \in \mathcal{O}_1, \tag{16.25}$$

parallel to (16.15), with partial products $h_k = y_{k-1} \cdots y_2 y_1$ lying along a smooth curve from e to h. Then

$$gh = x_{N-1} \cdots x_2 x_1 y_{M-1} \cdots y_2 y_1 \tag{16.26}$$

has the form developed above, so the construction of ρ in (16.24) yields

$$\rho(gh) = \rho(x_{N-1}) \cdots \rho(x_2)\rho(x_1)\rho(y_{M-1}) \cdots \rho(y_2)\rho(y_1). \tag{16.27}$$

However the right side of (16.27) is equal to $\rho(g)\rho(h)$, so indeed

$$\rho(gh) = \rho(g)\rho(h), \quad \forall\, g, h \in G. \tag{16.28}$$

Finally, the analyticity of ρ on $\mathcal{O}_1$ follows from (16.11) 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.

Thus we have a converse to Proposition 16.1:

Corollary 16.4. *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 $d\rho = \sigma$.*

17. 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

$$\mathfrak{U}(\mathfrak{g}) = \bigotimes \mathfrak{g}_{\mathbb{C}}/J, \tag{17.1}$$

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

$$\{XY - YX - [X, Y] : X, Y \in \mathfrak{g}\}. \tag{17.2}$$

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, but we will not need the proof of this. See Appendix F 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

$$P = \sum_{\mu \le m} c_{i_1\cdots i_\mu} X_{i_1} \cdots X_{i_\mu}, \quad X_j \in \mathfrak{g}, \tag{17.3}$$

with $c_{i_1\cdots i_\mu} \in \mathbb{C}$, we have

$$d\pi(P) = \sum_{\mu \le m} c_{i_1\cdots i_\mu} d\pi(X_{i_1}) \cdots d\pi(X_{i_\mu}). \tag{17.4}$$

The following result is an immediate consequence of Proposition 13.5. As we will see it will be quite useful.

Proposition 17.1. *Suppose G is connected. Let $P \in \mathfrak{U}(\mathfrak{g})$ and assume*

$$PX = XP, \quad \forall\ X \in \mathfrak{g}. \tag{17.5}$$

If π is an irreducible unitary representation of G on V, then $d\pi(P)$ is a scalar multiple of the identity:

$$d\pi(P) = \lambda I.$$

18. 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.,

$$\text{SU}(2) = \left\{ \begin{pmatrix} z_1 & z_2 \\ -\overline{z}_2 & \overline{z}_1 \end{pmatrix} : |z_1|^2 + |z_2|^2 = 1,\ z_j \in \mathbb{C} \right\}. \tag{18.1}$$

As a set, SU(2) is naturally identified with the unit sphere S^3 in $\mathbb{C}^2$. Its Lie algebra su(2) consists of 2×2 complex skew adjoint matrices of trace zero. A basis of su(2) is formed by

$$X_1 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad X_2 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad X_3 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{18.2}$$

Note the commutation relations

$$[X_1, X_2] = X_3, \quad [X_2, X_3] = X_1, \quad [X_3, X_1] = X_2. \tag{18.3}$$

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 (18.3). Thus SU(2) and SO(3) have isomorphic Lie algebras. There is an explicit homomorphism

$$p : \text{SU}(2) \longrightarrow \text{SO}(3), \tag{18.4}$$

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 (18.2) 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 $p(g) = gXg^{-1}$ preserves this inner product and gives (18.4). Note that $\operatorname{Ker} p = \{I, -I\}$.

If we regard X_j as left-invariant vector fields on SU(2), set

$$\Delta = X_1^2 + X_2^2 + X_3^2, \tag{18.5}$$

a second-order, left-invariant differential operator. It follows easily from (18.3) that X_j and Δ commute:

$$\Delta X_j = X_j \Delta, \quad 1 \le j \le 3. \tag{18.6}$$

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 su(2), and an algebraic representation of the universal enveloping algebra. By (18.6), $d\pi(\Delta)$ commutes with $d\pi(X_j),\ j = 1, \dots, 3$. Thus, if π is irreducible, Proposition 17.1 implies

$$d\pi(\Delta) = -\lambda^2 I, \tag{18.7}$$

for some $\lambda \in \mathbb{R}$. (Since $d\pi(\Delta)$ is a sum of squares of skew-adjoint operators, it must be negative.) Let

$$L_j = d\pi(X_j). \tag{18.8}$$

Now we will diagonalize L_1 on V. Set

$$V_\mu = \{v \in V : L_1 v = i\mu v\}, \quad V = \bigoplus_{i\mu \in \text{ spec } L_1} V_\mu. \tag{18.9}$$

The structure of π is defined by how L_2 and L_3 behave on V_μ. It is convenient to set

$$L_\pm = L_2 \mp i L_3, \tag{18.10}$$

i.e., $L_\pm = d\pi(X_\pm)$ where

$$X_+ = X_2 - iX_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_- = X_2 + iX_3 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}.$$

We have the following key identities, as a direct consequence of (18.3):

$$[X_1, X_\pm] = \pm i X_\pm, \quad \text{hence} \quad [L_1, L_\pm] = \pm i L_\pm. \tag{18.11}$$

Using this, we can establish the following:

Lemma 18.1. *We have*

$$L_\pm : V_\mu \longrightarrow V_{\mu \pm 1}. \tag{18.12}$$

In particular, if $i\mu \in$ spec L_1, then either $L_+ = 0$ on V_μ or $\mu + 1 \in$ spec L_1, and also either $L_- = 0$ on V_μ or $\mu - 1 \in$ spec L_1.

Proof. Let $v \in V_\mu$. By (18.11) we have

$$L_1 L_\pm v = L_\pm L_1 v \pm i L_\pm v = i(\mu \pm 1) L_\pm v,$$

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

$$\text{spec } \frac{1}{i} L_1 = \{\mu_0, \mu_0 + 1, \dots, \mu_0 + k = \mu_1\}, \tag{18.13}$$

with

$$L_+ : V_{\mu_0+j} \to V_{\mu_0+j+1} \quad \text{isomorphism, for } \ 0 \le j \le k-1, \tag{18.14}$$

and

$$L_- : V_{\mu_1-j} \to V_{\mu_1-j-1} \quad \text{isomorphism, for } \ 0 \le j \le k-1. \tag{18.15}$$

In fact, we can compute

$$L_-L_+ = L_2^2 + L_3^2 + i[L_3, L_2] = -\lambda^2 - L_1^2 - iL_1 \tag{18.16}$$

on V, and

$$L_+L_- = -\lambda^2 - L_1^2 + iL_1 \tag{18.17}$$

on V, so

$$\begin{aligned} L_-L_+ &= \mu(\mu+1) - \lambda^2 \ \text{ on } \ V_\mu, \\ L_+L_- &= \mu(\mu-1) - \lambda^2 \ \text{ on } \ V_\mu. \end{aligned} \tag{18.18}$$

Note that, since L_2 and L_3 are skew-adjoint, $L_+ = -L_-^*$, so

$$L_+L_- = -L_-^* L_-, \quad L_-L_+ = -L_+^* L_+.$$

Thus

$$\text{Ker } L_+ = \text{ Ker } L_-L_+, \quad \text{Ker } L_- = \text{ Ker } L_+L_-.$$

These observations establish (18.13)–(18.15).

Considering that $d\pi$ acts on 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

$$\dim V_\mu = 1, \quad \mu_0 \le \mu \le \mu_1. \tag{18.19}$$

From (18.18) we see that $\mu_1(\mu_1+1) = \lambda^2 = \mu_0(\mu_0-1)$. Hence,

$$\mu_1 - \mu_0 = k \Longrightarrow \mu_0 = -k/2, \ \mu_1 = k/2, \tag{18.20}$$

and we have

$$\dim V = k+1, \quad \lambda^2 = \frac{1}{4}k(k+2) = \frac{1}{4}(\dim V^2 - 1). \tag{18.21}$$

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 18.2. *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

$$\mathcal{P}_k = \{p(z) : p \text{ homogeneous polynomial of degree } k \text{ on } \mathbb{C}^2\}, \tag{18.22}$$

with SU(2) acting on $\mathcal{P}_k$ by

$$D_{k/2}(g)f(z) = f(g^{-1}z), \quad g \in SU(2),\ z \in \mathbb{C}^2. \tag{18.23}$$

Note that, for $X \in \text{su}(2)$,

$$dD_{k/2}(X)f(z) = \frac{d}{dt} f\big(e^{-tX}z\big)\big|_{t=0} = -(\partial_1 f, \partial_2 f) \cdot X \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \tag{18.24}$$

where $\partial_j f = \partial f / \partial z_j$. A calculation gives

$$\begin{aligned} L_1 f(z) &= -\frac{i}{2}(z_1\partial_1 f - z_2\partial_2 f), \\ L_2 f(z) &= -\frac{1}{2}(z_2\partial_1 f - z_1\partial_2 f), \\ L_3 f(z) &= -\frac{i}{2}(z_2\partial_1 f + z_1\partial_2 f). \end{aligned} \tag{18.25}$$

In particular, for

$$\varphi_{kj}(z) = z_1^{k-j} z_2^j \in \mathcal{P}_k, \quad 0 \le j \le k, \tag{18.26}$$

we have

$$L_1 \varphi_{kj} = i\Big(-\frac{k}{2} + j\Big)\varphi_{kj}, \tag{18.27}$$

so

$$V = \mathcal{P}_k \Longrightarrow \text{ span } \varphi_{kj} = V_{-k/2+j}, \quad 0 \le j \le k. \tag{18.28}$$

Note that

$$L_+ f(z) = -z_2 \partial_1 f(z), \quad L_- f(z) = z_1 \partial_2 f(z), \tag{18.29}$$

so

$$L_+ \varphi_{kj} = -(k-j)\varphi_{k,j+1}, \quad L_- \varphi_{kj} = j\varphi_{k,j-1}. \tag{18.30}$$

We see that the structure of the representation $D_{k/2}$ of SU(2) on $\mathcal{P}_k$ is as described in (18.12)–(18.21). 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.

We can deduce the classification of irreducible unitary representations of SO(3) from the result above as follows. We have the covering homomorphism (18.4), 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

$$\widetilde{D}_j\big(p(g)\big) = D_j(g), \quad g \in \mathrm{SU}(2). \tag{18.31}$$

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

$$dD_j(\Delta) = -j(j+1). \tag{18.32}$$

Also we can classify the irreducible representations of U(2), using the results on SU(2). To do this, use the exact sequence

$$1 \to K \to S^1 \times \mathrm{SU}(2) \to \mathrm{U}(2) \to 1, \tag{18.33}$$

where "1" denotes the trivial multiplicative group, and

$$K = \{(\omega, g) \in S^1 \times \mathrm{SU}(2) : g = \omega^{-1} I, \omega^2 = 1\}. \tag{18.34}$$

The irreducible representations of $S^1 \times \mathrm{SU}(2)$ are given by

$$\pi_{mk}(\omega, g) = \omega^m D_{k/2}(g) \ \text{ on } \ \mathcal{P}_k, \tag{18.35}$$

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.

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

$$(g_1, g_2) \cdot x = g_1 x g_2^{-1}, \quad g_j \in \mathrm{SU}(2), \quad x \in \mathrm{SU}(2) \approx S^3. \tag{18.36}$$

Thus we have a map

$$\tau : \mathrm{SU}(2) \times \mathrm{SU}(2) \longrightarrow \mathrm{SO}(4). \tag{18.37}$$

This is a group homomorphism. Note that $(g_1, g_2) \in \operatorname{Ker} \tau$ implies $g_1 = g_2 = \pm I$. Furthermore, a dimension count shows τ must be surjective, so

$$\mathrm{SO}(4) \approx \mathrm{SU}(2) \times \mathrm{SU}(2)/\{\pm(I, I)\}. \tag{18.38}$$

As shown in Proposition 11.11, 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

$$\{\pi(g) = \pi_1(g_1) \otimes \pi_2(g_2) : \pi_j \in \widehat{G}_j\}, \tag{18.39}$$

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

$$\gamma_{k\ell}(g) = D_{k/2}(g_1) \otimes D_{\ell/2}(g_2), \quad k, \ell \in \{0, 1, 2, \dots\}, \tag{18.40}$$

acting on $\mathcal{P}_k \otimes \mathcal{P}_\ell \approx \mathbb{C}^{k+1} \otimes \mathbb{C}^{\ell+1}$. By (18.38), the irreducible unitary representations of SO(4) are given by all $\gamma_{k\ell}$ such that $k + \ell$ is even, since, for $p_0 = (-I, -I) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$, $\gamma_{k\ell}(p_0) = (-1)^{k+\ell} I$.

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 (18.40) 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

$$\begin{aligned} \mathcal{P}_{k\ell} = \{ f(z, w) &: \text{polynomial on } \mathbb{C}^2 \times \mathbb{C}^2, \\ &\text{homogeneous of degree } k \text{ in } z, \ \ell \text{ in } w\}, \end{aligned} \tag{18.41}$$

as

$$\pi_{k\ell}(g) f(z, w) = f(g^{-1} z, g^{-1} w). \tag{18.42}$$

Parallel to (18.25) and (18.29), we have, on $\mathcal{P}_{k\ell}$,

$$\begin{aligned} L_1 f &= -\frac{i}{2}(z_1 \partial_{z_1} f - z_2 \partial_{z_2} f + w_1 \partial_{w_1} f - w_2 \partial_{w_2} f), \\ L_+ f &= -z_2 \partial_{z_1} f - w_2 \partial_{w_1} f, \\ L_- f &= z_1 \partial_{z_2} f + w_1 \partial_{w_2} f. \end{aligned} \tag{18.43}$$

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

$$f(z,w) = g(z_2, w_2, w_2 z_1 - z_2 w_1), \tag{18.44}$$

and the kernel of L_+ in $\mathcal{P}_{k\ell}$ is the linear span of

$$\psi_{k\ell\mu}(z,w) = z_2^{k-\mu} w_2^{\ell-\mu} (w_2 z_1 - z_2 w_1)^\mu, \quad 0 \le \mu \le \ell. \tag{18.45}$$

A calculation gives

$$L_1 \psi_{k\ell\mu} = \frac{i}{2}(k + \ell - 2\mu)\psi_{k\ell\mu}. \tag{18.46}$$

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

$$D_{k/2} \otimes D_{\ell/2} \approx \bigoplus_{\mu=0}^{\ell} D_{(k+\ell-2\mu)/2} = D_{(k-\ell)/2} \oplus D_{(k-\ell)/2+1} \oplus \cdots \oplus D_{(k+\ell)/2}. \tag{18.47}$$

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

$$K = \operatorname{Ker} L_+ \subset V. \tag{18.48}$$

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

$$L_1 : K \longrightarrow K, \tag{18.49}$$

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 $\mathcal{P}_{k\ell}$.

19. 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

$$\mathrm{u}(n) = \{X \in \mathrm{M}(n, \mathbb{C}) : X^* = -X\}. \tag{19.1}$$

The complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$ of $\mathfrak{g} = \mathrm{u}(n)$ is just M($n, \mathbb{C}$), which is the Lie algebra of Gl($n, \mathbb{C}$), which in turn can be regarded as the complexification of U(n). We can write

$$\mathrm{M}(n, \mathbb{C}) = \mathbb{C}\mathfrak{h} \oplus \mathfrak{n}_+ \oplus \mathfrak{n}_-. \tag{19.2}$$

Here

$$\mathfrak{h} = \{\mathrm{diag}(ia_1, \dots, ia_n) : a_j \in \mathbb{R}\} \tag{19.3}$$

is the Lie algebra of

$$\mathbb{T} = \{\mathrm{diag}(e^{ia_1}, \dots, e^{ia_n}) : a_j \in \mathbb{R}\} \subset \mathrm{U}(n). \tag{19.4}$$

In addition $\mathfrak{n}_+$ consists of strictly upper triangular matrices and $\mathfrak{n}_-$ of strictly lower triangular matrices, in M($n, \mathbb{C}$). It is clear that each of the three factors on the right side of (19.2) is a Lie algebra. The Lie algebra $\mathbb{C}\mathfrak{h}$ generates

$$D = \{\mathrm{diag}(c_1, \dots, c_n) : c_j \in \mathbb{C} \setminus 0\} \subset \mathrm{Gl}(n, \mathbb{C}), \tag{19.5}$$

while $\mathfrak{n}_+$ generates N_+, the group of upper triangular matrices in Gl($n, \mathbb{C}$) with ones on the diagonal, and $\mathfrak{n}_-$ generates N_-, the group of lower triangular matrices in Gl($n, \mathbb{C}$) with ones on the diagonal. There is the Gauss decomposition:

$$N_- D N_+ = G_{\mathrm{reg}} \ \text{ is dense in } \ \mathrm{Gl}(n, \mathbb{C}), \tag{19.6}$$

or (in a weaker form) G_{reg} contains a neighborhood of the identity. The latter result follows fairly easily from the spanning property (19.2).

Convenient bases for the factors in (19.2) 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

$$e_{jk} u_\ell = \delta_{k\ell}\, u_j. \tag{19.7}$$

Then $\{e_{jk} : j < k\}$ spans $\mathfrak{n}_+$, $\{e_{jk} : j > k\}$ spans $\mathfrak{n}_-$, and, with

$$e_j = i e_{jj}, \tag{19.8}$$

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,

$$V = \bigoplus_{\lambda \in \mathfrak{h}'} V_\lambda, \tag{19.9}$$

where, for $\lambda \in \mathfrak{h}'$,

$$V_\lambda = \{v \in V : d\pi(X)v = i\lambda(X)v, \ \forall\ X \in \mathfrak{h}\}. \tag{19.10}$$

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 (19.9) are mutually orthogonal.

Let us apply this notion to the adjoint representation of U(n) on $\mathrm{u}(n)_{\mathbb{C}} = \mathrm{M}(n, \mathbb{C})$. It is convenient to use the basis e_{jk} defined in (19.7). A computation gives

$$[e_{ij}, e_{k\ell}] = \delta_{jk} e_{i\ell} - \delta_{i\ell} e_{kj}. \tag{19.11}$$

In particular,

$$X = \sum_j x_j e_j \in \mathfrak{h} \Longrightarrow [X, e_{jk}] = i(x_j - x_k) e_{jk}. \tag{19.12}$$

In other words, if we define

$$\omega_{jk} \in \mathfrak{h}', \quad \omega_{jk}(X) = x_j - x_k, \tag{19.13}$$

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 (19.12) and the commutator relation $[X_1, X_\pm] = \pm i X_\pm$, from (18.11).

Let us return to a general unitary representation π of $U(n)$ on V. The following can be compared with Lemma 18.1.

Proposition 19.1. *Set $E_{jk} = d\pi(e_{jk})$. Then*

$$E_{jk} : V_\lambda \longrightarrow V_{\lambda + \omega_{jk}}. \tag{19.14}$$

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 (19.12), which can be rewritten as

$$[X, e_{jk}] = i\omega_{jk}(X) e_{jk}, \quad X \in \mathfrak{h}, \tag{19.15}$$

leads to the identity

$$(19.16)\qquad d\pi(X)E_{jk} = E_{jk}\big(d\pi(X) + i\omega_{jk}(X)I\big), \quad X \in \mathfrak{h},$$

which implies (19.14).

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

$$(19.17)\qquad \begin{aligned} \lambda + \omega_{jk} &> \lambda \ \text{ if } \ j < k, \\ \lambda + \omega_{jk} &< \lambda \ \text{ if } \ j > k. \end{aligned}$$

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 (19.9) 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 19.2. *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 $\mathcal{A}(\pi)$.

Let us note the following. Set

$$(19.18)\qquad \mathcal{H}(\pi) = \bigcap_{j<k} \operatorname{Ker} E_{jk}.$$

From (19.16) it follows that

$$(19.19)\qquad X \in \mathfrak{h} \Longrightarrow d\pi(X) : \mathcal{H}(\pi) \to \mathcal{H}(\pi),$$

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

$$(19.20)\qquad \langle v, \overline{\pi}(g)w\rangle = \langle \pi(g^{-1})v, w\rangle, \quad v \in V,\ w \in V',$$

as in (7.12). 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

$$\psi(X) = \langle d\pi(X)\xi_0, \eta_0\rangle = -\langle \xi_0, d\overline{\pi}(X)\eta_0\rangle, \quad X \in \mathrm{M}(n, \mathbb{C}). \tag{19.21}$$

Note that

$$\begin{aligned} X_+ \in \mathfrak{n}_+ &\Longrightarrow \psi(X_+) = 0, \\ X_- \in \mathfrak{n}_- &\Longrightarrow \psi(X_-) = 0, \\ H \in \mathfrak{h} &\Longrightarrow \psi(H) = i\lambda(H)\langle \xi_0, \eta_0\rangle = i\mu(H)\langle \xi_0, \eta_0\rangle. \end{aligned} \tag{19.22}$$

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 (19.21). 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

$$\alpha(g) = \langle \pi(g)\xi_0, \eta_0\rangle = \langle \xi_0, \overline{\pi}(g^{-1})\eta_0\rangle. \tag{19.23}$$

As we will show in §22, 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 Appendix G.) Hence $\alpha(g)$ is well defined for $g \in \mathrm{Gl}(n, \mathbb{C})$ and is holomorphic in g. Parallel to (19.22), we have, for all $g \in \mathrm{Gl}(n, \mathbb{C})$,

$$\begin{aligned} \alpha(g\zeta_+) &= \alpha(g), \quad \zeta_+ \in N_+, \\ \alpha(\zeta_- g) &= \alpha(g), \quad \zeta_- \in N_-, \end{aligned} \tag{19.24}$$

since $\pi(\zeta_+)\xi_0 = \xi_0$ and $\overline{\pi}(\zeta_-^{-1})\eta_0 = \eta_0$. Also

$$\begin{aligned} \alpha(g\delta) &= e^{i\lambda(H)}\alpha(g), \quad \delta = e^H \in \mathbb{T}, \\ \alpha(\delta g) &= e^{i\mu(H)}\alpha(g), \end{aligned} \tag{19.25}$$

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,

$$\begin{aligned} \alpha(\delta g) &= e^{i(\lambda(H_1) + i\lambda(H_2))}\alpha(g), \quad \delta = e^{H_1 + iH_2} \in D, \\ \alpha(g\delta) &= e^{i(\mu(H_1) + i\mu(H_2))}\alpha(g). \end{aligned} \tag{19.26}$$

We have from (19.24)–(19.26) that

$$\zeta_\pm \in N_\pm,\ \delta = e^{H_1 + iH_2} \in D \Longrightarrow \alpha(\zeta_-\delta\zeta_+) = \alpha(\delta) = e^{i(\lambda(H_1) + i\lambda(H_2))}\alpha(e). \tag{19.27}$$

We are now prepared to prove:

Lemma 19.3. *Given that π is irreducible,*

$$\langle \xi_0, \eta_0 \rangle \neq 0. \tag{19.28}$$

Hence $\lambda = \mu$.

Proof. We have $\langle \xi_0, \eta_0 \rangle = \alpha(e)$. By (19.27), (19.6) and holomorphy, if $\alpha(e) = 0$ then $\alpha(g) \equiv 0$. Consider

$$V_0 = \{\xi \in V : \langle \pi(g)\xi, \eta_0 \rangle = 0, \ \forall\, g \in \mathrm{Gl}(n, \mathbb{C})\}. \tag{19.29}$$

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 (19.28) to hold.

Having $\lambda = \mu$, we can rewrite (19.26) as

$$\alpha(g\delta) = \alpha(\delta g) = e^{i(\lambda(H_1) + i\lambda(H_2))}\alpha(g), \quad \delta = e^{H_1 + iH_2} \in D. \tag{19.30}$$

We next prove:

Proposition 19.4. *If π is an irreducible representation of* $\mathrm{U}(n)$ *on V, then $\mathcal{H}(\pi)$ 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 (19.21)–(19.27) give $\langle \pi(g)\xi_1, \eta_0 \rangle \equiv \alpha(g)$, so

$$\langle \pi(g)(\xi_1 - \xi_0), \eta_0 \rangle = 0, \quad \forall\, g, \tag{19.31}$$

or

$$\langle \xi_1 - \xi_0, \overline{\pi}(g)\eta_0 \rangle = 0, \forall\, g. \tag{19.32}$$

Since $\overline{\pi}$ is irreducible, this implies $\xi_1 = \xi_0$.

We next show that inequivalent irreducible representations of $\mathrm{U}(n)$ have distinct highest weights.

Proposition 19.5. *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 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

$$\beta(g) = \langle \pi'(g)\xi_0', \eta_0' \rangle. \tag{19.33}$$

We have $\beta(e) = \alpha(e) \neq 0$, and results parallel to (19.24)–(19.27) for β imply

$$\beta(\zeta_-\delta\zeta_+) = \alpha(\zeta_-\delta\zeta_+), \quad \forall\, \zeta_- \in N_-,\ \delta \in D,\ \zeta_+ \in N_+. \tag{19.34}$$

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 §21.

20. Representations of U(n), II: some basic examples

Here we now consider some basic examples of representations of U(n). First, define representations S^ℓ and $\overline{S}^\ell$ of U(n) on

$$\mathcal{P}_\ell = \text{ space of polynomials on } \mathbb{C}^n \text{ homogeneous of degree } \ell, \tag{20.1}$$

by

$$S^\ell(g)f(z) = f(g^t z), \quad \overline{S}^\ell(g)f(z) = f(g^{-1}z). \tag{20.2}$$

Note that (20.2) extends to $g \in \mathrm{Gl}(n, \mathbb{C})$, and we have

$$d\overline{S}^\ell(X)f(z) = \frac{d}{dt} f(e^{-tX}z)\big|_{t=0} = \frac{d}{dt} f(z - tXz)\big|_{t=0}, \quad X \in \mathrm{M}(n, \mathbb{C}). \tag{20.3}$$

Hence

$$\begin{aligned} d\overline{S}^\ell(e_{jk})p(z) &= \frac{d}{dt} p(z_1, \dots, z_j - tz_k, \dots, z_n)\big|_{t=0} \\ &= -z_k \frac{\partial p}{\partial z_j}, \end{aligned} \tag{20.4}$$

and in particular

$$d\overline{S}^\ell(e_j)p(z) = -iz_j \frac{\partial p}{\partial z_j}. \tag{20.5}$$

We see that, for $|\alpha| = \ell$,

$$d\overline{S}^\ell(e_j)z^\alpha = -i\alpha_j z^\alpha. \tag{20.6}$$

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 (20.4) that the only weight vector annihilated by all raising operators is z_n^ℓ. Hence $\overline{S}^\ell$ is irreducible.

Note that

$$dS^\ell(X)f(z) = \frac{d}{dt} f(e^{tX^t}z)\big|_{t=0} = -d\overline{S}^\ell(X^t). \tag{20.7}$$

Hence

$$dS^\ell(e_{jk})p(z) = -d\overline{S}^\ell(e_{kj})p(z) = z_j \frac{\partial p}{\partial z_k}, \quad dS^\ell(e_j)p(z) = iz_j \frac{\partial p}{\partial z_j}. \tag{20.8}$$

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 $\mathrm{U}(n)$ on $\Lambda^\ell\mathbb{C}^n$ $(0 \le \ell \le n)$ by

$$\Lambda^\ell(g)\, v_1 \wedge \cdots \wedge v_\ell = gv_1 \wedge \cdots \wedge gv_\ell. \tag{20.9}$$

This is also well defined for $g \in \mathrm{Gl}(n, \mathbb{C})$, and we have, for $X \in \mathrm{M}(n, \mathbb{C})$,

$$\begin{aligned} d\Lambda^\ell(X)\, v_1 \wedge \cdots \wedge v_\ell &= \frac{d}{dt}\, e^{tX} v_1 \wedge \cdots \wedge e^{tX} v_\ell \big|_{t=0} \\ &= Xv_1 \wedge v_2 \wedge \cdots \wedge v_\ell + \cdots + v_1 \wedge \cdots \wedge v_{\ell-1} \wedge Xv_\ell. \end{aligned} \tag{20.10}$$

In this case, with $u_1, \dots, u_n$ as before denoting the standard basis of $\mathbb{C}^n$, if we set

$$u_J = u_{j_1} \wedge \cdots \wedge u_{j_\ell}, \quad J = (j_1, \dots, j_\ell), \tag{20.11}$$

with $j_1 < \cdots < j_\ell$, then

$$\begin{aligned} E_{jk} u_J = u_{j_1} \wedge \cdots \wedge u_{j_{\nu-1}} \wedge\, u_j \wedge u_{j_{\nu+1}} \wedge \cdots \wedge u_{j_\ell}, &\quad \text{if } k = j_\nu, \\ 0, &\quad \text{if } k \notin \{j_1, \dots, j_\ell\}, \end{aligned} \tag{20.12}$$

and

$$\begin{aligned} d\Lambda^\ell(e_j) u_J = i u_J &\ \text{ if }\ j \in \{j_1, \dots, j_\ell\}, \\ 0 &\ \text{ if }\ j \notin \{j_1, \dots, j_\ell\}. \end{aligned} \tag{20.13}$$

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 (20.12) 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 compute the dimensions of the representation spaces described above. A look at a standard basis shows that

$$\dim \Lambda^\ell\mathbb{C}^n = \binom{n}{\ell}. \tag{20.14}$$

As for $\mathcal{P}_\ell \approx S^\ell\mathbb{C}^n$, we have

$$\dim S^\ell\mathbb{C}^n = \#\{\beta \ge 0 : z^\beta = z_1^{\beta_1} \cdots z_n^{\beta_n},\ |\beta| = \ell\}. \tag{20.15}$$

If we set $\vartheta_n(\ell) = \dim S^\ell\mathbb{C}^n$, we can see that

$$\vartheta_{n+1}(\ell) = \vartheta_n(\ell) + \vartheta_n(\ell - 1) + \cdots + \vartheta_n(0). \tag{20.16}$$

Since clearly $\vartheta_1(\ell) = 1$, we see inductively that

$$\dim S^\ell \mathbb{C}^{n+1} = \binom{n+\ell}{n}. \tag{20.17}$$

We next reconsider the adjoint representation of $\mathrm{U}(n)$ on $\mathrm{M}(n, \mathbb{C})$, given by

$$\mathrm{Ad}(g)X = gXg^{-1}, \tag{20.18}$$

and the derived representation ad of $\mathrm{u}(n)$ on $\mathrm{M}(n, \mathbb{C})$, and its extension to the representation $\mathbb{C}\,\mathrm{u}(n) = \mathrm{M}(n, \mathbb{C})$ on $\mathrm{M}(n, \mathbb{C})$, given by

$$\mathrm{ad}\, X(Y) = [X, Y]. \tag{20.19}$$

These representations are not irreducible. We have a decomposition into invariant subspaces

$$M(n, \mathbb{C}) = \{cI\} \oplus \mathrm{M}_0(n, \mathbb{C}), \quad \mathrm{M}_0(n, \mathbb{C}) = \{X \in \mathrm{M}(n, \mathbb{C}) : \mathrm{Tr}\, X = 0\}. \tag{20.20}$$

Ad acts trivially on $\{cI\}$. We claim it acts irreducibly on $\mathrm{M}_0(n, \mathbb{C})$. The analysis below will establish this.

Using (19.11)–(19.13), we have the weight space (aka root space) decomposition

$$\begin{aligned} \mathrm{M}(n, \mathbb{C}) &= \mathbb{C}\mathfrak{h} \oplus \bigoplus_{j \neq k} \mathrm{Span}(e_{jk}) \\ &= \mathfrak{g}_0 \oplus \bigoplus_{j \neq k} \mathfrak{g}_{\omega_{jk}}, \end{aligned} \tag{20.21}$$

where, for $X = \sum x_j e_j$, $\omega_{jk}(X) = x_j - x_k$. Recall from (19.14) that $E_{\ell m} = \mathrm{ad}\, e_{\ell m}$ satisfies

$$E_{\ell m} : \mathfrak{g}_{\omega_{jk}} \longrightarrow \mathfrak{g}_{\omega_{jk} + \omega_{\ell m}}. \tag{20.22}$$

Now $\omega_{jk}(X) + \omega_{\ell m}(X) = x_j - x_k + x_\ell - x_m$, so, given that $\ell < m$ and $j \neq k$,

$$\omega_{jk} + \omega_{\ell m} \text{ is a root } \Longleftrightarrow k = \ell \text{ or } j = m. \tag{20.23}$$

Furthermore,

$$\begin{aligned} E_{\ell m} e_{jk} = [e_{\ell m}, e_{jk}] &= \delta_{mj} e_{\ell k} - \delta_{\ell k} e_{jm} \\ &= 0 \text{ provided } m \neq j \text{ and } \ell \neq k, \end{aligned} \tag{20.24}$$

and also

$$\begin{aligned} m = j \Rightarrow E_{\ell m} e_{jk} = e_{\ell k} - \delta_{\ell k} e_{jj} &= e_{\ell k} \text{ if } \ell \neq k \\ &\quad e_{\ell\ell} - e_{jj} \text{ if } \ell = k, \end{aligned} \tag{20.25}$$

and

$$\ell = k \Rightarrow E_{\ell m} e_{jk} = \delta_{mj} e_{\ell\ell} - e_{jm} = \begin{array}{l} -\, e_{jm} \text{ if } m \neq j, \\ e_{\ell\ell} - e_{jj} \text{ if } m = j. \end{array} \tag{20.26}$$

In conclusion, we deduce that

$$\begin{aligned} E_{\ell m} e_{jk} = 0,\ \forall\, \ell < m &\Longleftrightarrow m \neq j \text{ and } \ \ell \neq k,\ \forall\, \ell < m \\ &\Longleftrightarrow j = 1 \ \text{ and } \ k = n. \end{aligned} \tag{20.27}$$

Hence the subspace of $\oplus_{j \neq k} \mathfrak{g}_{\omega_{jk}}$ annihilated by all raising operators is $\mathfrak{g}_{\omega_{1n}} = \operatorname{Span}(e_{jn})$, 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 (19.12),

$$E_{\ell m} \Bigl(\sum x_j e_j \Bigr) = -i(x_\ell - x_m) e_{\ell m}, \tag{20.28}$$

which is 0 for all $\ell < m$ if and only if $x_1 = \cdots = x_n$, so $\mathcal{H}(\mathrm{Ad})$ is spanned by e_{1n} and $e_{11} + \cdots + e_{nn}$. These are weight vectors with weights

$$(1, 0, \dots, 0, -1) \ \text{ and } \ (0, \dots, 0). \tag{20.29}$$

This establishes the irreducibility of Ad on each of the two factors in (20.20).

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

$$X \in \mathfrak{g},\ Y \in \mathfrak{h} \Longrightarrow [X, Y] \in \mathfrak{h}, \tag{20.30}$$

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

$$\mathrm{M}_0(n, \mathbb{C}) \ \text{ is a simple Lie algebra.} \tag{20.31}$$

21. 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 §19, 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 21.1. *The elements of* $\mathfrak{h}'$ *that are highest weights of an irreducible representation of* U(n) *are precisely given by*

$$\{(k_1, \dots, k_n) : k_\nu \in \mathbb{Z},\ k_1 \ge \cdots \ge k_n\}. \tag{21.1}$$

Hence the set of equivalence classes of irreducible unitary representations of U(n) *is in natural one-to-one correspondence with the set (21.1).*

First we show that if $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathfrak{h}'$ is a highest weight, then it must have the form (21.1). 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 21.2. *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

$$E_\sigma^{-1}\, \mathrm{diag}\,(c_1, \dots, c_n)\, E_\sigma = \mathrm{diag}(c_{\sigma(1)}, \dots, c_{\sigma(n)}). \tag{21.2}$$

Now, given $\lambda = (\lambda_1, \dots, \lambda_n)$, the weight space V_λ has the characterization

$$v \in V_\lambda \Longleftrightarrow \pi\big(\mathrm{diag}(c_1, \dots, c_n)\big)v = (c_1^{\lambda_1} \cdots c_n^{\lambda_n})v. \tag{21.3}$$

It follows that

$$\pi(E_\sigma^{-1}) : V_\lambda \longrightarrow V_{\lambda(\sigma)} \tag{21.4}$$

is an isomorphism.

It remains to show that each element of the form (21.1) 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 $j \in \mathbb{Z}$,

$${}^j\pi(g) = (\det g)^j\, \pi(g) \ \text{ has highest weight } (k_1 + j, \dots, k_n + j), \tag{21.5}$$

with the same weight vector as π, as follows readily from (21.3).

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 $k_1 \geq \cdots \geq k_n \geq 0$. In this case we can write

$$(k_1, \dots, k_n) = j_1\gamma_1 + \cdots + j_n\gamma_n, \quad j_\nu \in \mathbb{Z}^+, \tag{21.6}$$

where γ_ℓ is the highest weight of the representation Λ^ℓ of $\mathrm{U}(n)$ discussed in (20.9)–(20.13), i.e.,

$$\gamma_\ell = (1, \dots, 1, 0, \dots, 0) \quad \text{(with } \ell \text{ ones)}. \tag{21.7}$$

The following gives the key construction.

Proposition 21.3. *A weight of the type (21.6) occurs as the highest weight of an irreducible component of the representation*

$$(\Lambda^1)^{\otimes j_1} \otimes \cdots \otimes (\Lambda^n)^{\otimes j_n} \tag{21.8}$$

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 21.3, we have the following.

Proposition 21.4. *Suppose* π_j *is a unitary representation of* $U(n)$ *on* V_j*, with highest weight* λ_j*. Then the representation*

$$\pi_1 \otimes \cdots \otimes \pi_K \quad \textit{on} \quad V_1 \otimes \cdots \otimes V_K \tag{21.9}$$

has highest weight $\lambda_1 + \cdots + \lambda_K$.

Proof. Indeed, suppose we have weight space decompositions

$$V_j = \bigoplus_{\mu \in \mathcal{S}_j \subset \mathfrak{h}'} V_{j\mu} \tag{21.10}$$

for π_j. Then $V_1 \otimes \cdots \otimes V_K$ is spanned by

$$V_{1\mu_1} \otimes \cdots \otimes V_{K\mu_K}, \quad \mu_\nu \in \mathcal{S}_\nu, \tag{21.11}$$

which consists of weight vectors for $\pi_1 \otimes \cdots \otimes \pi_K$, of weight $\mu_1 + \cdots + \mu_K$.

To return to Proposition 21.3, we have that (21.6) is the highest weight of the representation (21.8). 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 (21.8) has the desired highest weight. This finishes the proof of Theorem 21.1.

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 (21.1). In particular, from §20 we have

$$S^\ell \approx \mathcal{D}_{(\ell,0,\dots,0)}, \quad \overline{S}^\ell \approx \mathcal{D}_{(0,\dots,0,-\ell)}, \tag{21.12}$$

and

$$\Lambda^\ell \approx \mathcal{D}_{(1,\dots,1,0,\dots,0)}, \quad \text{(with } \ell \text{ ones)}, \quad 0 \le \ell \le n. \tag{21.13}$$

It is useful to record explicitly the content of (21.5) in this notation:

$$\mathcal{D}_{(k_1+j,\dots,k_n+j)}(g) = (\det g)^j\, \mathcal{D}_{(k_1,\dots,k_n)}(g). \tag{21.14}$$

Also from §20 we have

$$\mathrm{Ad} \approx D_{(1,0,\dots,0,-1)} \oplus D_{(0,\dots,0)}. \tag{21.15}$$

One simple corollary of Theorem 21.1 and Lemma 21.2 is the following.

Proposition 21.4. *All the one-dimensional representations of* $\mathrm{U}(n)$ *are equivalent to the representations*

$$\theta_j(g) = (\det g)^j, \tag{21.16}$$

for some $j \in \mathbb{Z}$*, in turn equivalent to* $D_{(j,\dots,j)}$.

Proof. A representation of $\mathrm{U}(n)$ 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 21.2, each $(k_{\sigma(1)},\dots,k_{\sigma(n)})$ must also be a weight. This forces $k_1 = \cdots = k_n = j$ (say), which gives (21.16).

22. 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 Appendix G 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

$$d\pi : \mathrm{u}(n) \longrightarrow \mathrm{End}(V), \tag{22.1}$$

which extends to a Lie algebra representation

$$d\pi : \mathrm{M}(n,\mathbb{C}) \longrightarrow \mathrm{End}(V), \tag{22.2}$$

which is also $\mathbb{C}$-linear. By Corollary 16.4, this exponentiates to a representation

$$\pi : \widetilde{\mathrm{Gl}}(n,\mathbb{C}) \longrightarrow \mathrm{Gl}(V), \tag{22.3}$$

where $\widetilde{\mathrm{Gl}}$ is the universal cover of Gl($n,\mathbb{C}$). In order to obtain

$$\pi : \mathrm{Gl}(n,\mathbb{C}) \longrightarrow \mathrm{Gl}(V), \tag{22.4}$$

we need to show that π in (22.3) has the property

$$\pi(g) = I, \quad \forall\, g \in \mathrm{Ker}\,\beta, \tag{22.5}$$

where

$$\beta : \widetilde{\mathrm{Gl}}(n,\mathbb{C}) \longrightarrow \mathrm{Gl}(n,\mathbb{C}) \tag{22.6}$$

is the natural covering map. To see this let

$$\alpha : \widetilde{\mathrm{U}}(n) \longrightarrow \mathrm{U}(n) \tag{22.7}$$

denote the natural projection of the universal group $\widetilde{\mathrm{U}}(n)$ onto U(n). We have a commutative diagram

$$\begin{array}{ccc} \widetilde{\mathrm{U}}(n) & \longrightarrow & \widetilde{\mathrm{Gl}}(n,\mathbb{C}) \\ {\scriptstyle\alpha}\downarrow & & {\scriptstyle\beta}\downarrow \\ \mathrm{U}(n) & \longrightarrow & \mathrm{Gl}(n,\mathbb{C}) \end{array}$$

The following result is key:

Lemma 22.1. *We have*

$$\operatorname{Ker}\alpha = \operatorname{Ker}\beta. \tag{22.8}$$

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:

$$\mathrm{U}(n) \hookrightarrow \mathrm{Gl}(n,\mathbb{C}) \ \text{ is a deformation retract.} \tag{22.9}$$

As for this result, this follows by polar decomposition:

$$\mathrm{Gl}(n,\mathbb{C}) \approx \mathrm{U}(n) \times \mathcal{P}(n), \tag{22.10}$$

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

$$A = UP, \quad U \in \mathrm{U}(n), \quad P \in \mathcal{P}(n), \tag{22.11}$$

uniquely defined by

$$P = (A^*A)^{1/2}, \quad U = AP^{-1}. \tag{22.12}$$

Returning to (22.5), 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 (22.8) holds. Thus (22.4) is established. Since

$$\mathrm{Exp} : \mathrm{M}(n,\mathbb{C}) \longrightarrow \mathrm{Gl}(n,\mathbb{C}) \tag{22.13}$$

is holomorphic, we also have that π in (22.3) 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

$$1 \longrightarrow K_n \longrightarrow S^1 \times \mathrm{SU}(n) \longrightarrow \mathrm{U}(n) \longrightarrow 1, \tag{22.14}$$

where $(\omega, g) \mapsto \omega g$ and

$$K_n = \{(\omega, g) \in S^1 \times \mathrm{SU}(n) : g = \omega^{-1} I,\ \omega^n = 1\}, \tag{22.15}$$

a cyclic group of order n, generated by

$$(\zeta^{-1}, \zeta I), \quad \zeta = e^{2\pi i/n}. \tag{22.16}$$

Let $\{\sigma_\alpha : \alpha \in \mathcal{I}\}$ denote a complete set of irreducible unitary representations of $\mathrm{SU}(n)$. By Proposition 11.11, 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

$$\pi_{m\alpha}(\omega, g) = \omega^m \sigma_\alpha(g). \tag{22.17}$$

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

$$\sigma_\alpha(\zeta I) = \zeta^m I, \tag{22.18}$$

when $\zeta = e^{2\pi i/n}$. Now for any $\alpha \in \mathcal{I}$, $\sigma_\alpha(\zeta I)$ is a scalar that is an nth root of unity, i.e.,

$$\sigma_\alpha(\zeta I) = \zeta^\mu I, \quad \mu = \mu(\alpha) \in \mathbb{Z}. \tag{22.19}$$

Then $\pi_{m\alpha}$ in (22.17) gives a representation of $\mathrm{U}(n)$ if and only if

$$m = \mu(\alpha) \mod n. \tag{22.20}$$

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 22.2. *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* $j \in \mathbb{Z}$*,*

$$\pi_2(g) = (\det g)^j\, \pi_1(g), \quad \forall\, g \in \mathrm{U}(n). \tag{22.21}$$

Hence the set of equivalence classes of irreducible unitary representations of $\mathrm{SU}(n)$ *is parametrized by*

$$\{(d_1, \dots, d_{n-1}, 0) : d_\nu \in \mathbb{Z},\ d_1 \ge d_2 \ge \cdots \ge d_{n-1} \ge 0\}. \tag{22.22}$$

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 (22.21) holds. To see this, suppose π_ℓ are as in (22.17), which we rewrite as

$$\pi_\ell(\omega g) = \omega^{m_\ell} \sigma_\ell(g), \quad \omega \in S^1,\ g \in \mathrm{SU}(n), \tag{22.23}$$

where σ_ℓ is an irreducible representation of $\mathrm{SU}(n)$, so, as in (22.18),

$$\sigma_\ell(\zeta I) = \zeta^{m_\ell} I, \tag{22.24}$$

where $\zeta = e^{2\pi i/n}$. We have

$$\pi_1 = \pi_2 \text{ on } \mathrm{SU}(n) \Longleftrightarrow \sigma_1 \equiv \sigma_2, \tag{22.25}$$

which implies

$$\pi_1(\omega g) = \omega^{m_1 - m_2} \pi_2(\omega g), \quad \forall\, \omega \in S^1,\ g \in \mathrm{SU}(n). \tag{22.26}$$

We claim

$$n|(m_1 - m_2), \quad \text{i.e., } m_1 - m_2 = nj, \ j \in \mathbb{Z}. \tag{22.27}$$

Since $\det \omega g = \omega^n$, this would give (22.21). To verify (22.27), we note that (22.24)–(22.25) give

$$\zeta^{m_1 - m_2} = 1, \tag{22.28}$$

from which (22.27) follows. This proves (22.21).

Recall from §21 the notation $\mathcal{D}_{(k_1,\dots,k_n)}$ for an irreducible representation of $\mathrm{U}(n)$ with highest weight $(k_1, \dots, k_n)$. We keep this notation for the restriction to $\mathrm{SU}(n)$, noting that

$$\mathcal{D}_{(k_1,\dots,k_n)} \approx \mathcal{D}_{(k_1+j,\dots,k_n+j)} \quad \text{on} \quad \mathrm{SU}(n), \quad \forall\, j \in \mathbb{Z}. \tag{22.29}$$

Note that the representations $D_{k/2}$ of $\mathrm{SU}(2)$ produced in §18, by (18.23), are given in this nomenclature as

$$\begin{aligned} D_{k/2} &= \mathcal{D}_{(0,-k)} \\ &\approx \mathcal{D}_{(k,0)}, \quad \text{on} \quad \mathrm{SU}(2), \end{aligned} \tag{22.30}$$

the last equivalence (special to $n = 2$) by (22.29).

23. Decomposition of $S^k \otimes \overline{S}^\ell$

Here we consider how to decompose the representation

$$S^k_\ell = S^k \otimes \overline{S}^\ell \tag{23.1}$$

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

$$S^k_\ell(g)p(z,w) = p(g^t z, g^{-1}w), \tag{23.2}$$

for $g \in \mathrm{U}(n)$, extending holomorphically to $g \in \mathrm{Gl}(n,\mathbb{C})$. This induces an action $dS^k_\ell(X)$ on such polynomials, for X in $\mathrm{u}(n)$, and its complexification $\mathrm{M}(n,\mathbb{C})$. Parallel to (20.4) and (20.8) we have

$$dS^k_\ell(e_{\mu\nu}) = z_\mu \frac{\partial}{\partial z_\nu} - w_\nu \frac{\partial}{\partial w_\mu}. \tag{23.3}$$

To decompose S^k_ℓ 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 (23.3) with $\mu < \nu$. The following result accomplishes this.

Lemma 23.1. *If $p(z,w)$ is a polynomial annihilated by all operators of the form (23.3) with $\mu < \nu$, then*

$$p(z,w) = q(z_1, w_n, z \cdot w), \tag{23.4}$$

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:

$$z_\nu \mapsto z_\nu + t z_\mu, \quad w_\mu \mapsto w_\mu - t w_\nu, \quad \mu < \nu. \tag{23.5}$$

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

$$p(z,w) = p((z_1, 0, \dots, 0), \tilde{w}), \tag{23.6}$$

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 (23.5), 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 (23.4) implies

$$p(z, w) = p((z_1, 0, \dots, 0, \zeta), (0, \dots, 0, w_n)). \tag{23.7}$$

Again $z \cdot w = (z_1, 0, \dots, 0, \zeta) \cdot (0, \dots, 0, w_n) = \zeta w_n$, so

$$\zeta = \frac{z \cdot w}{w_n}. \tag{23.8}$$

Since p is a polynomial in z and w, we have (23.3).

It follows from Lemma 23.1 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

$$\psi_{k\ell\mu}(z, w) = z_1^{k-\mu} w_n^{\ell-\mu} (z_1 w_1 + \cdots + z_n w_n)^\mu, \quad 0 \le \mu \le k \wedge \ell. \tag{23.9}$$

Each of these elements is a weight vector for S^k_ℓ. In fact,

$$\begin{aligned} dS^k_\ell(e_\nu) \psi_{k\ell\mu}(z, w) &= i\Big(z_\nu \frac{\partial}{\partial z_\nu} - w_\nu \frac{\partial}{\partial w_\nu}\Big) \psi_{k\ell\mu}(z, w) \\ &= i[(k-\mu)\delta_{\nu 1} - (\ell - \mu)\delta_{\nu n}] \psi_{k\ell\mu}(z, w). \end{aligned} \tag{23.10}$$

The weight so obtained is

$$(k-\mu, 0, \dots, 0, \mu - \ell), \quad 0 \le \mu \le k \wedge \ell. \tag{23.11}$$

These calculations establish the following.

Proposition 23.2. *For $k, \ell \ge 0$ we have*

$$S^k \otimes \overline{S}^\ell \approx \bigoplus_{0 \le \mu \le k \wedge \ell} \mathcal{D}_{(k-\mu, 0, \dots, 0, \mu - \ell)}, \tag{23.12}$$

as representations of $\mathrm{U}(n)$. The highest weight vectors for the irreducible components on the right side of (23.12) are given by (23.9).

Remark 1. For $n = 2$, this result captures the Clebsch-Gordon series (18.47), in view of the identities in (22.24).

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 (23.12) with (21.15).

Note that the highest weight that occurs in (23.12) 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 23.3. *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*

$$\mathcal{P}^\#_{k\ell} = \Big\{ p(z,w) \in \mathcal{P}_k \otimes \mathcal{P}_\ell : \sum_{j=1}^n \frac{\partial^2 p}{\partial z_j \partial w_j} = 0 \Big\}. \tag{23.13}$$

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 (23.3). 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 (23.9). A calculation gives

$$\sum_{j=1}^n \frac{\partial^2}{\partial z_j \partial w_j} \psi_{k\ell\mu}(z,w) = \mu(n-1+k+\ell-\mu)\, \frac{\psi_{k\ell\mu}(z,w)}{z\cdot w}. \tag{23.14}$$

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

$$\mathcal{D}_{(k,0,\dots,0,-\ell)} \ \text{ is realized on } \ \mathcal{P}^\#_{k\ell}. \tag{23.15}$$

Let us specialize to $n=3$. We have representations

$$\mathcal{D}_{(k,0,-\ell)} \ \text{ of } \ \mathrm{U}(3) \ \text{ on } \ \mathcal{P}^\#_{k\ell}. \tag{23.16}$$

Multiplying by $(\det g)^j$ gives representations

$$\mathcal{D}_{(k+j,j,j-\ell)} \ \text{ of } \ \mathrm{U}(3) \ \text{ on } \ \mathcal{P}^\#_{k\ell}. \tag{23.17}$$

The results of §21 show that (23.17) produces a complete set of irreducible representations of U(3).

We can produce an alternative realization of (23.15) as follows. An element of $\mathcal{P}_k \otimes \mathcal{P}_\ell$ can be written

$$p(z,w) = \sum A^{j_1\cdots j_k}_{i_1\cdots i_\ell} z_{j_1}\cdots z_{j_k} w_{i_1}\cdots w_{i_\ell}, \tag{23.18}$$

with $A^{\cdots}_{\cdots}$ symmetric in the js and the is. A computation gives

$$\sum_{j=1}^n \frac{\partial^2}{\partial z_j \partial w_j} p(z,w) = \sum_{\nu,i,j} A^{\nu j_2\cdots j_k}_{\nu i_2\cdots i_\ell} z_{j_2}\cdots z_{j_k} w_{i_2}\cdots w_{i_\ell}. \tag{23.19}$$

In other words, we have $\mathrm{U}(n)$ acting on

$$\mathcal{P}_k \otimes \mathcal{P}_\ell \approx (S^k\mathbb{C}^n) \otimes (S^\ell \mathbb{C}^n)', \tag{23.20}$$

and

$$\mathcal{P}^\#_{k\ell} \approx \big\{ A^{j_1\cdots j_k}_{i_1\cdots i_\ell} \in (S^k\mathbb{C}^n)\otimes (S^\ell\mathbb{C}^n)' : A^{\nu j_2\cdots j_k}_{\nu i_2\cdots i_\ell} \equiv 0 \big\}, \tag{23.21}$$

where the summation convention is indicated over ν.

24. 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

$$\mathcal{A} = \text{algebra generated by } \pi(g) : g \in G, \quad \mathcal{A} \subset \operatorname{End}(V), \tag{24.1}$$

and its *commutant*, defined by

$$\mathcal{A}' = \{B \in \operatorname{End}(V) : BA = AB, \ \forall\ A \in \mathcal{A}\}. \tag{24.2}$$

As we know, V can be decomposed into irreducible subspaces. Say

$$V = \bigoplus_{j=1}^{k} n_j V_j, \tag{24.3}$$

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 (24.3) 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.,

$$V = \bigoplus_{j=1}^{k} V_j \otimes W_j, \tag{24.4}$$

with $\pi_j = \pi|_{V_j \otimes W_j}$ given by

$$\pi_j(g) = U_j(g) \otimes I, \tag{24.5}$$

and

$$U_j \ \text{ irreducible on } \ V_j, \quad j_1 \neq j_2 \Rightarrow U_{j_1} \ \text{ inequivalent to } \ U_{j_2}. \tag{24.6}$$

The following result records some important structure.

Proposition 24.1. *In the set-up described above,*

$$\mathcal{A} = \Big\{ \bigoplus_{j=1}^{k} A_j \otimes I : A_j \in \operatorname{End}(V_j) \Big\}. \tag{24.7}$$

If we set

$$\mathcal{B} = \mathcal{A}', \tag{24.8}$$

then

$$\mathcal{B} = \Big\{ \bigoplus_{j=1}^{k} I \otimes B_j : B_j \in \operatorname{End}(W_j) \Big\}. \tag{24.9}$$

Furthermore,

$$\mathcal{B}' = \mathcal{A}. \tag{24.10}$$

Proof. It is immediate from (24.5) that every element of $\mathcal{A}$ has the form given on the right side of (24.7). 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 $e_{\ell m}^{(j)} \in \operatorname{End}(V_j)$ 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 $U_j(g) \in \operatorname{End}(V_j)$ with respect to this basis. It follows from the Weyl orthogonality relations (6.6)–(6.7) that

$$d_j \int_G U_{\ell m}^{(j)}(g)\, \pi(g)\, dg = e_{\ell m}^{(j)} \otimes I, \tag{24.11}$$

so every element on the right side of (24.7) 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 $\operatorname{End}(V)$ must be closed). Thus we have (24.7).

To prove (24.9), first note that whenever $\mathcal{A}$ is given by (24.7), then clearly the right side of (24.9) 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 (24.7), $\mathcal{P}_j \in \mathcal{A}$. 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

$$\widetilde{B}_j : V_j \otimes W_j \to V_j \otimes W_j, \quad \widetilde{B}_j (A \otimes I) = (A \otimes I) \widetilde{B}_j, \quad \forall\ A \in \operatorname{End}(V_j). \tag{24.12}$$

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

$$\widetilde{B}_j (u_\ell^{(j)} \otimes w_m^{(j)}) = \sum_\mu \beta_{\ell m}^{\mu}\, u_\ell^{(j)} \otimes w_\mu^{(j)}. \tag{24.13}$$

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 $B_j \in \operatorname{End}(W_j)$, proving (24.9). The way we got (24.9) from (24.7) immediately gives (24.10).

The result (24.10) is a special case of a result known as the double commutant theorem. It holds when $\mathcal{A}$ is a subalgebra of $\operatorname{End}(V)$ 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 infinite dimensional. See [Sak].

Now we add structure by bringing in two groups, acting on V.

Proposition 24.2. *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*

$$\begin{aligned}\mathcal{A} &= \textit{algebra generated by } \pi(g),\ g \in G,\\ \mathcal{B} &= \textit{algebra generated by } \tau(k),\ k \in K.\end{aligned} \tag{24.14}$$

Assume

$$\mathcal{A}' = \mathcal{B}. \tag{24.15}$$

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

$$V = \bigoplus_{j=1}^{k} V_j \otimes W_j, \quad k = \#(\mathcal{S}_\pi) = \#(\mathcal{S}_\tau), \tag{24.16}$$

such that

$$\pi(g) = \bigoplus_{j=1}^{k} \alpha_j(g) \otimes I, \quad \tau(k) = \bigoplus_{j=1}^{k} I \otimes \beta_{Q(j)}(k). \tag{24.17}$$

Proof. The representation π of G decomposes as in (24.3). 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 24.1, $\mathcal{B}' = \mathcal{A}$. 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

$$V \approx V_1 \otimes W_1, \tag{24.18}$$

with

$$\pi(g) = \alpha_1(g) \otimes I, \quad \tau(k) = I \otimes \beta_1(k), \tag{24.19}$$

given (24.15). In fact Proposition 24.1 gives a tensor product decomposition (24.18) such that (24.19) holds for $\pi(g)$. Then the fact that $\mathcal{A}'$ is given by (24.9) also puts $\tau(k)$ in the general form indicated in (24.19), 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 24.1, 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.

When compact G and K act on V as in Proposition 24.2, with (24.15) 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 §25.

Note that, in the setting of Proposition 24.2, $\pi(g)\tau(k)$ gives a representation of $G \times K$ on V, and (24.17) gives

$$\pi(g)\tau(k) = \bigoplus_{j=1}^{k} \alpha_j(g) \otimes \beta_{Q(j)}(k). \tag{24.20}$$

In particular, taking traces gives

$$\mathrm{Tr}\, \pi(g)\tau(k) = \sum_{j=1}^{k} \chi_{\alpha_j}(g)\, \chi_{\beta_{Q(j)}}(k). \tag{24.21}$$

25. The first fundamental theorem of invariant theory

The group $\mathrm{U}(n)$ acts on $\otimes^k \mathbb{C}^n$ via

$$\otimes^k g(v_1 \otimes \cdots \otimes v_k) = gv_1 \otimes \cdots \otimes gv_k, \quad g \in \mathrm{U}(n),\ v_\nu \in \mathbb{C}^n. \tag{25.1}$$

In addition, the permutation group S_k acts on $\otimes^k \mathbb{C}^n$ via

$$\tau(\sigma)(v_1 \otimes \cdots \otimes v_k) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}, \quad \sigma \in S_k. \tag{25.2}$$

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 25.1. *The groups S_k and $\mathrm{U}(n)$ act as a dual pair on $\otimes^k \mathbb{C}^n$.*

To restate this, let

$$\begin{aligned} \mathcal{A} &= \text{algebra generated by } \tau(\sigma),\ \sigma \in S_k, \quad \mathcal{A} \subset \mathrm{End}(\otimes^k \mathbb{C}^n), \\ \mathcal{B} &= \text{algebra generated by } \otimes^k g,\ g \in \mathrm{U}(n). \end{aligned} \tag{25.3}$$

It is clear that $\mathcal{B} \subset \mathcal{A}'$ and $\mathcal{A} \subset \mathcal{B}'$, as we have already mentioned. To prove Proposition 25.1, we will show that

$$\mathcal{A}' = \mathcal{B}. \tag{25.4}$$

In view of Proposition 24.1, this gives also

$$\mathcal{B}' = \mathcal{A}. \tag{25.5}$$

Our treatment follows [Si].

To begin our analysis of $\mathcal{A}'$, we note that

$$\mathrm{End}(\otimes^k \mathbb{C}^n) \approx \otimes^k \mathrm{End}(\mathbb{C}^n), \tag{25.6}$$

via

$$A_1 \otimes \cdots \otimes A_k(v_1 \otimes \cdots \otimes v_k) = A_1 v_1 \otimes \cdots \otimes A_k v_k. \tag{25.7}$$

In fact, (25.7) 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 (25.6) are equal. We let $\sigma \in S_k$ act on $\otimes^k \mathrm{End}(\mathbb{C}^n)$ by

$$T(\sigma)\, A_1 \otimes \cdots \otimes A_k = A_{\sigma(1)} \otimes \cdots \otimes A_{\sigma(k)}. \tag{25.8}$$

Lemma 25.2. *Given* $X \in \mathrm{End}(\otimes^k \mathbb{C}^n),\ \sigma \in S_k$,

$$T(\sigma)X = \tau(\sigma)X\tau(\sigma)^{-1}. \tag{25.9}$$

Proof. It suffices to check (25.9) for $X = A_1 \otimes \cdots \otimes A_k$. Then

$$\begin{aligned} \tau(\sigma)(A_1 \otimes \cdots \otimes A_k)\tau(\sigma)^{-1}&(v_1 \otimes \cdots \otimes v_k) \\ &= \tau(\sigma)\, A_1 \otimes \cdots \otimes A_k(v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}) \\ &= \tau(\sigma)\, A_1 v_{\sigma^{-1}(1)} \otimes \cdots \otimes A_k v_{\sigma^{-1}(k)} \\ &= A_{\sigma(1)} v_1 \otimes \cdots \otimes A_{\sigma(k)} v_k, \end{aligned} \tag{25.10}$$

which gives (25.9).

At this point we have

$$\begin{aligned} \mathcal{A}' &= \{X : \tau(\sigma)X = X\tau(\sigma),\ \forall\, \sigma \in S_k\} \\ &= \{X : T(\sigma)X = X,\ \forall\, \sigma \in S_k\} \\ &= S^k\, \mathrm{End}(\mathbb{C}^n). \end{aligned} \tag{25.11}$$

The next lemma is an exercise in linear algebra.

Lemma 25.3. *If* W *is a finite-dimensional vector space,*

$$S^k W = \mathit{Span}\{a \otimes \cdots \otimes a : a \in W\}. \tag{25.12}$$

Hence we have from (25.11)

$$\mathcal{A}' = \mathrm{Span}\,\{A \otimes \cdots \otimes A : A \in \mathrm{End}(\mathbb{C}^n)\}. \tag{25.13}$$

By comparison,

$$\mathcal{B} = \mathrm{Span}\,\{g \otimes \cdots \otimes g : g \in \mathrm{U}(n)\}. \tag{25.14}$$

To prove (25.4), it remains to show that the spaces (25.13) and (25.14) coincide. To see this, note that, for $Y \in \mathrm{u}(n)$,

$$d \otimes^k (Y)(v_1 \otimes \cdots \otimes v_k) = Yv_1 \otimes v_2 \otimes \cdots \otimes v_k + \cdots + v_1 \otimes \cdots \otimes v_{k-1} \otimes Yv_k \tag{25.15}$$

has the property that

$$d \otimes^k (Y) \in \mathcal{B}, \tag{25.16}$$

since this is a limit of difference quotients of elements of $\mathcal{B}$, by (25.14). Then (25.16) holds for all $Y \in \mathbb{C}\,\mathrm{u}(n) = \mathrm{End}(\mathbb{C}^n)$, and exponentiating this gives

$$A \otimes \cdots \otimes A \in \mathcal{B}, \quad \forall\ A \in \mathrm{Gl}(n, \mathbb{C}). \tag{25.17}$$

Since $\mathrm{Gl}(n, \mathbb{C})$ is dense in $\mathrm{End}(\mathbb{C}^n)$, we have $A \otimes \cdots \otimes A \in \mathcal{B}$ for all $A \in \mathrm{End}(\mathbb{C}^n)$, so in view of (25.13) we now have $\mathcal{A}' = \mathcal{B}$, as advertised in (25.4), and Proposition 25.1 is proven.

It is useful to restate the result (25.5) as follows. Define the representation ϑ_{nk} of $\mathrm{U}(n)$ on $\mathrm{End}(\otimes^k \mathbb{C}^n)$ by

$$\vartheta_{nk}(g)A = (\otimes^k g)A(\otimes^k g^{-1}), \quad g \in \mathrm{U}(n),\ A \in \mathrm{End}(\otimes^k \mathbb{C}^n). \tag{25.18}$$

Denote by $\mathcal{E}_{nk}$ the subspace of $\mathrm{End}(\otimes^k \mathbb{C}^n)$ on which ϑ_{nk} acts trivially. Then (25.5) implies that $\mathcal{E}_{nk}$ is spanned by the operators $\tau(\sigma),\ \sigma \in S_k$, given by (25.2). To elaborate, (25.2) yields a linear map

$$\tau^{\#}_{nk} : \ell^1(S_k) \longrightarrow \mathrm{End}(\otimes^k \mathbb{C}^n), \tag{25.19}$$

as a special case of the construction in §10, and

$$\mathcal{E}_{nk} = \text{Range of } \ \tau^{\#}_{nk}. \tag{25.20}$$

Note that

$$\vartheta_{nk} \approx (\otimes^k) \otimes (\overline{\otimes}^k), \tag{25.21}$$

acting on $(\otimes^k \mathbb{C}^n) \otimes (\otimes^k \mathbb{C}^n)'$, via

$$g \cdot (v_1 \otimes \cdots \otimes v_k \otimes w_1 \otimes \cdots \otimes w_k) = gv_1 \otimes \cdots \otimes gv_k \otimes g'w_1 \otimes \cdots \otimes g'w_k, \tag{25.22}$$

where $g' = (g^t)^{-1}$, so $g \in \mathrm{U}(n) \Rightarrow g' = \overline{g}$. Hence $\mathcal{E}_{nk}$ is isomorphic to the space

$$E_{nk} \subset (\otimes^k \mathbb{C}^n) \otimes (\otimes^k \mathbb{C}^n)' \ \text{ where } \ \mathrm{U}(n) \ \text{ acts trivially.} \tag{25.23}$$

The following restatement of (25.20) is known as the first fundamental theorem of invariant theory (for unitary invariants).

Proposition 25.4. *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*

$$t_\sigma(w_1 \otimes \cdots \otimes w_k, v_1 \otimes \cdots \otimes v_k) = \langle v_1, w_{\sigma(1)} \rangle \cdots \langle v_k, w_{\sigma(k)} \rangle, \tag{25.24}$$

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 5.5 that the orthogonal projection of $(\otimes^k \mathbb{C}^n) \otimes (\otimes^k \mathbb{C}^n)'$ onto E_{nk} is given by

$$P_{nk} = \int\limits_{\mathrm{U}(n)} (\otimes^k g) \otimes (\otimes^k \overline{g})\, dg. \tag{25.25}$$

Hence

$$\dim E_{nk} = \mathrm{Tr}\, P_{nk} = \int\limits_{\mathrm{U}(n)} |\,\mathrm{Tr}\, g|^{2k}\, dg. \tag{25.26}$$

A computation of this dimension is of interest; "half" the cases are elementary:

Proposition 25.5. *If $k \le n$, then the map $\tau_{nk}^{\#}$ in (25.19) is injective. Hence*

$$k \le n \Longrightarrow \dim E_{nk} = k! \tag{25.27}$$

Proof. Let $\{u_1, \dots, u_n\}$ denote the standard basis of $\mathbb{C}^n$. If $k \le n$, then the elements

$$\tau(\sigma)\, u_1 \otimes \cdots \otimes u_k = u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(k)}, \quad \sigma \in S_k, \tag{25.28}$$

are linearly independent in $\otimes^k \mathbb{C}^n$, which implies injectivity of $\tau_{nk}^{\#}$.

If $k > n$, neither the left nor the right side of (25.26) is easy to evaluate. We will make further comments on this in §26.

26. Decomposition of $\otimes^k \mathbb{C}^n$

Since S_k and $\mathrm{U}(n)$ act as a dual pair on $\otimes^k \mathbb{C}^n$, Proposition 24.2 is applicable. Hence $\otimes^k \mathbb{C}^n$ has a decomposition of the form (24.16), with S_k acting irreducibly on V_j and $\mathrm{U}(n)$ 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 §21 irreducible representations of $\mathrm{U}(n)$ that were produced on subspaces of $\otimes^k \mathbb{C}^n$. Namely we have the representation $\mathcal{D}_\lambda$ of $\mathrm{U}(n)$ on the linear subspace of

$$(\Lambda^1 \mathbb{C}^n)^{\otimes j_1} \otimes \cdots \otimes (\Lambda^n \mathbb{C}^n)^{\otimes j_n}, \tag{26.1}$$

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

$$\lambda = j_1 \gamma_1 + \cdots + j_n \gamma_n = (r_1, \dots, r_n), \quad r_\nu = j_\nu + \cdots + j_n. \tag{26.2}$$

The cases that arise for which (26.1) is a subspace of $\otimes^k \mathbb{C}^n$ are those for which

$$r_1 \geq \cdots \geq r_n \geq 0, \quad r_1 + \cdots + r_n = k, \quad r_\nu \in \mathbb{Z}^+. \tag{26.3}$$

We denote by F_{nk} the set of n-tuples $\lambda = (r_1, \dots, r_n)$ satisfying (26.3). Such λ is called a Young frame on k for $\mathrm{U}(n)$. It turns out that precisely these representations $\mathcal{D}_\lambda$ of $\mathrm{U}(n)$ occur in the decomposition (24.16) 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 26.1. 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 $\mathcal{F}_{nk}$ denote the Young diagram just described. We set

$$\begin{aligned} \mathcal{P}_\lambda &= \{\sigma \in S_k : \sigma \text{ preserves each row of } \mathcal{F}_{nk}\}, \\ \mathcal{Q}_\lambda &= \{\sigma \in S_k : \sigma \text{ preserves each column of } \mathcal{F}_{nk}\}. \end{aligned} \tag{26.4}$$

We define $p_\lambda, q_\lambda, c_\lambda \in \ell^1(S_k)$ by

$$p_\lambda = \sum_{\sigma \in \mathcal{P}_\lambda} \sigma, \quad q_\lambda = \sum_{\sigma \in \mathcal{Q}_\lambda} (\operatorname{sgn} \sigma) \sigma, \quad c_\lambda = p_\lambda * q_\lambda. \tag{26.5}$$

Proposition 26.1. *In the convolution algebra* $\ell^1(S_k)$,

$$c_\lambda * c_\lambda = \mu\, c_\lambda, \tag{26.6}$$

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 26.2. *The representation* $\mathcal{S}_\lambda$ *of* S_k *on* E_λ *is irreducible.*

The following result is due to H. Weyl.

Proposition 26.3. *For the representations* τ *of* S_k *and* $\otimes^k$ *of* $\mathrm{U}(n)$ *on* $\otimes^k \mathbb{C}^n$, *we have*

$$\tau \cdot \otimes^k \approx \bigoplus_{\lambda \in F_{nk}} \mathcal{S}_\lambda \otimes \mathcal{D}_\lambda. \tag{26.7}$$

Complete proofs of Propositions 26.1–26.3 can be found in [FH] and [Si]. Let us explicitate how much of Proposition 26.3 has been proven in these notes. That $\tau \cdot \otimes^k$ has a decomposition of the form (26.7), as λ ranges over *some* set of maximal weights for irreducible representations of $\mathrm{U}(n)$, is a consequence of Proposition 24.2, combined with Proposition 25.1. That each $\lambda \in F_{nk}$ arises in this decomposition follows from our observations about (26.1). For a complete proof of Proposition 26.3, it remains to establish two things:

(i) There are no other highest weights that should appear in (26.7).
(ii) The irreducible representation of S_k that is paired with $\mathcal{D}_\lambda$, whose existence is established in Proposition 24.2, is in fact the representation $\mathcal{S}_\lambda$ described above.

Proofs of these results, which can be found on p. 251 if [Si], rely strictly on the representation theory of S_k.

We obtain some consequences of (26.7) for characters. Let us set

$$\chi^U_\lambda(g) = \mathrm{Tr}\, \mathcal{D}_\lambda(g), \quad \chi^S_\lambda(\sigma) = \mathrm{Tr}\, \mathcal{S}_\lambda(\sigma). \tag{26.8}$$

Then (26.7) implies

$$\mathrm{Tr}\big(\tau(\sigma) \cdot \otimes^k g\big) = \sum_{\lambda \in F_{nk}} \chi^S_\lambda(\sigma)\, \chi^U_\lambda(g), \quad \sigma \in S_k,\ g \in \mathrm{U}(n). \tag{26.9}$$

In particular, taking $\sigma = e$, the identity element of S_k, gives

$$(\mathrm{Tr}\, g)^k = \sum_{\lambda \in F_{nk}} f^\lambda\ \chi^U_\lambda(g), \tag{26.10}$$

where

$$f^\lambda = \chi^S_\lambda(e) \tag{26.11}$$

is the dimension of the representation space for $\mathcal{S}_\lambda$. The Weyl orthogonality relations imply

$$\int\limits_{\mathrm{U}(n)} |\operatorname{Tr} g|^{2k}\, dg = \sum_{\lambda \in F_{nk}} (f^\lambda)^2. \tag{26.12}$$

Recall that the left side of (26.12) also satisfies (25.26). In other words, (26.12) is equal to

$$\dim E_{nk} = k! - \dim \operatorname{Ker} \tau^\#_{nk}, \tag{26.13}$$

where $\tau^\#_{nk} : \ell^1(S_k) \to \operatorname{End}(\otimes^k \mathbb{C}^n)$ is as in (25.19).

We illustrate the decomposition (26.7) in the case $k = n = 3$. The three Young diagrams that arise in $\mathcal{F}_{33}$ are pictured in Figure 26.2. They correspond, respectively, to

$$\lambda = (1,1,1), \quad \lambda = (2,1,0), \quad \lambda = (3,0,0). \tag{26.14}$$

The representations of S_3 so obtained are

$$\mathcal{S}_{(1,1,1)} = \operatorname{sgn}, \quad \mathcal{S}_{(2,1,0)} = \pi^3_S, \quad \mathcal{S}_{(3,0,0)} = 1. \tag{26.15}$$

The representation π^3_S, discussed in Lemma 9.1, represents S_3 as the group of symmetries of an equilateral triangle. Hence (26.7) leads to

$$\otimes^3 \mathbb{C}^3 \approx \Lambda^3 \mathbb{C}^3 \oplus V_8 \oplus V_8 \oplus S^3 \mathbb{C}^3, \tag{26.16}$$

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

$$\dim S^3 \mathbb{C}^3 = 10, \tag{26.17}$$

as a special case of the general result

$$\dim S^k \mathbb{C}^{n+1} = \binom{n+k}{n}, \tag{26.18}$$

as shown in §20. Hence

$$\dim V_8 = 8. \tag{26.19}$$

In fact, more can be said about V_8. The adjoint representation of U(3) on $\mathrm{M}(3,\mathbb{C})$ is

$$S^1 \otimes \overline{S}^1 \approx \mathcal{D}_{(0,0,0)} \oplus \mathcal{D}_{(1,0,-1)}, \tag{26.20}$$

by (23.12). Here $\mathcal{D}_{(0,0,0)}$ is a one-dimensional representation and $\mathcal{D}_{(1,0,-1)}$ is an 8-dimensional representation, acting on

$$\{A \in \mathrm{M}(3,\mathbb{C}) : \operatorname{Tr} A = 0\}. \tag{26.21}$$

On the other hand, by (21.5),

$$\mathcal{D}_{(2,1,0)}(g) = (\det g)\mathcal{D}_{(1,0,-1)}(g). \tag{26.22}$$

Let us turn to $\otimes^4\mathbb{C}^3$. The five Young diagrams that arise in $\mathcal{F}_{44}$ are pictured in Figure 26.3, but the first one does not belong to $\mathcal{F}_{34}$, though the others do. They correspond, respectively, to

$$\lambda = (2,1,1), \quad \lambda = (2,2,0), \quad \lambda = (3,1,0), \quad \lambda = (4,0,0). \tag{26.23}$$

Recall the representations of S_4 as described in §9:

$$1, \quad \mathrm{sgn}, \quad \pi_S^4, \quad \pi_Q^4, \quad \pi_S^3 \circ \beta, \tag{26.24}$$

where $\beta : S_4 \to S_3$ is as in (9.20). Of these, of course

$$\mathcal{S}_{(4,0,0)} = 1. \tag{26.25}$$

As we have noted, the representation sgn of S_4 does not arise in $\otimes^4\mathbb{C}^3$. We claim that

$$\mathcal{S}_{(2,2,0)} \approx \pi_S^3 \circ \beta, \tag{26.26}$$

and that

$$\mathcal{D}_{(2,2,0)} \text{ acts on } S^2(\Lambda^2\mathbb{C}^3). \tag{26.27}$$

Note that

$$\mathcal{D}_{(2,2,0)}(g) = (\det g)^2\, \overline{S}^2(g). \tag{26.28}$$

We ask the reader to pair π_S^4 and π_Q^4 with the two remaining weights listed in (26.23), 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 (26.12) for $k = 4, n = 3$. The right side involves all the representations of S_4 but sgn, which is one dimensional, so we have

$$\int_{\mathrm{U}(3)} |\operatorname{Tr} g|^8\, dg = 23. \tag{26.29}$$

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 $k \geq 4$.

27. 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:

$$\int_G f(x)\,dx = \frac{1}{w}\int_T \Big(\int_G f(g^{-1}kg)\,dg\Big) |\det(I - \operatorname{Ad} k)_{\mathfrak{g}/\mathfrak{h}}|\,dk. \tag{27.1}$$

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 (27.1) arises from a study of

$$F : G \times T \longrightarrow G, \quad F(g, h) = ghg^{-1}, \tag{27.2}$$

and its induced action

$$\widetilde{F} : (G/T) \times T \longrightarrow G. \tag{27.3}$$

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

$$DL_{gh^{-1}g^{-1}}(ghg^{-1}) \circ DF(g,h) \circ (DL_g(e) \times DL_h(e)), \tag{27.4}$$

where $L_g(x) = gx$. Note that (27.4) is equal to $DG(e,e)$, where

$$\begin{aligned} G(x,z) &= L_{gh^{-1}g^{-1}} \circ F \circ (L_g \times L_h)(x,z) \\ &= gh^{-1}xhzx^{-1}g^{-1}. \end{aligned} \tag{27.5}$$

Note that $G(e,e) = e$; we compute

$$DG(e,e) : \mathfrak{g} \oplus \mathfrak{h} \longrightarrow \mathfrak{g}. \tag{27.6}$$

First, with $Z \in \mathfrak{h}$, $z(t)$ a curve in T such that $z(0) = e$, $z'(0) = Z$, we have

$$\begin{aligned} D_2G(e,e)Z &= \frac{d}{dt}\,gz(t)g^{-1}\big|_{t=0} \\ &= \operatorname{Ad} g(Z), \end{aligned} \tag{27.7}$$

the last identity following from (14.2), or alternatively from (14.9). Next, with $X \in \mathfrak{g}$, $x(t)$ a curve in G such that $x(0) = e$, $x'(0) = X$, we have

$$\begin{aligned} D_1 G(e,e)X &= \frac{d}{dt}\, gh^{-1}x(t)hx(t)^{-1}g^{-1}\big|_{t=0} \\ &= \operatorname{Ad} g\, DK(e)X, \end{aligned} \tag{27.8}$$

where

$$K(x) = h^{-1}xhx^{-1}, \tag{27.9}$$

so

$$\begin{aligned} DK(e)X &= \frac{d}{dt}\, h^{-1}x(t)hx(t)^{-1}\big|_{t=0} \\ &= h^{-1}Xh - X \\ &= (\operatorname{Ad} h^{-1} - I)X. \end{aligned} \tag{27.10}$$

(Here we take $G \subset \operatorname{End}(\mathbb{C}^m)$, to simplify the calculation.) Putting together (27.7), (27.8), and (27.10), we have

$$DG(e,e)(X,Z) = \operatorname{Ad} g(\operatorname{Ad} h^{-1} - I)X + \operatorname{Ad} g\, Z. \tag{27.11}$$

Now we can take $X \in \mathfrak{g}/\mathfrak{h}$. Thus we have

$$\det D\widetilde{F}(g,h) = \det(\operatorname{Ad} h^{-1} - I)_{\mathfrak{g}/\mathfrak{h}} = \det(I - \operatorname{Ad} h)_{\mathfrak{g}/\mathfrak{h}}. \tag{27.12}$$

The formula (27.1) is now a consequence of the following assertion:

Lemma 27.1. *The map $\widetilde{F}$ in (27.3) 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*

$$\widetilde{F}^{-1}(g) \subset (G/T) \times T \quad \text{has } w \text{ elements}, \quad \forall\, g \in \mathcal{O}.$$

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 (19.4). 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 27.1 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 [DK], [Si], or another source for a treatment of the general case.

We give an explicit formula for the right side of (27.12) when $G = \mathrm{U}(n)$. In such a case, $\mathfrak{g}_{\mathbb{C}} = \operatorname{End}(\mathbb{C}^n)$. As in §19, 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

$$H = \sum t_j e_j \Longrightarrow [H, e_{jk}] = i(t_j - t_k)e_{jk}. \tag{27.13}$$

Using this, we have that, when $G = \mathrm{U}(n)$, $h = \mathrm{diag}(e^{i\theta_1}, \dots, e^{i\theta_n}) \in T$,

$$\mathrm{Ad}\, h(e_{jk}) = e^{i(\theta_j - \theta_k)}\, e_{jk}. \tag{27.14}$$

Thus

$$\begin{aligned} \det(I - \mathrm{Ad}\, h)_{\mathfrak{g}/\mathfrak{h}} &= \det(I - \mathrm{Ad}\, h)_{\mathfrak{g}_{\mathbb{C}}/\mathfrak{h}_{\mathbb{C}}} \\ &= \prod_{j \ne k} (1 - e^{i(\theta_j - \theta_k)}) \\ &= \prod_{j \ne k} e^{-i\theta_k}(e^{i\theta_k} - e^{i\theta_j}), \end{aligned} \tag{27.15}$$

and hence

$$|\det(I - \mathrm{Ad}\, h)_{\mathfrak{g}/\mathfrak{h}}| = \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2. \tag{27.16}$$

We explicitly specialize (27.1) 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 27.2. *If $f : \mathrm{U}(n) \to \mathbb{C}$ is a central function, then*

$$\int\limits_{\mathrm{U}(n)} f(g)\, dg = \frac{1}{(2\pi)^n n!} \int\limits_{\mathbb{T}^n} f(D(\theta)) J(\theta)\, d\theta_1 \cdots d\theta_n, \tag{27.17}$$

where $D(\theta)$ is the diagonal matrix with diagonal entries $e^{i\theta_1}, \dots, e^{i\theta_n}$, and

$$J(\theta) = \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2. \tag{27.18}$$

Here we take $\mathbb{T}^n = (\mathbb{R}/2\pi\mathbb{Z})^n$. We mention another way of writing $J(\theta)$, namely as

$$J(\theta) = A(\theta)\overline{A(\theta)}, \quad A(\theta) = \prod_{j<k} \bigl(1 - e^{-i\omega_{jk}(\theta)}\bigr), \tag{27.19}$$

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

$$J(\theta) = V(\zeta)V(\overline{\zeta}), \quad V(\zeta) = \prod_{j<k} (\zeta_k - \zeta_j). \tag{27.20}$$

Now $V(\zeta)$ is a Vandermonde determinant:

$$V(\zeta) = \det \begin{pmatrix} 1 & \cdots & 1 \\ \zeta_1 & \cdots & \zeta_n \\ \vdots & & \vdots \\ \zeta_1^{n-1} & \cdots & \zeta_n^{n-1} \end{pmatrix}. \tag{27.21}$$

Hence

$$V(\zeta) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, \zeta_1^{\sigma(1)-1} \cdots \zeta_n^{\sigma(n)-1}. \tag{27.22}$$

Now $\overline{\zeta}_j = \zeta_j^{-1}$ for $\zeta_j \in S^1$, so

$$J(\theta) = \sum_{\sigma,\tau \in S_n} (\operatorname{sgn} \sigma)(\operatorname{sgn} \tau)\, \zeta_1^{\sigma(1)-\tau(1)} \cdots \zeta_n^{\sigma(n)-\tau(n)}. \tag{27.23}$$

Note that

$$\begin{aligned} (2\pi)^{-n} \int_{\mathbb{T}^n} J(\theta)\, d\theta &= \text{constant term in (27.23)} \\ &= \sum \big\{ (\operatorname{sgn} \sigma)(\operatorname{sgn} \tau) : \sigma = \tau \in S_n \big\} \\ &= n!, \end{aligned} \tag{27.24}$$

which gives a check on the coefficient on the right side of (27.17).

28. 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 §19 that the representation space V of $\mathcal{D}_\lambda$ has a decomposition

$$V = \bigoplus_\mu V_\mu \tag{28.1}$$

into spaces of weight vectors. It follows from (19.9) that

$$\chi_\lambda(D(\theta)) = \sum m_\mu \, e^{i\mu(\theta)}, \quad m_\mu = \dim V_\mu. \tag{28.2}$$

Our goal is to get a more explicit formula for this object. To be sure, (28.2) 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 21.2 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,

$$\sigma \cdot \mu = (\mu_{\sigma(1)}, \dots, \mu_{\sigma(n)}). \tag{28.3}$$

In view of (21.4), we have the identity

$$m_{\sigma\cdot\mu} = m_\mu, \quad \forall\, \sigma \in S_n. \tag{28.4}$$

Another identity arises by rewriting the identity

$$\int_{\mathrm{U}(n)} \chi_\lambda(g) \, \overline{\chi_\lambda(g)} \, dg = 1, \tag{28.5}$$

using the Weyl integration formula (27.17):

$$(2\pi)^{-n} \int_{\mathbb{T}^n} A(\theta)\chi_\lambda(D(\theta)) \, \overline{A(\theta)} \, \overline{\chi_\lambda(D(\theta))} \, d\theta = n! \tag{28.6}$$

To exploit this, we consider

$$\varphi(\theta) = A(\theta) \, \chi_\lambda(D(\theta)). \tag{28.17}$$

Recall from (27.19) that

$$A(\theta) = \prod_{j<k} \bigl(1 - e^{-i\omega_{jk}(\theta)}\bigr). \tag{28.18}$$

In particular, we have a finite sum

$$\varphi(\theta) = \sum c_\mu \, e^{i\mu(\theta)}, \quad c_\mu \in \mathbb{Z}, \tag{28.19}$$

and the identity (28.6) implies

$$\sum c_\mu^2 = n! \tag{28.20}$$

As we will see, this will help place strong constraints on the coefficients c_μ, particularly in concert with the following observation.

Lemma 28.1. *For the highest weight λ of $\mathcal{D}_\lambda$, we have*

$$c_\lambda = m_\lambda = 1. \tag{28.21}$$

Proof. Since the elements $\omega_{jk} \in \mathfrak{h}'$ are all positive for $j < k$ (with respect to the ordering defined in §19), it is clear from (28.17)–(28.18) that $c_\lambda = m_\lambda$. That $m_\lambda = 1$ follows from Proposition 19.4.

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

$$A(\sigma^t \cdot \theta) = \prod_{j<k} \bigl(1 - e^{-i\omega_{\sigma(j)\sigma(k)}(\theta)}\bigr). \tag{28.22}$$

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

$$1 - e^{-i\omega_{\sigma(j)\sigma(k)}(\theta)} = -e^{-i\omega_{\sigma(j)\sigma(k)}(\theta)} \bigl(1 - e^{-i\omega_{\sigma(k)\sigma(j)}(\theta)}\bigr). \tag{28.23}$$

Recombining the two products gives

$$A(\sigma^t \cdot \theta) = \alpha \, e^{-i\beta} \, A(\theta), \tag{28.24}$$

with

$$\alpha = \prod_{\{(j,k): j<k, \sigma(j)>\sigma(k)\}} (-1), \quad \beta = \sum_{\{(j,k): j<k, \sigma(j)>\sigma(k)\}} \omega_{\sigma(j)\sigma(k)}(\theta). \tag{28.25}$$

A calculation gives $\alpha = \operatorname{sgn}\sigma$. Also, if we set

$$\rho = \frac{1}{2}\sum_{j<k}\omega_{jk} \in \mathfrak{h}', \tag{28.26}$$

then

$$\beta = \frac{1}{2}\sum_{j<k}\big[\omega_{\sigma(j)\sigma(k)} - \omega_{jk}\big] = \sigma\cdot\rho - \rho. \tag{28.27}$$

Hence (28.24) becomes

$$A(\sigma^t\cdot\theta) = (\operatorname{sgn}\sigma)\, e^{i(\rho(\theta)-\sigma\cdot\rho(\theta))}\, A(\theta). \tag{28.28}$$

In view of the conjugation invariance of χ_λ and (21.2), this gives

$$\varphi(\sigma^t\cdot\theta) = (\operatorname{sgn}\sigma)\, e^{i(\rho(\theta)-\sigma\cdot\rho(\theta))}\, \varphi(\theta). \tag{28.29}$$

Equivalently, the coefficients c_μ in (28.19) satisfy

$$c_{\sigma\cdot\mu+\sigma\cdot\rho-\rho} = (\operatorname{sgn}\sigma)c_\mu. \tag{28.30}$$

We are led to consider a “shifted” action of S_n on $\mathfrak{h}'$:

$$\tilde\sigma(\mu) = \sigma\cdot\mu + \sigma\cdot\rho - \rho, \tag{28.31}$$

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 28.2. *The orbit of the highest weight λ under the action of S_n given by (28.31) has $n!$ elements.*

Proof. Consider $\sigma\in S_n$ such that $\tilde\sigma(\lambda)=\lambda$. This implies

$$\sigma\cdot\lambda + \sigma\cdot\rho = \lambda + \rho. \tag{28.32}$$

Now $\lambda = (\lambda_1,\dots,\lambda_n)$ with $\lambda_1\ge\cdots\ge\lambda_n$ and

$$\rho = (\rho_1,\dots,\rho_n) = \frac{1}{2}(n-1,n-3,\dots,3-n,1-n), \tag{28.33}$$

so $\rho_1>\cdots>\rho_n$, and hence $\lambda+\rho = (\xi_1,\dots,\xi_n)$ with $\xi_1>\cdots>\xi_n$. Thus (28.2) can hold only if σ is the identity element of S_n.

Thus taking $\mu = \lambda$ in (28.30) gives $n!$ coefficients c_μ that are equal to ± 1. In view of (28.20), these are all the nonzero coefficients in (28.19). We have established

$$A(\theta)\chi_\lambda(D(\theta)) = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, e^{i(\sigma\cdot\lambda(\theta)+\sigma\cdot\rho(\theta)-\rho(\theta))}. \tag{28.34}$$

In view of the formula (28.18) 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 (28.34) yields

$$A(\theta) = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, e^{i(\sigma\cdot\rho(\theta)-\rho(\theta))}, \tag{28.35}$$

which one might also try to derive directly from (28.18). This suggests writing the character formula in terms of

$$A_\mu(\theta) = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, e^{i\sigma\cdot\mu(\theta)}. \tag{28.36}$$

Proposition 28.3. *The irreducible representation of* U(n) *with highest weight* λ *has character satisfying*

$$\chi_\lambda(D(\theta)) = \frac{A_{\lambda+\rho}(\theta)}{A_\rho(\theta)}. \tag{28.37}$$

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 (28.37) 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 (28.35) by $e^{i\rho(\theta)}$, the product (28.18) for $A(\theta)$ yields

$$A_\rho(\theta) = \prod_{j<k} \bigl(e^{i\omega_{jk}(\theta)/2} - e^{-i\omega_{jk}(\theta)/2}\bigr). \tag{28.38}$$

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 (28.36) 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

$$A_\mu(t\rho) = A_\rho(t\mu) = \prod_{j<k} \bigl(e^{it\langle\omega_{jk},\mu\rangle/2} - e^{-it\langle\omega_{jk},\mu\rangle/2}\bigr). \tag{28.39}$$

Here we have used (28.38) 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 (28.39) we can prove the following dimension formula.

Proposition 28.4. *The irreducible representation of* $\mathrm{U}(n)$ *with highest weight* λ *acts on a space* $V(\lambda)$ *whose dimension is*

$$\dim V(\lambda) = d_\lambda = \prod_{j<k} \frac{\langle \omega_{jk}, \lambda + \rho \rangle}{\langle \omega_{jk}, \rho \rangle}. \tag{28.40}$$

Proof. Clearly $d_\lambda = \chi_\lambda(I)$, and hence

$$d_\lambda = \lim_{t \to 0} \chi_\lambda(D(t\rho)) = \lim_{t \to 0} \frac{A_{\lambda+\rho}(t\rho)}{A_\rho(t\rho)}, \tag{28.41}$$

by (28.37). Now we can apply (28.39) to obtain

$$d_\lambda = \lim_{t \to 0} \prod_{j<k} \frac{\sin t\langle \omega_{jk}, \lambda + \rho \rangle/2}{\sin t\langle \omega_{jk}, \rho \rangle/2}, \tag{28.42}$$

which yields (28.24), granted that $\prod_{j<k} \langle \omega_{jk}, \rho \rangle \neq 0$. In fact,

$$j < k \Longrightarrow \langle \omega_{jk}, \rho \rangle = \rho_j - \rho_k = k - j, \tag{28.43}$$

by (28.33), which leads to the following explicit formula for the denominator that arises in (28.40):

$$\prod_{j<k} \langle \omega_{jk}, \rho \rangle = \prod_{1 \le j < k \le n} (k - j) = \prod_{\ell=1}^{n-1} \ell!. \tag{28.44}$$

29. Examples of characters

Let us first specialize the character formula of §28 to the case of U(2). We consider $\chi_\lambda(D(\theta))$ with $\lambda = (\lambda_1, \lambda_2),\ \lambda_1 \geq \lambda_2,\ \lambda_\nu \in \mathbb{Z}$. In this case, $\omega_{12} = (1, -1),\ \rho = (1/2, -1/2)$, and hence $A_\mu(\theta)$, defined by (28.35), takes the form

$$A_\mu(\theta) = e^{i(\mu_1\theta_1+\mu_2\theta_2)} - e^{i(\mu_2\theta_1+\mu_1\theta_2)}. \tag{29.1}$$

In particular, the denominator $A_\rho(\theta)$ in (28.37) becomes

$$A_\rho(\theta) = e^{i(\theta_1-\theta_2)/2} - e^{i(\theta_2-\theta_1)/2} = 2i \sin \frac{\theta_1-\theta_2}{2}. \tag{29.2}$$

To evaluate the numerator in (28.37), we take $\mu = \lambda + \rho = (\lambda_1 + 1/2, \lambda_2 - 1/2)$ in (29.1). For simplicity, let us take

$$\lambda = (k, 0). \tag{29.3}$$

Then

$$A_{\lambda+\rho}(\theta) = e^{i(k+1/2)\theta_1 - i\theta_2/2} - e^{i(k+1/2)\theta_2 - i\theta_1/2}. \tag{29.4}$$

If we also take $\theta_2 = -\theta_1$, so $D(\theta) \in \mathrm{SU}(2)$, then

$$A_{\lambda+\rho}(\theta) = e^{i(k+1)\theta_1} - e^{-i(k+1)\theta_1}, \tag{29.5}$$

and hence the character formula (21.37) gives

$$\chi_{(k,0)}(D(\theta, -\theta)) = \frac{\sin(k+1/2)\theta}{\sin\theta}. \tag{29.6}$$

Taking the limit $\theta \to 0$ gives the dimension formula

$$d_{(k,0)} = k + 1, \tag{29.7}$$

familiar from our previous discussion of the representations of SU(2).

Note that it is a direct consequence of (18.9), (18.13), (18.19), and (18.20) that

$$\chi_{(k,0)}(D(\theta, -\theta)) = \sum_{j=-k}^{k} e^{ij\theta}, \tag{29.8}$$

which sums to the right side of (29.6).

Let us generalize these calculations to the representations $S^k = \mathcal{D}_{(k,0,\dots,0)}$ of $\mathrm{U}(n)$. Using (20.1)–(20.8), we see that

$$\operatorname{Tr} S^k(D(\theta)) = \sum_{|\alpha|=k} e^{i\langle\alpha,\theta\rangle}, \tag{29.9}$$

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 (28.37) then asserts that

$$A_\rho(\theta)\ \operatorname{Tr} S^k(D(\theta)) = \sum_{|\alpha|=k} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, e^{i\langle\sigma\cdot\rho+\alpha,\theta\rangle} \tag{29.10}$$

is equal to

$$A_{(k,0,\dots,0)+\rho}(\theta) = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, e^{i\langle\sigma\cdot\rho+\sigma\cdot(k,0,\dots,0),\theta\rangle}. \tag{29.11}$$

To check this for $n = 2$, note that (29.10) becomes

$$\sum_{\alpha_1+\alpha_2=k} \Bigl(e^{i(\alpha_1+1/2)\theta_1+i(\alpha_2-1/2)\theta_2} - e^{i(\alpha_1-1/2)\theta_1+i(\alpha_2+1/2)\theta_2}\Bigr). \tag{29.12}$$

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 (29.12) telescopes to the right side of (29.4), verifying identity of (29.10) and (29.11) when $n = 2$.

Note that if we reverse the order of summation in (29.10) and sum over $\mathcal{E} = \{(\sigma, \alpha) : \alpha = \sigma\cdot(k, 0, \dots, 0)\}$, we get (29.11). Also note that all the frequencies that arise in (29.10) 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 (28.20), an analogue of which also holds for $A_\rho(\theta)\chi_\lambda(\theta)$, that the rest of the sum in (29.10) vanishes, due to cancellations. The reader is invited to find a more direct demonstration of this vanishing.

Next consider the representation Λ^ℓ of $\mathrm{U}(n)$ on $\Lambda^\ell\mathbb{C}^n$, discussed in (20.9)–(20.14). 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

$$\operatorname{Tr}\Lambda^\ell D(\theta) = \sigma_\ell(e^{i\theta_1}, \dots, e^{i\theta_n}), \tag{29.13}$$

where σ_ℓ is the ℓth elementary symmetric polynomial. We can write this as

$$\begin{aligned} \operatorname{Tr}\Lambda^\ell(D(\theta)) &= \frac{1}{\ell!(n-\ell)!} \sum_{\sigma\in S_n} e^{i(\theta_{\sigma(1)}+\cdots+\theta_{\sigma(\ell)})} \\ &= \frac{1}{n!}\binom{n}{\ell} \sum_{\sigma\in S_n} e^{i\langle\sigma\cdot\gamma_\ell,\theta\rangle}, \end{aligned} \tag{29.14}$$

where $\gamma_\ell = (1, \ldots, 1, 0, \ldots, 0)$, with ℓ ones, is the highest weight of Λ^ℓ. Recall that $\Lambda^\ell = \mathcal{D}_{(1,\ldots,1,0,\ldots,0)}$. According to the character formula (28.37), the quantity

$$A_\rho(\theta)\ \mathrm{Tr}\,\Lambda^\ell(D(\theta)) = \frac{1}{n!}\binom{n}{\ell} \sum_{\sigma,\tau\in S_n} (\operatorname{sgn}\tau) e^{i\langle \tau\cdot\rho+\sigma\cdot\gamma_\ell,\theta\rangle} \tag{29.15}$$

is equal to

$$A_{\gamma_\ell+\rho}(\theta) = \sum_{\tau\in S_n} (\operatorname{sgn}\tau) e^{i\langle \tau\cdot(\rho+\gamma_\ell),\theta\rangle}. \tag{29.16}$$

Note that by taking $\sigma \mapsto \tau\sigma$, we can rewrite the right side of (29.15) as

$$\frac{1}{n!}\binom{n}{\ell} \sum_{\sigma,\tau\in S_n} (\operatorname{sgn}\tau) e^{i\langle \tau\cdot(\rho+\sigma\cdot\gamma_\ell),\theta\rangle}. \tag{29.17}$$

The sum in (29.17) over $\{(\sigma,\tau) : \sigma \text{ fixes } \gamma_\ell\}$ is equal to (29.16). An argument involving (29.20), similar to that made above comparing (29.10) and (29.11), can be used to show that all the other terms in (29.17) must cancel out. Again the reader is invited to find a direct demonstration of this cancellation.

30. 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 (26.7), i.e.,

$$\tau \cdot \otimes^k = \bigoplus_{\lambda \in F_{nk}} \mathcal{S}_\lambda \otimes \mathcal{D}_\lambda, \tag{30.1}$$

and its implication for characters,

$$\mathrm{Tr}(\tau(\sigma) \cdot \otimes^k g) = \sum_{\lambda \in F_{nk}} \chi_\lambda^S(\sigma)\, \chi_\lambda(g), \quad \sigma \in S_k,\ g \in \mathrm{U}(n). \tag{30.2}$$

Here $\chi_\lambda(g) = \mathrm{Tr}\, \mathcal{D}_\lambda$ is the character for which (28.37) 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 (30.2).

Lemma 30.1. *Suppose $\sigma \in S_k$ has cycles of length $\ell_1, \dots, \ell_r$ (so $\ell_1 + \cdots + \ell_r = k$). Then*

$$\mathrm{Tr}(\tau(\sigma) \cdot \otimes^k g) = \prod_{\nu=1}^{r} \mathrm{Tr}(g^{\ell_\nu}). \tag{30.3}$$

Proof. Since the left side of (30.3) 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 (30.3) is given by

$$\begin{aligned} \sum_{1 \le j_1, \dots, j_k \le n} & \langle \tau(\sigma) \cdot \otimes^k g(u_{j_1} \otimes \cdots \otimes u_{j_k}), u_{j_1} \otimes \cdots \otimes u_{j_k} \rangle \\ &= \sum \langle g\, u_{j_{\sigma(1)}}, u_{j_1} \rangle \cdots \langle g\, u_{j_{\sigma(k)}}, u_{j_k} \rangle \\ &= \sum \zeta_{j_1} \cdots \zeta_{j_k}\, \delta_{j_1 j_{\sigma(1)}} \cdots \delta_{j_k j_{\sigma(k)}}. \end{aligned} \tag{30.4}$$

Meanwhile the right side of (30.3) is equal to

$$\sum_{j_1, \dots, j_r} \zeta_{j_1}^{\ell_1} \cdots \zeta_{j_r}^{\ell_r}. \tag{30.5}$$

Now under our stated hypotheses on σ, the nonzero terms in the last sum in (30.4) 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 (30.4) does indeed coincide with (30.5).

We will denote the quantity (30.3) by $\Xi(\sigma, g)$, so (30.2) reads

$$\Xi(\sigma, g) = \sum_{\lambda \in F_{nk}} \chi_\lambda^S(\sigma)\chi_\lambda(g). \tag{30.6}$$

Given $\mu \in F_{nk}$, we can multiply both sides of (30.6) by $\overline{\chi}_\mu^S(\sigma)$ and average over $\sigma \in S_k$, obtaining (upon switching notation from μ to λ)

$$\chi_\lambda(g) = \frac{1}{k!} \sum_{\sigma \in S_k} \Xi(\sigma, g)\overline{\chi}_\lambda^S(\sigma). \tag{30.7}$$

Similarly,

$$\chi_\lambda^S(\sigma) = \int\limits_{\mathrm{U}(n)} \Xi(\sigma, g)\overline{\chi}_\lambda(g)\, dg. \tag{30.8}$$

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

$$P_\sigma(\zeta) = P_{\ell_1}(\zeta) \cdots P_{\ell_r}(\zeta) \tag{30.9}$$

if σ consists of cycles of length $\ell_1, \dots, \ell_r$ (so $\ell_1 + \cdots + \ell_r = k$). Then

$$\Xi(\sigma, g) = P_\sigma(\zeta) \tag{30.10}$$

provided the eigenvalues of g are $\zeta_1, \dots, \zeta_n$.

If we insert the character formula (28.37) into (30.8), 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 (30.8) as an integral over $\mathbb{T}^n$, using the Weyl formula (27.17). Note that, in place of $J(\theta) = A(\theta)\overline{A(\theta)}$ as in (27.19), we can write

$$J(\theta) = A_\rho(\theta)\overline{A_\rho(\theta)}, \tag{30.11}$$

with $A_\rho(\theta)$ the denominator in (28.37). Hence (30.8) yields

$$\chi_\lambda^S(\sigma) = \frac{1}{(2\pi)^n n!} \int\limits_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\, \overline{A_{\lambda+\rho}(\theta)}\, A_\rho(\theta)\, d\theta, \tag{30.12}$$

Next, we write

$$A_\rho(\theta) = e^{i\langle\rho,\theta\rangle}\Bigl(\prod_{j<k} e^{-i\theta_j}\Bigr)\Delta(\theta), \quad \Delta(\theta) = \prod_{j<k}(e^{i\theta_j} - e^{i\theta_k}), \tag{30.13}$$

and note that

$$\sum_{j<k}\theta_j = (n-1)\theta_1 + (n-2)\theta_2 + \cdots + \theta_{n-1} = \langle\Gamma,\theta\rangle, \tag{30.14}$$

with

$$\Gamma = (n-1, n-2, \dots, 1, 0). \tag{30.15}$$

Hence

$$\begin{aligned} A_\rho(\theta) &= e^{i\langle\rho-\Gamma,\theta\rangle}\,\Delta(\theta) \\ &= e^{i(n-1)(\theta_1+\cdots+\theta_n)/2}\,\Delta(\theta). \end{aligned} \tag{30.16}$$

We have (via (28.36) with $\mu = \lambda + \rho$)

$$\chi^S_\lambda(\sigma) = \frac{1}{n!}\sum_{\tau\in S_n}(\operatorname{sgn}\tau) I^\tau_\lambda(\sigma), \tag{30.17}$$

with

$$I^\tau_\lambda(\sigma) = \frac{1}{(2\pi)^n}\int_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\Delta(\theta)\, e^{i\langle\rho-\Gamma,\theta\rangle} e^{-i\langle\tau\cdot(\lambda+\rho),\theta\rangle}\,d\theta. \tag{30.18}$$

Now $\Xi(\sigma, D(\theta))$ and $e^{i\langle\rho-\Gamma,\theta\rangle}$ are symmetric in $(\theta_1,\dots,\theta_n)$ (the latter by (30.16)), while applying a permutation τ to θ multiplies $\Delta(\theta)$ by $\operatorname{sgn}\tau$. Hence we have

$$\begin{aligned} I^\tau_\lambda(\sigma) &= \frac{\operatorname{sgn}\tau}{(2\pi)^n}\int_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\Delta(\theta)\, e^{i\langle\rho-\Gamma,\theta\rangle} e^{-i\langle\lambda+\rho,\theta\rangle}\,d\theta \\ &= \frac{\operatorname{sgn}\tau}{(2\pi)^n}\int_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\Delta(\theta) e^{-i\langle\lambda+\Gamma,\theta\rangle}\,d\theta. \end{aligned} \tag{30.19}$$

Plugging this into (30.17), we have Schur's formula:

Proposition 30.2. *Given $\lambda \in F_{nk}$, the associated representation $\mathcal{S}_\lambda$ of S_k has character*

$$\chi^S_\lambda(\sigma) = (2\pi)^{-n}\int_{\mathbb{T}^n} \Xi(\sigma, D(\theta))\Delta(\theta) e^{-i\langle\lambda+\Gamma,\theta\rangle}\,d\theta. \tag{30.20}$$

Equivalently, $\chi^S_\lambda(\sigma)$ *is equal to the coefficient of* $\zeta_1^{\ell_1}\cdots\zeta_n^{\ell_n}$ *in*

$$P_\sigma(\zeta)\prod_{j<k}(\zeta_j-\zeta_k), \tag{30.21}$$

where

$$\ell_1=\lambda_1+n-1,\ \ell_2=\lambda_2+n-2,\dots,\ell_n=\lambda_n. \tag{30.22}$$

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 (30.9),

$$P_e(\zeta)=(\zeta_1+\cdots+\zeta_n)^k=\sum_{|\beta|=k}\frac{k!}{\beta_1!\cdots\beta_n!}\,\zeta_1^{\beta_1}\cdots\zeta_n^{\beta_n}. \tag{30.23}$$

Using the Vandermonde determinant, as in (27.22), we have

$$\prod_{i<j}(\zeta_i-\zeta_j)=(-1)^{n(n-1)/2}\sum_{\sigma\in S_n}(\operatorname{sgn}\sigma)\,\zeta_1^{\sigma(1)-1}\cdots\zeta_n^{\sigma(n)-1}. \tag{30.24}$$

After some computation, there results the dimension formula

$$d^S_\lambda=\frac{k!}{\ell_1!\cdots\ell_n!}\prod_{i<j}(\ell_i-\ell_j), \tag{30.25}$$

for the representation $\mathcal{S}_\lambda$ of S_k, with $\ell_1,\dots,\ell_n$ given by (30.22). For details, see [FH], pp. 49–50.

31. Integral of $|\operatorname{Tr} g^k|^2$ and variants

Integrals of the form

$$\mathcal{I}(\sigma_1, \sigma_2) = \int_{\mathrm{U}(n)} \Xi(\sigma_1, g)\overline{\Xi(\sigma_2, g)}\, dg \tag{31.1}$$

are of interest in random matrix theory (cf., e.g., [DS], [DE]). Here $\Xi(\sigma, g)$ is as in (30.10). Note that

$$\sigma \in S_k,\ \vartheta \in \mathbb{R} \Longrightarrow \Xi(\sigma, e^{i\vartheta} g) = e^{ik\vartheta}\Xi(\sigma, g), \tag{31.2}$$

and since $g \mapsto e^{i\vartheta} g$ is a measure preserving map on $\mathrm{U}(n)$ it easily follows that

$$\sigma_\nu \in S_{k_\nu},\ k_1 \neq k_2 \Longrightarrow \mathcal{I}(\sigma_1, \sigma_2) = 0. \tag{31.3}$$

on the other hand, if $\sigma_1, \sigma_2 \in S_k$, one can use (30.6) to write

$$\mathcal{I}(\sigma_1, \sigma_2) = \sum_{\lambda \in F_{nk}} \chi^S_\lambda(\sigma_1)\, \overline{\chi}^S_\lambda(\sigma_2). \tag{31.4}$$

Such an identity is applied to random matrix theory in [DE].

Cases of (31.4) where $\sigma_1 = \sigma_2 = \sigma$ are of particular interest. One example, which has already been mentioned in (26.12), arises from $\sigma = e$, the identity element of S_k:

$$\int_{\mathrm{U}(n)} |\operatorname{Tr} g|^{2k}\, dg = \sum_{\lambda \in F_{nk}} (f^\lambda)^2, \tag{31.5}$$

where f^λ is the dimension of the representation space of $\mathcal{S}_\lambda$. Another interesting example is

$$\int_{\mathrm{U}(n)} |\operatorname{Tr} g^k|^2\, dg = \sum_{\lambda \in F_{nk}} |\chi^S_\lambda(c_k)|^2, \quad c_k = (12\cdots k) \in S_k. \tag{31.6}$$

See [DE] for a direct evaluation of the right side of (31.6), using results on the symmetric group. Here we will make a direct calculation of the left side of (31.6), using the Weyl integration formula.

We have

$$I_{nk} = \int_{\mathrm{U}(n)} |\operatorname{Tr} g^k|^2\, dM = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} |e^{ik\theta_1} + \cdots + e^{ik\theta_n}|^2 J(\theta)\, d\theta. \tag{31.7}$$

We re-state this as follows. Set $\zeta_j = e^{i\theta_j}$, so

$$|e^{ik\theta_1} + \cdots + e^{ik\theta_n}|^2 = |\zeta_1^k + \cdots + \zeta_n^k|^2 = \sum_{\mu,\nu} \zeta_\mu^k \zeta_\nu^{-k}. \tag{31.8}$$

Also, we use (27.23) for $J(\theta)$.

Thus I_{nk} is equal to the constant term in

$$\frac{1}{n!} \sum_{\mu,\nu,\sigma,\tau} (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)\zeta_\mu^k \zeta_\nu^{-k} \zeta_1^{\sigma(1)-\tau(1)} \cdots \zeta_n^{\sigma(n)-\tau(n)}, \tag{31.9}$$

which we write as

$$\frac{1}{n!}(S_1 + S_2), \tag{31.10}$$

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

$$S_1 = n \cdot n!. \tag{31.11}$$

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 (31.9) 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

$$\begin{aligned} \sigma(\mu) - \tau(\mu) &= -k, \\ \sigma(\nu) - \tau(\nu) &= k. \end{aligned} \tag{31.12}$$

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

$$\psi(j_1) = j_1 + k, \quad \psi(j_2) = j_2 - k. \tag{31.13}$$

This requires $\psi(j_1) = j_2,\ \psi(j_2) = j_1$, with

$$j_1 = j_2 - k. \tag{31.14}$$

Then

$$S_2 = \sum (\operatorname{sgn}\sigma)(\operatorname{sgn}\psi\sigma), \tag{31.15}$$

the sum running over such allowable (μ, ν, σ, ψ). Note that (31.14) constrains j_1; we require $k+1 \le j_1 \le n$. Thus if $k \ge n$ the sum in (31.15) is empty and $S_2 = 0$. If $1 \le k \le n-1$, then there are $(n-k) \cdot n!$ terms in the sum (31.15). In fact, if we pick $\sigma \in S_n$ and then pick one of the $n-k$ permutations $\psi \in S_n$ for which (31.13) holds, then for each such (σ, ψ), the pair (μ, ν) is uniquely determined. Furthermore, each term in (31.15) is equal to $\operatorname{sgn} \psi = -1$. Hence

$$S_2 = -(n-k) \cdot n!, \quad 1 \leq k \leq n-1. \tag{31.16}$$

Putting together these computations, we have, for integers $k \geq 1$,

$$\int\limits_{\mathrm{U}(n)} |\operatorname{Tr} g^k|^2 \, dg = k \wedge n. \tag{31.17}$$

The formula (31.17) 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 (31.17), one can show that

$$\int\limits_{\mathrm{U}(n)} X_u(g) X_v(g) \, dg = \sum_{k=-\infty}^{\infty} a_{nk} \, \hat{u}(k) \hat{v}(-k), \tag{31.18}$$

where $\hat{u}(k)$ are the Fourier coefficients of u, $a_{n0} = n^2$, and $a_{nk} = (|k| \wedge n)$ for $k \neq 0$.

To compare the derivation (31.7)–(31.17) with a treatment via (31.6), note that, by Proposition 30.2, $\chi_\lambda^S(c_k)$ is the coefficient of $\zeta_1^{\ell_1} \cdots \zeta_n^{\ell_n}$ in

$$\begin{aligned} &(\zeta_1^k + \cdots + \zeta_n^k) \prod_{i<j} (\zeta_i - \zeta_j) \\ &= (-1)^{n(n-1)/2} \sum_{j=1}^{n} \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \zeta_j^k \zeta_1^{\sigma(1)-1} \cdots \zeta_n^{\sigma(n)-1}. \end{aligned} \tag{31.19}$$

Using this one can show that $\chi_\lambda^S(c_k)$ is either 0 or ± 1. Computing the right side of (31.6) then apparently involves calculations somewhat similar to those done in (31.7)–(31.17).

32. The Laplace operator on $\mathrm{U}(n)$

If $\{X_j\}$ is an orthonormal basis of the Lie algebra $\mathfrak{g} = \mathrm{u}(n)$, regarded as an algebra of left-invariant vector fields, then the Laplace operator on $\mathrm{U}(n)$ is the second order differential operator

$$\Delta = \sum X_j^2. \tag{32.1}$$

It is an exercise to show that Δ is independent of the choice of orthonormal basis. 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

$$x_{jk} = \frac{1}{\sqrt{2}}(e_{jk} - e_{kj}), \quad y_{jk} = -\frac{i}{\sqrt{2}}(e_{jk} + e_{kj}), \tag{32.2}$$

then we have an orthonormal basis of $\mathfrak{g}$, so

$$\Delta = \sum_j e_j^2 + \sum_{j<k}(x_{jk}^2 + y_{jk}^2). \tag{32.3}$$

Proposition 32.1. *For all $X \in \mathfrak{g}$, $[\Delta, X] = 0$.*

Proof. It suffices to show that, for all j,k,

$$[\Delta, e_{jk}] = 0, \tag{32.4}$$

This is a straightforward computation using (19.10), i.e.,

$$[e_{ij}, e_{k\ell}] = \delta_{jk}e_{i\ell} - \delta_{i\ell}e_{kj}. \tag{32.5}$$

Details are an exercise.

We produce some more useful formulas for Δ. Note that

$$x_{jk} + iy_{jk} = \sqrt{2}\, e_{jk}, \quad x_{jk} - iy_{jk} = -\sqrt{2}\, e_{kj}. \tag{32.5}$$

Also, using (32.5), we have

$$[x_{jk}, y_{jk}] = -(e_j - e_k). \tag{32.7}$$

Hence we can rewrite (32.3) as

$$\begin{aligned} \Delta &= \sum_j e_j^2 + \sum_{j<k}\big\{(x_{jk} - iy_{jk})(x_{jk} + iy_{jk}) - i[x_{jk}, y_{jk}]\big\} \\ &= \sum_j e_j^2 + i\sum_{j<k}(e_j - e_k) - 2\sum_{j<k} e_{kj}e_{jk}. \end{aligned} \tag{32.8}$$

The significance for representation theory is highlighted by the following:

Proposition 32.2. *For the irreducible representation $\mathcal{D}_\lambda$ of* U(n) *on* $V(\lambda)$, *we have*

$$d\mathcal{D}_\lambda(\Delta) = -(|\lambda+\rho|^2 - |\rho|^2)I. \tag{32.9}$$

Proof. It follows from Proposition 17.1 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

$$d\mathcal{D}_\lambda(\Delta)v = \sum_j d\mathcal{D}_\lambda(e_j)^2 v + i\sum_{j<k} d\mathcal{D}_\lambda(e_j - e_k)v. \tag{32.10}$$

Now $d\mathcal{D}_\lambda(e_j)v = i\lambda(e_j)v$, so

$$\begin{aligned} d\mathcal{D}_\lambda(\Delta)v &= -\sum_j \lambda(e_j)^2 v - \sum_{j<k} \lambda(e_j - e_k)v \\ &= -(|\lambda|^2 + 2\langle\lambda,\rho\rangle)v, \end{aligned} \tag{32.11}$$

which gives (32.8).

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

$$\mathcal{V}_\alpha = \text{span of } \{\pi^\alpha_{jk} : 1 \le j,k \le d_\alpha\}, \tag{32.12}$$

and let $\mathcal{P}_\alpha$ denote the orthogonal projection of $L^2(G)$ onto $\mathcal{V}_\alpha$, as in Proposition 11.5. We have

$$\mathcal{P}_\alpha u(x) = d_\alpha \sum_{j,k} \pi^\alpha_{jk}(x) \int_G u(y)\overline{\pi}^\alpha_{jk}(y)\,dy. \tag{32.13}$$

Writing $\sum \pi^\alpha_{jk}(x)\overline{\pi}^\alpha_{jk}(y) = \mathrm{Tr}(\pi^\alpha(x)\pi^\alpha(y)^*) = \chi_\alpha(xy^{-1})$, we see that

$$\mathcal{P}_\alpha u(x) = d_\alpha\, \chi_\alpha * u(x). \tag{32.14}$$

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

$$\Gamma_\alpha(g,h)u(x) = u(g^{-1}xh), \quad u \in \mathcal{V}_\alpha. \tag{32.15}$$

A brief calculation shows that

$$\Gamma_\alpha(g,h)\pi^\alpha_{jk}(x) = \sum_{\ell,m} \overline{\pi}^\alpha_{\ell j}(g)\pi^\alpha_{mk}(h)\pi^\alpha_{\ell m}(x). \tag{32.16}$$

It follows readily that

$$\operatorname{Tr}\Gamma_\alpha(g,h) = \overline{\chi}_\alpha(g)\chi_\alpha(h), \tag{32.17}$$

and in particular

$$\int\limits_{G\times G} |\operatorname{Tr}\Gamma_\alpha(g,h)|^2\,dg\,dh = \int\limits_G |\overline{\chi}_\alpha(g)|^2\,dg \int\limits_G |\chi_\alpha(h)|^2\,dh = 1, \tag{32.18}$$

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 $\mathcal{V}_\alpha$:

$$\Delta\pi^\alpha_{jk}(x) = c(\alpha)\,\pi^\alpha_{jk}(x). \tag{32.19}$$

The formula (32.1) 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 (32.16), we see that

$$d\pi^\alpha(\Delta) = c(\alpha)I. \tag{32.20}$$

We mention that (32.9) generalizes from $\mathrm{U}(n)$ to a general compact G; see, e.g., [T1], 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 (32.19) and (32.20), we see that (32.9) gives

$$\Delta\pi^\lambda_{jk}(x) = -(|\lambda+\rho|^2 - |\rho|^2)\pi^\lambda_{jk}(x), \quad 1, \le j,k \le d_\lambda, \tag{32.21}$$

if (π^λ_{jk}) denotes a matrix form of $\mathcal{D}_\lambda$.

33. 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 §32. We consider the heat equation for $u(t, x)$ on $\mathbb{R}^+ \times G$:

$$\frac{\partial u}{\partial t} = \Delta u, \quad u(0, x) = f(x). \tag{33.1}$$

As seen in §7, we can write

$$f(x) = \sum_{\alpha \in \mathcal{I}} d_\alpha \sum_{j,k} \hat{f}_{jk}(\alpha)\, \pi^\alpha_{jk}(x), \tag{33.2}$$

where $\{\pi^\alpha : \alpha \in \mathcal{I}\}$ is a complete set of irreducible unitary representations of G and

$$\hat{f}_{jk}(\alpha) = \overline{\pi}^\alpha_{jk}(f) = \int_G f(y) \overline{\pi}^\alpha_{jk}(y)\, dy. \tag{33.3}$$

In view of (32.19), we then have

$$u(t, x) = \sum_{\alpha \in \mathcal{I}} d_\alpha e^{c(\alpha)t} \sum_{j,k} \hat{f}_{jk}(\alpha) \pi^\alpha_{jk}(x). \tag{33.4}$$

We can write

$$u(t, x) = H_t * f(x), \tag{33.5}$$

with $H_t(x)$, known as the heat kernel, given as follows. By (10.5), (33.5) is equivalent to

$$\widehat{H}_t(\alpha) = e^{c(\alpha)t}\, I, \tag{33.6}$$

so, parallel to (33.2),

$$H_t(x) = \sum_{\alpha \in \mathcal{I}} d_\alpha e^{c(\alpha)t}\, \chi_\alpha(x). \tag{33.7}$$

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

$$H_t(x) = \sum_{\lambda \in \mathcal{P}_+} d_\lambda e^{-(|\lambda+\rho|^2 - |\rho|^2)t}\, \chi_\lambda(x). \tag{33.8}$$

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

$$H_t(D(\theta)) = \frac{e^{|\rho|^2 t}}{MA_\rho(\theta)} \sum_{\lambda\in\mathcal{P}_+} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\Delta(\lambda+\rho)e^{i\langle\sigma\cdot(\lambda+\rho),\theta\rangle}e^{-|\lambda+\rho|^2 t}. \tag{33.9}$$

Here $M = \prod_{\ell=1}^{n-1} \ell!$ is the denominator calculated in (28.44) and

$$\Delta(\lambda) = \prod_{j<k} \langle\omega_{jk},\lambda\rangle = \prod_{j<k} (\lambda_j - \lambda_k). \tag{33.10}$$

We can rewrite (33.9) as

$$H_t(D(\theta)) = \frac{e^{|\rho|^2 t}}{MA_\rho(\theta)} \sum_{\lambda\in\widetilde{\mathcal{P}}_+} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\Delta(\lambda)e^{i\langle\sigma\cdot\lambda,\theta\rangle}e^{-|\lambda|^2 t}, \tag{33.11}$$

where

$$\widetilde{\mathcal{P}}_+ = \{\lambda \in (\mathbb{Z}+\gamma)^n : \lambda_1 > \cdots > \lambda_n\}, \tag{33.12}$$

and where

$$\gamma = 0 \ \text{ for } \ n \ \text{ odd, } \ 1/2 \ \text{ for } \ n \ \text{ even.} \tag{33.13}$$

Note that

$$\Delta(\sigma\cdot\lambda) = (\operatorname{sgn}\sigma)\Delta(\lambda), \tag{33.14}$$

and that $\Delta(\lambda) = 0$ whenever $\lambda_\mu = \lambda_\nu$ for some $\mu \neq \nu$. Hence

$$H_t(D(\theta)) = \frac{e^{|\rho|^2 t}}{MA_\rho(\theta)} \sum_{\lambda\in(\mathbb{Z}+\gamma)^n} \Delta(\lambda)e^{i\langle\lambda,\theta\rangle}e^{-|\lambda|^2 t}. \tag{33.15}$$

Let us set

$$E_\gamma(t,\theta) = \sum_{\lambda\in(\mathbb{Z}+\gamma)^n} e^{i\langle\lambda,\theta\rangle}e^{-|\lambda|^2 t}. \tag{33.16}$$

Then (33.15) yields

$$H_t(D(\theta)) = \frac{e^{|\rho|^2 t}}{MA_\rho(\theta)}\, Q(D)E_\gamma(t,\theta), \tag{33.17}$$

where $Q(D)$ is the differential operator on functions of θ:

$$Q(D)v(\theta) = \prod_{j<k} \frac{1}{i}\Big(\frac{\partial}{\partial\theta_j} - \frac{\partial}{\partial\theta_k}\Big). \tag{33.18}$$

Note that $E_\gamma(t,\theta)$ satisfies the heat equation on $\mathbb{R}^+ \times \mathbb{R}^n$:

$$\frac{\partial E_\gamma}{\partial t} = \sum_{j=1}^n \frac{\partial^2 E_\gamma}{\partial \theta_j^2}. \tag{33.19}$$

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,

$$\begin{aligned} E_0(0,\theta) &= (2\pi)^n \sum_{\nu\in\mathbb{Z}^n} \delta_{2\pi\nu}(\theta), \\ E_{1/2}(0,\theta) &= (2\pi)^n \sum_{\nu\in\mathbb{Z}^n} (-1)^{\nu_1+\cdots+\nu_n}\, \delta_{2\pi\nu}(\theta). \end{aligned} \tag{33.20}$$

Hence, for $t > 0$,

$$E_0(t,\theta) = \Big(\frac{\pi}{t}\Big)^{n/2} \sum_{\nu\in\mathbb{Z}^n} e^{-|\theta-2\pi\nu|^2/4t}, \tag{33.21}$$

and

$$E_{1/2}(t,\theta) = \Big(\frac{\pi}{t}\Big)^{n/2} \sum_{\nu\in\mathbb{Z}^n} (-1)^{\nu_1+\cdots+\nu_n}\, e^{-|\theta-2\pi\nu|^2/4t}. \tag{33.22}$$

Both functions have the same asymptotic behavior for $|\theta| \le \pi$ as $t \searrow 0$:

$$E_\gamma(t,\theta) \sim \Big(\frac{\pi}{t}\Big)^{n/2} e^{-|\theta|^2/4t}. \tag{33.23}$$

Regarding the heat kernel $H_t(x)$, we also have

$$H_t(x) \sim (4\pi t)^{-n^2/2}\, e^{-d(x)^2/4t}, \quad t \searrow 0, \tag{33.24}$$

where $d(x)$ denotes the distance from x to the identity element in the Riemannian metric on $\mathrm{U}(n)$. In particular, for $X \in \mathrm{u}(n)$, $|X| \le \pi/2$,

$$H_t(e^X) \sim (4\pi t)^{-n^2/2}\, e^{-|X|^2/4t}, \quad t \searrow 0. \tag{33.25}$$

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 [T2], Chapter 7 for a proof, and Chapter 10 for important geometrical applications. The reader might try to obtain (33.25) from (33.17) and (33.23).

34. The Harish-Chandra/Itzykson-Zuber integral

The integral

$$\int_{\mathrm{U}(n)} e^{s\,\mathrm{Tr}(gXg^{-1}Y)}\,dg = \mathcal{H}(s,X,Y), \quad X,Y\in \mathrm{M}(n,\mathbb{C}), \tag{34.1}$$

is of great interest in random matrix theory. Here we give several formulas for this, with arguments adapted from [AI] and [IZ]. 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

$$\begin{aligned}\mathcal{H}(s,x,y) &= \sum_{k=0}^{\infty}\frac{s^k}{k!}\int_{\mathrm{U}(n)} \bigl(\mathrm{Tr}(gxg^{-1}y)\bigr)^k\,dg \\ &= \sum_{k=0}^{\infty}\frac{s^k}{k!}\int_{\mathrm{U}(n)} \mathrm{Tr}\otimes^k(gxg^{-1}y)\,dg.\end{aligned} \tag{34.2}$$

To proceed, we use the following.

Lemma 34.1. *Let π,π' be irreducible representations of a compact Lie group G, with characters $\chi_\pi,\chi_{\pi'}$. Then*

$$\int_G \chi_\pi(xy)\overline{\chi}_{\pi'}(y)\,dy = d_\pi^{-1}\chi_\pi(x)\,\delta_{\pi\pi'}, \tag{34.3}$$

where $\delta_{\pi\pi'}=1$ if $\pi\approx\pi'$, 0 otherwise. Furthermore,

$$\int_G \chi_\pi(gxg^{-1}y)\,dg = d_\pi^{-1}\chi_\pi(x)\chi_\pi(y). \tag{34.4}$$

Proof. If we write the integrand in the left side of (34.3) as $\mathrm{Tr}(\pi(x)\pi(y))\overline{\chi}_{\pi'}(y)$ and apply Proposition 8.2 to $\int_G \pi(y)\overline{\chi}_{\pi'}(y)\,dy$, we get the identity (34.3). As for (34.4), one easily shows the left side is invariant under $y\mapsto h^{-1}yh,\ h\in G$. Hence

$$\int_G \chi_\pi(gxg^{-1}y)\,dg = \sum_{\alpha\in\mathcal{I}}\psi_\alpha(x)\,\chi_\alpha(y), \tag{34.5}$$

with

$$\psi_\alpha(x) = \iint_{G\,G} \chi_\pi(gxg^{-1}y)\overline{\chi}_\alpha(y)\,dy\,dg = d_\pi^{-1}\,\delta_{\pi\pi_\alpha}, \tag{34.6}$$

the last identity by (34.3).

To apply (34.4) to (34.2), we break up $\otimes^k$ into irreducibles, using Proposition 26.3. We get

$$\int_{\mathrm{U}(n)} \operatorname{Tr} \otimes^k (gxg^{-1}y)\, dy = \sum_{\lambda \in F_{nk}} \frac{f^\lambda}{d_\lambda}\, \chi_\lambda(x)\chi_\lambda(y), \tag{34.7}$$

where f^λ is the dimension of the representation space for $\mathcal{S}_\lambda$. Hence

$$\mathcal{H}(s,x,y) = \sum_{k=0}^{\infty} \frac{s^k}{k!} \sum_{\lambda \in F_{nk}} \frac{f^\lambda}{d_\lambda}\, \chi_\lambda(x)\chi_\lambda(y). \tag{34.8}$$

For a second approach, we apply (34.4) to each term in the series (33.8) for the heat kernel $H_t(x)$, obtaining

$$\int_{\mathrm{U}(n)} H_t(gxg^{-1}y)\, dg = \sum_{\lambda \in \mathcal{P}_+} e^{-(|\lambda+\rho|^2 - |\rho|^2)t}\, \chi_\lambda(x)\chi_\lambda(y). \tag{34.9}$$

We analyze this in a fashion parallel to (33.9)–(33.15). Denoting the quantity (34.9) by $K_t(x,y)$, we have

$$\begin{aligned} &K_t(D(\theta), D(\varphi)) \\ &= \frac{e^{|\rho|^2 t}}{A_\rho(\theta)A_\rho(\varphi)} \sum_{\lambda \in \mathcal{P}_+} \sum_{\sigma,\tau} (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau) e^{i\langle \sigma\cdot(\lambda+\rho),\theta\rangle} e^{i\langle \tau\cdot(\lambda+\rho),\varphi\rangle} e^{-|\lambda+\rho|^2 t} \\ &= \frac{e^{|\rho|^2 t}}{A_\rho(\theta)A_\rho(\varphi)} \sum_{\lambda \in \widetilde{\mathcal{P}}_+} \sum_{\sigma,\tau} (\operatorname{sgn}\tau) e^{i\langle \sigma\cdot\lambda,\theta+\tau^t\varphi\rangle} e^{-|\lambda|^2 t}, \end{aligned} \tag{34.10}$$

where we take $\tau \mapsto \tau\sigma$ to produce the last identity, and we define $\widetilde{\mathcal{P}}_+$ as in (33.16). To proceed further, we note that

$$B_\lambda(\theta + \tau^t\varphi) = \sum_{\sigma,\tau} (\operatorname{sgn}\tau) e^{i\langle \sigma\cdot\lambda,\theta+\tau^t\varphi\rangle} = A_\lambda(\theta)A_\lambda(\varphi), \tag{34.11}$$

and

$$A_{\sigma\cdot\lambda}(\theta) = (\operatorname{sgn}\sigma)A_\lambda(\theta). \tag{34.12}$$

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 (34.10) as

$$K_t(D(\theta), D(\varphi)) = \frac{e^{|\rho|^2 t}}{A_\rho(\theta)A_\rho(\varphi)} \sum_{\lambda \in (Z+\gamma)^n} \sum_{\tau \in S_n} (\operatorname{sgn}\tau) e^{i\langle \lambda,\theta+\tau^t\varphi\rangle} e^{-|\lambda|^2 t}. \tag{34.13}$$

Equivalently, with $E_\gamma(t,\theta)$ as in (33.16),

$$K_t(D(\theta),D(\varphi)) = \frac{e^{|\rho|^2 t}}{A_\rho(\theta)A_\rho(\varphi)} \sum_{\tau\in S_n} (\operatorname{sgn}\tau) E_\gamma(t,\theta+\tau^t\varphi). \tag{34.14}$$

The relevance of (34.14) to a calculation of (34.1) arises from the heat kernel asymptotics

$$H_t(g) \sim (4\pi t)^{-n^2/2}\, e^{-d(g)^2/4t}, \quad t \searrow 0, \tag{34.15}$$

discussed in §33. (Actually, an extra factor A_n enters, which we will discuss at the end of this section.) Let us take

$$x = e^{\eta X}, \quad y = e^{\eta Y}, \quad X,Y \in \mathrm{u}(n), \tag{34.16}$$

with $|\eta|$ small. Then

$$gxg^{-1}y = e^{\eta\,\mathrm{Ad}(g)X} e^{\eta Y} = e^{\eta(\mathrm{Ad}(g)X+Y)} + O(\eta^2). \tag{34.17}$$

We take

$$\eta = 2\sqrt{t}. \tag{34.18}$$

Then

$$\begin{aligned} H_t(gxg^{-1}y) &\sim (4\pi t)^{-n^2/2}\, e^{-\eta^2|\mathrm{Ad}(g)X+Y|^2/4t} \\ &= (4\pi t)^{-n^2/2}\, e^{-(|X|^2+|Y|^2+2\langle \mathrm{Ad}(g)X,Y\rangle)}. \end{aligned} \tag{34.19}$$

Hence, with (34.16) and (34.18) in effect,

$$\begin{aligned} K_t(x,y) &\sim (4\pi t)^{-n^2/2} e^{-(|X|^2+|Y|^2)} \int_{\mathrm{U}(n)} e^{-2\langle \mathrm{Ad}(g)X,Y\rangle}\, dg \\ &= (4\pi t)^{-n^2/2} e^{-(|X|^2+|Y|^2)} \mathcal{H}(2,X,Y), \end{aligned} \tag{34.20}$$

since $\langle X,Y\rangle = -\operatorname{Tr}(XY)$ in this case, so, for $X,Y\in \mathrm{u}(n)$,

$$\mathcal{H}(2,X,Y) = e^{|X|^2+|Y|^2} \lim_{t\searrow 0} (4\pi t)^{n^2/2}\, K_t(e^{2\sqrt{t}X}, e^{2\sqrt{t}Y}). \tag{34.21}$$

Say

$$e^{\eta X} \sim D(\eta\theta), \quad e^{\eta Y} \sim D(\eta\varphi), \tag{34.22}$$

i.e., these matrices are similar. Then (34.14) gives

$$
\begin{aligned}
&K_t(e^{2\sqrt{t}X}, e^{2\sqrt{t}Y}) \\
&= \frac{e^{|\rho|^2 t}}{A_\rho(2\sqrt{t}\theta)A_\rho(2\sqrt{t}\varphi)} \sum_{\tau\in S_n} (\operatorname{sgn}\tau) E_\gamma(t, 2\sqrt{t}(\theta + \tau^t\varphi)).
\end{aligned} \tag{34.23}
$$

Now, as $t \searrow 0$,

$$
A_\rho(2\sqrt{t}\theta) \sim (2i\sqrt{t})^{n(n-1)/2}\, \Delta(\theta), \tag{34.24}
$$

and, by (33.22),

$$
E_\gamma(t, 2\sqrt{t}\theta) \sim \Big(\frac{\pi}{t}\Big)^{n/2} e^{-|\theta|^2}. \tag{34.25}
$$

Thus, as $t \searrow 0$,

$$
K_t(e^{2\sqrt{t}X}, e^{2\sqrt{t}Y}) \sim t^{-n/2} \frac{\pi^{n/2}}{(2i)^{n(n-1)}} \frac{1}{\Delta(\theta)\Delta(\varphi)} \sum_{\tau\in S_n} (\operatorname{sgn}\tau) e^{-|\theta+\tau^t\varphi|^2}. \tag{34.26}
$$

Hence, by (34.21),

$$
\mathcal{H}(2, X, Y) = \frac{C_n}{\Delta(\theta)\Delta(\varphi)} \sum_{\tau\in S_n} (\operatorname{sgn}\tau) e^{-2\langle\tau\cdot\theta,\varphi\rangle}, \tag{34.27}
$$

where X and Y are related to θ and φ by (34.22). Since clearly $\mathcal{H}(s, X, Y) = \mathcal{H}(1, sX, Y) = \mathcal{H}(2, (s/2)X, Y)$, we have

$$
\mathcal{H}(s, X, Y) = s^{-n(n-1)/2} \frac{C_n'}{\Delta(\theta)\Delta(\varphi)} \sum_{\tau\in S_n} (\operatorname{sgn}\tau) e^{-s\langle\tau\cdot\theta,\varphi\rangle}, \quad X, Y \in \mathrm{u}(n). \tag{34.28}
$$

Our renaming of the constants C_n, C_n' derives from the fact that (34.15) holds when $\mathrm{U}(n)$ has the Riemannian metric in which the norm on $T_I\,\mathrm{U}(n) = \mathrm{u}(n)$ 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 (34.28) is to consider the $s \to 0$ limit, using $\mathcal{H}(0, X, Y) = 1$. In fact,

$$
C_n' = \prod_{\ell=1}^{n-1} \ell!. \tag{34.29}
$$

We mention that if instead of X and Y being skew-adjoint, as in (34.28), we take X and Y self-adjoint, with eigenvalues θ_j and φ_j, respectively, then (34.28) holds with $-s$ changed to s in the exponent. This is the form in which the identity commonly appears. See [EM] for a recent application of such an identity.

35. Roots and weights for general compact Lie groups

The notions of roots and weights, described for $\mathrm{U}(n)$ in §19, 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 $g \in G$, $\mathrm{Ad}\, g$ is an orthogonal transformation on $\mathfrak{g}$, and hence for each $X \in \mathfrak{g}$, $\mathrm{ad}\, X$ 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 subgeoup; 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 $\mathrm{ad}\, h_j$ 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 $\mathrm{ad}\, h_j$ and it suffices to put $\mathrm{ad}\, h^b$ 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

$$\mathrm{ad}\, h(x_j \pm i y_j) = \pm i \alpha_j(h)(x_j \pm i y_j), \quad \forall\, h \in \mathfrak{h}, \tag{35.1}$$

for certain $\alpha_j \in \mathfrak{h}'$. Hence we can decompose the complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$ as

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus \bigoplus_{\alpha} \mathfrak{g}_\alpha, \tag{35.2}$$

where, given $\alpha \in \mathfrak{h}'$,

$$\mathfrak{g}_\alpha = \{z \in \mathfrak{g}_{\mathbb{C}} : [h, z] = i\alpha(h) z,\ \forall\, h \in \mathfrak{h}\}. \tag{35.3}$$

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

$$\mathrm{ad}\, h([z_\alpha, z_\beta]) = [\mathrm{ad}\, h(z_\alpha), z_\beta] + [z_\alpha, \mathrm{ad}\, h(z_\beta)], \tag{35.4}$$

it follows that

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}. \tag{35.5}$$

Note from (35.1) 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 §19, 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

$$H_\alpha \in \mathfrak{h}, \quad \alpha(h) = \langle H_\alpha, h \rangle, \quad \forall\, h \in \mathfrak{h}. \tag{35.6}$$

Here $\langle \ , \ \rangle$ is the Ad-invariant inner product on $\mathfrak{h}$ mentioned above (restricted in (35.6) 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 35.1. *If $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, then*

$$[X, Y] = i\langle X, Y \rangle H_\alpha. \tag{35.7}$$

Proof. By (35.5), $[X, Y] \in \mathfrak{h}_{\mathbb{C}} = \mathfrak{g}_0$. Now, for any $H \in \mathfrak{h}$,

$$\langle [X, Y], H \rangle = \langle Y, [H, X] \rangle = i\alpha(H)\langle Y, X \rangle, \tag{35.8}$$

and, by (35.6), this identity is equivalent to (35.7).

Note also that

$$\alpha(H_\alpha) = \langle H_\alpha, H_\alpha \rangle > 0. \tag{35.9}$$

We are ready for the following key result.

Proposition 35.2. *For each root α,*

$$dim\, \mathfrak{g}_\alpha = 1. \tag{35.10}$$

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

$$\langle X, Y \rangle = 1, \quad \langle Z, Y \rangle = 0. \tag{35.11}$$

This implies

$$[X, Y] = iH_\alpha, \quad [Z, Y] = 0. \tag{35.12}$$

From here an inductive argument, which we leave as an exercise, shows

$$\operatorname{ad} Y (\operatorname{ad} X)^n Z = -\frac{n(n+1)}{2}\, \alpha(H_\alpha) (\operatorname{ad} X)^{n-1} Z. \tag{35.13}$$

By (35.9), it follows that all the elements $(\operatorname{ad} X)^n Z \in \mathfrak{g}_{(n+1)\alpha}$ are nonzero. This contradicts the fact that $\mathfrak{g}$ is finite dimensional.

It remains to show that $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are not orthogonal with respect to $\langle\ ,\ \rangle$. Indeed,

$$
\begin{aligned}
x+iy \in \mathfrak{g}_\alpha\ (x,y\in\mathfrak{g}) &\Longrightarrow x-iy \in \mathfrak{g}_{-\alpha} \\
&\Longrightarrow \langle x+iy, x-iy\rangle = \langle x,x\rangle + \langle y,y\rangle,
\end{aligned}
\tag{35.14}
$$

which is > 0 as long as $x+iy \neq 0$. This completes the proof of Proposition 35.2.

Hence we can pick nonzero vectors $e_\alpha \in \mathfrak{g}_\alpha$, and arrange that $e_{-\alpha}$ be the complex conjugate of e_α,

$$
e_{\pm\alpha} = x_\alpha \pm i y_\alpha, \quad x_\alpha, y_\alpha \in \mathfrak{g}. \tag{35.15}
$$

Furthermore, we can scale these elements so that $\langle e_\alpha, e_{-\alpha}\rangle = 1$. Thus

$$
[e_\alpha, e_{-\alpha}] = iH_\alpha. \tag{35.16}
$$

These commutation relations, together with $[H_\alpha, e_{\pm\alpha}] = \pm i\alpha(H_\alpha)e_{\pm\alpha}$ are equivalent to

$$
[x_\alpha, y_\alpha] = -\frac{1}{2}H_\alpha, \quad [H_\alpha, x_\alpha] = -\alpha(H_\alpha)y_\alpha, \quad [H_\alpha, y_\alpha] = \alpha(H_\alpha)x_\alpha. \tag{35.17}
$$

The following result bears on the size of the linear span of the set of roots α in $\mathfrak{h}'$.

Proposition 35.3. *We have*

$$
\bigcap_\alpha \ker\alpha = \mathfrak{z}, \tag{35.18}
$$

the center of $\mathfrak{g}$.

Proof. Given $h \in \mathfrak{h}$,

$$
\begin{aligned}
\alpha(h) = 0\ \forall\,\alpha &\Longleftrightarrow \operatorname{ad} h = 0 \ \text{ on each } \mathfrak{g}_\alpha \\
&\Longleftrightarrow h \in \mathfrak{z}.
\end{aligned}
\tag{35.19}
$$

Of course, $\mathfrak{z} \subset \mathfrak{h}$, so this gives (35.18).

Corollary 35.4. *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 $\mathfrak{g} = \mathrm{u}(n)$, then $\mathfrak{z} = \{iaI : a \in \mathbb{R}\}$. If $\mathfrak{g} = \mathrm{su}(n)$, then $\mathfrak{z} = 0$. If $\mathfrak{g} = \mathrm{so}(n)$, 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 §19, we will find it convenient to bring in the complexification $G_{\mathbb{C}}$ of G, and use the following fact:

Proposition 35.5. *The representation π of G on V extends to a holomorphic representation of $G_{\mathbb{C}}$ on V.*

See Appendix H for a description of $G_{\mathbb{C}}$ and a proof of Proposition 35.5.

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

$$V_\lambda = \{v \in V : d\pi(h) = i\lambda(h)v,\ \forall\, h \in \mathfrak{h}\}. \tag{35.20}$$

We have

$$V = \bigoplus_\lambda V_\lambda. \tag{35.21}$$

If $V_\lambda \neq 0$ we call λ a *weight*, and any nonzero $v \in V_\lambda$ a *weight vector.* The decomposition (35.20) is called the weight space decomposition of V.

For the root vectors e_α considered above, set

$$E_\alpha = d\pi(e_\alpha). \tag{35.22}$$

We call E_α a *raising operator* if $\alpha > 0$ and a lowering operator if $\alpha < 0$. The commutation relations

$$[h, e_\alpha] = i\alpha(h)e_\alpha, \quad \forall\, h \in \mathfrak{h} \tag{35.23}$$

imply

$$d\pi(h)E_\alpha = E_\alpha d\pi(h) + i\alpha(h)E_\alpha. \tag{35.24}$$

Using this we can prove the following.

Proposition 35.6. *For each root α, we have*

$$E_\alpha : V_\lambda \longrightarrow V_{\lambda+\alpha}. \tag{35.25}$$

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 $h \in \mathfrak{h}$,

$$\begin{aligned} d\pi(h)(E_\alpha\xi) &= E_\alpha d\pi(h)\xi + i\alpha(h)E_\alpha\xi \\ &= i\big(\lambda(h) + \alpha(h)\big)E_\alpha\xi, \end{aligned} \tag{35.26}$$

which proves the proposition.

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 thate will be a *highest weight* λ_m, and also a lowest weight λ_s. From Proposition 35.6 we see that

$$\begin{aligned} E_\alpha &= 0 \text{ on } V_{\lambda_m}, \text{ for all raising operators } E_\alpha, \\ E_\alpha &= 0 \text{ on } V_{\lambda_s}, \text{ for all lowering operators } E_\alpha. \end{aligned} \tag{35.27}$$

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 35.7. *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 35.8. *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 35.5 produces two linearly independent weight vectors $\xi_j \in V_j$, annihilated by all raising operators.

We note the following. Set

$$\mathcal{H}(\pi) = \bigcap_{\alpha>0} \operatorname{Ker} E_\alpha. \tag{35.28}$$

From (35.24) it follows that

$$h \in \mathfrak{h} \Longrightarrow d\pi(h) : \mathcal{H}(\pi) \to \mathcal{H}(\pi), \tag{35.29}$$

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 35.8 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 35.8, parallel to Propositions 19.4–19.5. Let $\overline{\pi}$ denote the representation of G on V' contragredient to π, given by

$$\langle \xi, \overline{\pi}(g)\eta\rangle = \langle \pi(g^{-1})\xi, \eta\rangle, \quad \xi \in V, \ \eta \in V'. \tag{35.30}$$

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 $-\mu \in \mathfrak{h}'$. Is in §19, we will study the function

$$\varphi(g) = \langle \pi(g)\xi_0, \eta_0\rangle. \tag{35.31}$$

As stated in Proposition 35.5, 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

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus \mathfrak{n}_+ \oplus \mathfrak{n}_-, \quad \mathfrak{n}_+ = \bigoplus_{\alpha>0} \mathfrak{g}_\alpha, \quad \mathfrak{n}_- = \bigoplus_{\alpha<0} \mathfrak{g}_\alpha. \tag{35.32}$$

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

$$N_- D N_+ = G_{\mathrm{reg}} \tag{35.33}$$

is a subset of $G_{\mathbb{C}}$ that contains an open neighborhood of the identity element e. Let

$$g = \zeta\delta z, \quad \zeta \in N_-,\ z \in N_+, \quad \delta = \exp(h) = \exp(h_1 + ih_2) \in D, \tag{35.34}$$

with $h_j \in \mathfrak{h}$. We see that

$$\begin{aligned} \pi(z)\xi_0 = \xi_0, \quad \pi(\delta)\xi_0 = e^{i\lambda(h)}\xi_0, \\ \overline{\pi}(\zeta)\eta_0 = \eta_0, \quad \overline{\pi}(\delta^{-1})\eta_0 = e^{i\mu(h)}\eta_0. \end{aligned} \tag{35.35}$$

Consequently,

$$\varphi(\zeta g) = \varphi(g), \quad \varphi(\delta g) = e^{i(\mu(h_1)+i\mu(h_2))}\varphi(g), \tag{35.36}$$

and

$$\varphi(gz) = \varphi(g), \quad \varphi(g\delta) = e^{i(\lambda(h_1)+i\lambda(h_2))}\varphi(g). \tag{35.37}$$

Hence

$$\begin{aligned} \varphi(\zeta\delta z) = \varphi(\delta) &= e^{i(\lambda(h_1)+i\lambda(h_2))}\varphi(e) \\ &= e^{i(\mu(h_1)+i\mu(h_2))}\varphi(e). \end{aligned} \tag{35.38}$$

This identity is very significant, in light of the following result.

Lemma 35.9. *Assume π is irreducible. Then the function φ has the property*

$$\varphi(e) \neq 0. \tag{35.39}$$

Proof. If $\varphi(e) = 0$, then (35.38) 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. But if we set

$$V_0 = \{\xi \in V : \langle \pi(g)\xi, \eta_0\rangle = 0\ \forall\, g \in G\}, \tag{35.40}$$

then V_0 is invariant, and since clearly $V_0 \neq V$, we have $V_0 = 0$ if π is irreducible. This proves the lemma.

From (35.38) and the lemma, we can deduce the following important result.

Theorem 35.10. *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 $\lambda = \mu$ 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:

$$\varphi \text{ is holomorphic on } G_{\mathbb{C}}, \tag{35.41}$$

$$\varphi(\zeta g z) = \varphi(g), \quad \forall\, \zeta \in N_-,\ z \in N_+,\ g \in G, \tag{35.42}$$

$$\varphi(\delta) = e^{i(\lambda(h_1)+i\lambda(h_2))}, \quad \forall\, \delta = \exp(h_1 + ih_2) \in D. \tag{35.43}$$

Thus, if ξ_1 were another highest weight vector, also normalized so $\langle \xi_1, \eta_0 \rangle = 1$, we would have

$$\langle \pi(g)\xi_1, \eta_0 \rangle = \varphi(g), \quad \forall\, g \in G, \tag{35.44}$$

so $\langle \pi(g)(\xi_1 - \xi_0), \eta_0 \rangle = 0$ for all g, or equivalently

$$\langle \xi_1 - \xi_0, \overline{\pi}(g)\eta_0 \rangle = 0, \quad \forall\, g \in G. \tag{35.45}$$

Since $\overline{\pi}$ is irreducible, this implies $\xi_1 = \xi_0$.

As for the final assertion of Theorem 35.10, 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

$$\varphi_2(g) = \langle \pi_2(g)\xi_2, \eta_2 \rangle. \tag{35.46}$$

Then φ_2 also satisfies the conditions (35.41)–(35.43). 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 $L^2(G)$.

36. Roots and weights for compact G, II: injections $\mathrm{su}(2) \hookrightarrow \mathfrak{g}$

Recall from (35.15)–(35.17) the construction of $e_{\pm\alpha} = x_\alpha \pm i y_\alpha$, spanning $\mathfrak{g}_{\pm\alpha}$, satisfying

$$[x_\alpha, y_\alpha] = -\frac{1}{2} H_\alpha, \quad [H_\alpha, x_\alpha] = -\alpha(H_\alpha) y_\alpha, \quad [H_\alpha, y_\alpha] = \alpha(H_\alpha) x_\alpha, \tag{36.1}$$

with $H_\alpha \in \mathfrak{h}$ given by (35.6). 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

$$\alpha(H_\alpha) = \langle \alpha, \alpha \rangle, \tag{36.2}$$

and more generally $\lambda(H_\alpha) = \langle \lambda, \alpha \rangle$ for each $\lambda \in \mathfrak{h}'$. Let us set

$$X_1^\alpha = \frac{1}{\langle \alpha, \alpha \rangle} H_\alpha, \quad X_2^\alpha = \sqrt{\frac{2}{\langle \alpha, \alpha \rangle}}\, y_\alpha, \quad X_3^\alpha = \sqrt{\frac{2}{\langle \alpha, \alpha \rangle}}\, x_\alpha. \tag{36.3}$$

Then the commutation relations (36.1) are equivalent to

$$[X_1^\alpha, X_2^\alpha] = X_3^\alpha, \quad [X_2^\alpha, X_3^\alpha] = X_1^\alpha, \quad [X_3^\alpha, X_1^\alpha] = X_2^\alpha. \tag{36.4}$$

Now these commutation relations are identical to those in (18.2). Hence each root α gives rise to an injective Lie algebra homomorphism $\mathrm{su}(2) \hookrightarrow \mathfrak{g}$, which in turn, since $SU(2)$ is simply connected, exponentiates to a Lie group homomorphism

$$\gamma^\alpha : SU(2) \longrightarrow G, \tag{36.5}$$

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 (36.5) 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 §18 applies. Suppose λ is a weight of π, with weight space $V_\lambda \subset V$. Then

$$\begin{aligned} v \in V_\lambda &\Longrightarrow \pi(e^{tX_1^\alpha}) v = e^{it\lambda(X_1^\alpha)} v \\ &\Longrightarrow \pi \circ \gamma^\alpha(e^{tX_1}) v = e^{it\lambda(X_1^\alpha)} v \\ &\Longrightarrow d\pi^\alpha(X_1) v = i\lambda(X_1^\alpha) v. \end{aligned} \tag{36.6}$$

Results of §18, analyzing (18.1), imply that $\lambda(X_1^\alpha) = n/2$ for some $n \in \mathbb{Z}$, hence

$$\frac{\lambda(H_\alpha)}{\langle \alpha, \alpha \rangle} = \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = \frac{n}{2}, \quad \text{for some } n \in \mathbb{Z}. \tag{36.7}$$

Note that

$$d\gamma^\alpha(X_2 \mp iX_3) = \mp i\sqrt{\frac{2}{\langle\alpha,\alpha\rangle}}(x_\alpha \pm iy_\alpha), \tag{36.8}$$

hence, taking into account (18.10), 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 (36.7), for $\alpha > 0$. We record the result.

Proposition 36.1. *Let π be a unitary representation of G on V. Then for each root α of $\mathfrak{g}$ and each weight λ of π,*

$$2\frac{\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} \tag{36.9}$$

is an integer. If $\alpha > 0$ and V_λ is annihilated by all raising operators (e.g., if λ is a highest weight), then (36.9) 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 (19.8), which is orthonormal with respect to the Ad-invariant Hilbert-Schmidt inner product on $\mathfrak{g} = \mathrm{u}(n)$. 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 (19.13), 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

$$2\frac{\langle\lambda,\omega_{jk}\rangle}{\langle\omega_{jk},\omega_{jk}\rangle} = d_j - d_k. \tag{36.10}$$

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 21.1.

Example 2. Take $G = \mathrm{SU}(n)$. If $\mathrm{u}(n)_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus (\oplus_{j\neq k}\mathfrak{g}_{\omega_{jk}})$, then $\mathrm{su}(n)_{\mathbb{C}}$ has the same form, with $\mathfrak{h}$ replaced by $\tilde{\mathfrak{h}}$, the codimension one subspace of $\mathfrak{h}$ defined by

$$\tilde{\mathfrak{h}} = \{h \in \mathfrak{h} : \mathrm{Tr}\, h = 0\}. \tag{36.11}$$

It is natural to take the order on $\tilde{\mathfrak{h}}$ induced from that on $\mathfrak{h}$ via inclusion. The roots for $\mathrm{su}(n)$ 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

$$\tilde{\mathfrak{h}}' = \mathfrak{h}'/\{(r, \dots, r) : r \in \mathbb{R}\}. \tag{36.12}$$

The weights are equivalence classes $(d_1, \dots, d_n) \sim (d_1 + r, \dots, d_n + r)$, and (36.10) holds in this context; note that $d_j - d_k = (d_j + r) - (d_k + r)$. Again the condition that

(36.10) 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 $(d_1, \ldots, d_{n-1}, 0)$, 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 SU(n) in Proposition 22.2.

REMARK. In Example 2 we see that the necessary condition given in Proposition 36.1 for an element $\lambda \in \mathfrak{h}'$ to be the highest weight for some irreducible representation of SU(n) is also sufficient. By contrast, in Example 1 the necessary condition given in Proposition 36.1 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 SU(n) has a trivial center, while the center of the Lie algebra of U(n) is $\{iaI : a \in \mathbb{R}\}$, which is nontrivial. The following result completes Proposition 36.1.

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 (36.9) 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 [Wal]. We discuss an approach to obtaining such a proof. Namely, one produces a certain finite set $\{\lambda_1, \ldots, \lambda_K\} \subset \mathfrak{h}'$ of dominant integral 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 36.2. *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 21.4.

Recall that this was the program used in §21 to classify the irreducible representations of U(n).

We turn our attention to the adjoint representation of G on $\mathfrak{g}_{\mathbb{C}}$. If we apply Proposition 36.1 to the adjoint representation, we get:

Corollary 36.3. *If α and β are roots, then*

$$n_{\alpha\beta} = 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}. \tag{36.13}$$

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

$$P_\alpha \beta = \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \alpha. \tag{36.14}$$

Hence Corollary 36.3 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 (36.13). Comparing the results implies the following. If $\theta_{\alpha\beta}$ denotes the angle between α and β in $\mathfrak{h}'$, then

$$\cos^2 \theta_{\alpha\beta} = \frac{n_{\alpha\beta} n_{\beta\alpha}}{4}, \tag{36.15}$$

and hence, since the numerator must be an integer,

$$\cos^2 \theta_{\alpha\beta} \in \Big\{ 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1 \Big\}. \tag{36.16}$$

It also follows that

$$n_{\alpha\beta} \in \{0, \pm 1, \pm 2, \pm 3, \pm 4\}. \tag{36.17}$$

More precisely, we have the following. Assume

$$\langle \alpha, \beta \rangle \neq 0, \quad \langle \alpha, \alpha \rangle \geq \langle \beta, \beta \rangle \quad (\text{so } n_{\alpha\beta} \leq n_{\beta\alpha}). \tag{36.18}$$

Then, with $\sigma = \pm 1$,

$$\cos^2 \theta_{\alpha\beta} = \frac{1}{4} \Longleftrightarrow n_{\alpha\beta} = n_{\beta\alpha} = \sigma \Longrightarrow \langle \alpha, \alpha \rangle = \langle \beta, \beta \rangle, \tag{36.19}$$

$$\cos^2 \theta_{\alpha\beta} = \frac{1}{2} \Longleftrightarrow n_{\alpha\beta} = \sigma, \ n_{\beta\alpha} = 2\sigma \Longrightarrow \langle \alpha, \alpha \rangle = 2\langle \beta, \beta \rangle, \tag{36.20}$$

$$\cos^2 \theta_{\alpha\beta} = \frac{3}{4} \Longleftrightarrow n_{\alpha\beta} = \sigma, \ n_{\beta\alpha} = 3\sigma \Longrightarrow \langle \alpha, \alpha \rangle = 3\langle \beta, \beta \rangle. \tag{36.21}$$

Here is another restriction on the set of roots.

Proposition 36.4. *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, (36.13) 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:

$$\text{If } \beta \text{ is a root, then } 2\beta \text{ is not a root.} \tag{36.22}$$

To see this, consider

$$W_\beta = \mathbb{C}\text{-Span}(H_\beta) \oplus \bigoplus_{k\in\mathbb{Z}} \mathfrak{g}_{k\beta}. \tag{36.23}$$

Then W_β is invariant under $\mathrm{Ad} \circ \gamma^\beta$, i.e., we have a representation π^β of 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 18.2, having weight space decomposition

$$V_{-\ell/2} \oplus V_{-\ell/2+1} \oplus \cdots \oplus V_{\ell/2}, \quad dD_{\ell/2}(X_1) = i\mu \ \text{ on } \ V_\mu. \tag{36.24}$$

Now if we take $\pi = \mathrm{Ad}$ in (36.6) (and replace α by β, and λ by $k\beta$), we get

$$d\pi^\beta(X_1) = ik\beta(X_1^\beta) = ik \ \text{ on } \ \mathfrak{g}_{k\beta}, \tag{36.25}$$

since, by (36.3), $\beta(X_1^\beta) = \beta(H_\beta)/\langle\beta,\beta\rangle = 1$. Hence the only representations $D_{\ell/2}$ of 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 (36.23) and (36.24), 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 $\pi^\beta = \mathrm{Ad} \circ \gamma^\beta$, we must have (36.23) equal to $\mathrm{Span}(H_\beta) \oplus \mathfrak{g}_\beta \oplus \mathfrak{g}_{-\beta}$. This establishes (36.22) and completes the proof of Proposition 36.4.

In light of Proposition 36.4, we can complement (36.19)–(36.21) with

$$\begin{aligned} \cos^2\theta_{\alpha\beta} = 1 &\Longleftrightarrow \alpha = \sigma\beta \\ &\Longrightarrow n_{\alpha\beta} = n_{\beta\alpha} = 2\sigma, \end{aligned} \tag{36.26}$$

where again $\sigma = \pm 1$. Also, we can sharpen (36.17) to

$$n_{\alpha\beta} \in \{0, \pm 1, \pm 2, \pm 3\}. \tag{36.27}$$

37. The Weyl group

In §21 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

$$E_\sigma u_k = u_{\sigma(k)}, \tag{37.1}$$

preserves the maximal torus $\mathbb{T} \subset \mathrm{U}(n)$, consisting of diagonal unitary matrices, and permutes their entries:

$$E_\sigma^{-1} \mathrm{diag}(c_1, \dots, c_n) E_\sigma = \mathrm{diag}(c_{\sigma(1)}, \dots, c_{\sigma(n)}). \tag{37.2}$$

It followed that if π is a unitary representation of $\mathrm{U}(n)$ on V, then applications of $\pi(E_\sigma)$ premute the weight spaces; cf. (21.4).

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

$$W(G) = N(\mathbb{T})/\mathbb{T}, \tag{37.3}$$

where $\mathbb{T}$ is a maximal torus of G and $N(\mathbb{T})$ is the normalizer of $\mathbb{T}$:

$$N(\mathbb{T}) = \{g \in G : g^{-1} x g \in \mathbb{T}, \ \forall\, x \in \mathbb{T}\}. \tag{37.4}$$

Note that

$$g \in N(\mathbb{T}) \Longrightarrow \mathrm{Ad}\, g : \mathfrak{h} \to \mathfrak{h}. \tag{37.5}$$

We define the representation $\mathcal{W}$ of $N(\mathbb{T})$ on $\mathfrak{h}$ by

$$\mathcal{W}(g) = \mathrm{Ad}(g)\big|_{\mathfrak{h}}, \quad \text{for} \ \ g \in N(\mathbb{T}). \tag{37.6}$$

Then $N(\mathbb{T})$ has the contragredient representation $\overline{\mathcal{W}}$ on $\mathfrak{h}'$:

$$\langle \mathcal{W}(g^{-1}) H, \lambda \rangle = \langle H, \overline{\mathcal{W}}(g) \lambda \rangle, \quad g \in N(\mathbb{T}), \ H \in \mathfrak{h}, \ \lambda \in \mathfrak{h}'. \tag{37.7}$$

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 (21.4).

Proposition 37.1. *Let π be a unitary representation of G on V, with weight space decomposition $V = \oplus V_\lambda$. Then*

$$g \in N(\mathbb{T}) \Longrightarrow \pi(g) : V_\lambda \to V_{\overline{\mathcal{W}}(g)\lambda}. \tag{37.8}$$

Proof. Recall that

$$\begin{aligned} V_\lambda &= \{v \in V : d\pi(h)v = i\lambda(h)v, \ \forall\, h \in \mathfrak{h}\} \\ &= \{v \in V : \pi(\operatorname{Exp} h)v = e^{i\lambda(h)}v, \ \forall\, e^h \in \mathbb{T}\}. \end{aligned} \tag{37.9}$$

Now

$$\begin{aligned} g \in N(\mathbb{T}), \ v \in V_\lambda \Longrightarrow \pi(g^{-1})\pi(\operatorname{Exp} h)\pi(g)v &= \pi(\operatorname{Exp} \operatorname{Ad} g^{-1}\, h)v \\ &= e^{i\lambda(\operatorname{Ad} g^{-1}\, h)}v, \end{aligned} \tag{37.10}$$

and

$$\lambda(\operatorname{Ad} g^{-1}\, h) = \langle \mathcal{W}(g^{-1})h, \lambda\rangle = \langle h, \overline{\mathcal{W}}(g)\lambda\rangle, \tag{37.11}$$

so

$$\begin{aligned} g \in N(\mathbb{T}), \ v \in V_\lambda &\Longrightarrow \pi(\operatorname{Exp} h)\pi(g)v = e^{\overline{\mathcal{W}}(g)\lambda(h)}\pi(g)v \\ &\Longrightarrow \pi(g)v \in V_{\overline{\mathcal{W}}(g)\lambda}, \end{aligned} \tag{37.12}$$

as stated in (37.8).

In particular, the Weyl group permutes the roots of $\mathfrak{g}$, at least if $W(G)$ is nontrivial. The following result gives valuable information on how $W(G)$ permutes the roots, and implies that $W(G)$ has lots of elements.

Proposition 37.2. *For each root α, there exists $g_\alpha \in N(\mathbb{T})$ such that*

$$\mathcal{W}(g_\alpha)H = H - 2\frac{\langle H_\alpha, H\rangle}{\langle H_\alpha, H_\alpha\rangle}H_\alpha, \quad \forall\, H \in \mathfrak{h}, \tag{37.13}$$

i.e., $\mathcal{W}(g_\alpha)$ is reflection across the hyperplane in $\mathfrak{h}$ orthogonal to H_α. Hence

$$\overline{\mathcal{W}}(g_\alpha)\lambda = \lambda - 2\frac{\langle \alpha, \lambda\rangle}{\langle \alpha, \alpha\rangle}\alpha, \quad \forall\, \lambda \in \mathfrak{h}'. \tag{37.14}$$

Proof. We will show that (37.13) holds with $g_\alpha = A_\alpha(\pi)$, where

$$A_\alpha(t) = \operatorname{Exp} tX_3^\alpha. \tag{37.15}$$

Here X_3^α is as in (36.3) and $\pi = 3.14159\cdots$. To begin, note that

$$\mathrm{Ad}(A_\alpha(t))H = e^{t\,\mathrm{ad}\,X_3^\alpha}H. \tag{37.16}$$

Now

$$\begin{aligned} H \in \ker\alpha \subset \mathfrak{h} &\Longrightarrow \mathrm{ad}\,e_{\pm\alpha}(H) = 0 \quad (\text{since } [H, e_{\pm\alpha}] = \pm i\alpha(H)e_{\pm\alpha}) \\ &\Longrightarrow \mathrm{ad}\,X_3^\alpha(H) = 0 \\ &\Longrightarrow e^{t\,\mathrm{ad}\,X_3^\alpha}(H) = H. \end{aligned} \tag{37.17}$$

We note parenthetically that by the same reasoning, $\alpha(H) = 0 \Rightarrow \mathrm{ad}\,X_2^\alpha(H) = 0$, and of course $\mathrm{ad}(H_\alpha)H = 0$; that is to say, more generally than (37.17), we have

$$H \in \ker\alpha \Longrightarrow \mathrm{Ad}\circ\gamma^\alpha(g)H = h, \quad \forall\, g \in SU(2). \tag{37.18}$$

Since $\alpha(H) = \langle H, H_\alpha\rangle$, we have $\ker\alpha = (H_\alpha)^\perp$. The result (37.13) (and the containment $A_\alpha(\pi) \in N(\mathbb{T})$) will hence follow from (37.17) together with the result

$$\mathrm{Ad}(A_\alpha(\pi))H_\alpha = -H_\alpha. \tag{37.19}$$

To establish (37.19), we analyze the action of $\mathrm{Ad}(A_\alpha(t))$ on X_1^α (which by (36.3) is parallel to H_α). The commutation relations (36.4) give

$$\mathrm{ad}\,X_3^\alpha(X_1^\alpha \pm iX_2^\alpha) = \mp i(X_1^\alpha \pm iX_2^\alpha), \tag{37.20}$$

hence

$$e^{t\,\mathrm{ad}\,X_3^\alpha}(X_1^\alpha \pm iX_2^\alpha) = e^{\mp it}(X_1^\alpha \pm iX_2^\alpha), \tag{37.21}$$

hence

$$\begin{aligned} \mathrm{Ad}(A_\alpha(t))X_1^\alpha &= e^{t\,\mathrm{ad}\,X_3^\alpha}X_1^\alpha \\ &= \frac{1}{2}e^{t\,\mathrm{ad}\,X_3^\alpha}[(X_1^\alpha + iX_2^\alpha) + (X_1^\alpha - iX_2^\alpha)] \\ &= \frac{1}{2}[e^{-it}(X_1^\alpha + iX_2^\alpha) + e^{it}(X_1^\alpha - iX_2^\alpha)], \end{aligned} \tag{37.22}$$

hence

$$\mathrm{Ad}(A_\alpha(\pi))X_1^\alpha = -X_1^\alpha, \tag{37.23}$$

which gives (37.19). This proves (37.13), and (37.14) follows.

Remark. If $\alpha, \beta \in \mathfrak{h}'$ are both roots, then (37.14) together with (36.13) gives

$$\overline{\mathcal{W}}(g_\alpha)\beta = \beta - n_{\alpha\beta}\alpha, \tag{37.24}$$

where $n_{\alpha\beta}$ are the Cartan integers. We use the notation

$$S_\alpha\beta = \overline{\mathcal{W}}(g_\alpha)\beta. \tag{37.25}$$

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 37.3. *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 [Si], 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.

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 35.3 and (35.6), we have

$$\mathfrak{h} = \mathfrak{z} \oplus \mathrm{Span}\,\{H_\alpha : \alpha \text{ root}\}. \tag{37.26}$$

Consequently,

$$\begin{aligned} g \in N(\mathbb{T}),\ \mathcal{W}(g)H_\alpha &= H_\alpha,\ \forall\, \alpha \\ &\Longrightarrow \mathcal{W}(g) = I \quad \text{on} \quad \mathfrak{h} \\ &\Longrightarrow [g] = [e] \quad \text{in} \quad W(G), \end{aligned} \tag{37.27}$$

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

$$\Delta = \{\alpha \in \mathfrak{h}' : \alpha \text{ root of } \mathfrak{g}\}. \tag{37.28}$$

In particular,

$$\#W(G) \,\big|\, (\#\Delta)!, \tag{37.29}$$

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 37.4. *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 (37.24)–(37.25).*

For a proof of this, see [Si], Chapter 8. It is easy enough to verify in case $G = \mathrm{U}(n)$. In that case,

$$\Delta = \{\omega_{jk} : j \neq k,\, 1 \leq j, k \leq n\}, \tag{37.30}$$

with ω_{jk} as in (19.13). Equivalently, with $\{e_j : 1 \leq j \leq n\}$ the basis of $\mathfrak{h}$ given by (19.8) and $\{e_j' : 1 \leq j \leq n\}$ the dual basis, $\omega_{jk} = e_j' - e_k'$. A calculation gives

$$S_{\omega_{jk}} \omega_{\ell m} = \omega_{\sigma(\ell)\sigma(m)}, \quad \text{where} \quad \sigma = (j\ k), \tag{37.31}$$

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

$$\{S_\sigma : \sigma \in S_n\}, \quad S_\sigma \omega_{\ell m} = \omega_{\sigma(\ell)\sigma(m)}. \tag{37.32}$$

That this exhausts $W(\mathrm{U}(n))$ follows from:

Proposition 37.5. *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 $\omega_{jk} = e'_j - e'_k$, $j \neq k$. Then S has the form (37.32), for some $\sigma \in S_n$.*

Proof. Left to the reader.

Given Proposition 37.5, we have

$$W(\mathrm{U}(n)) \approx S_n. \tag{37.33}$$

We also claim

$$W(\mathrm{SU}(n)) \approx S_n. \tag{37.34}$$

The action of $\sigma \in S_n$ on the maximal torus in $\mathrm{SU}(n)$ is given by a slight modification of (37.1)–(37.2), needed because $\det E_\sigma = \operatorname{sgn} \sigma$. More generally than (37.1), 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

$$E^\theta_\sigma u_k = \theta_k u_{\sigma(k)}. \tag{37.35}$$

Then

$$\det E^\theta_\sigma = \theta_1 \cdots \theta_n \cdot \operatorname{sgn} \sigma, \tag{37.36}$$

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 (37.2):

$$(E^\theta_\sigma)^{-1} \operatorname{diag}(c_1, \dots, c_n) E^\theta_\sigma = \operatorname{diag}(c_{\sigma(1)}, \dots, c_{\sigma(n)}). \tag{37.37}$$

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 $\mathrm{SU}(n)$.

38. 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 (35.31), 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 35.9 that $\langle \xi_\lambda, \eta_\lambda \rangle \neq 0$; normalize so that $\langle \xi_\lambda, \eta_\lambda \rangle = 1$, and set

$$\varphi_\lambda(g) = \langle \pi_\lambda(g)\xi_\lambda, \eta_\lambda \rangle. \tag{38.1}$$

This is just as in (35.31), except that here we record the dependence on λ. This family of functions on G has the following important property.

Proposition 38.1. *If λ and μ are highest weights, then*

$$\varphi_{\lambda+\mu}(g) = \varphi_\lambda(g)\varphi_\mu(g). \tag{38.2}$$

Proof. We know by Proposition 36.2 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 35.10),

$$\begin{aligned} \varphi_{\lambda+\mu}(g) &= \langle \pi_\lambda(g) \otimes \pi_\mu(g)(\xi_\lambda \otimes \xi_\mu), \eta_\lambda \otimes \eta_\mu \rangle \\ &= \varphi_\lambda(g)\varphi_\mu(g), \end{aligned} \tag{38.3}$$

as asserted.

Let us recall the conjugate linear map $C : V_\lambda \to V'_\lambda$ from (7.13)–(7.14), satisfying

$$(u, v) = \langle u, Cv \rangle, \quad \overline{\pi}_\lambda(g) = C\pi_\lambda(g)C^{-1}. \tag{38.4}$$

In this setting we have (up to scaling)

$$\eta_\lambda = C\xi_\lambda, \tag{38.5}$$

and hence

$$\varphi_\lambda(g) = (\pi_\lambda(g)\xi_\lambda, \xi_\lambda), \tag{38.6}$$

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

$$\varphi_\lambda(e) = 1. \tag{38.7}$$

We now demonstrate a connection between φ_λ and the character $\chi_\lambda(g) = \operatorname{Tr} \pi_\lambda(g)$.

Proposition 38.2. *We have*

$$\chi_\lambda(x) = d_\lambda \int_G \varphi_\lambda(g^{-1}xg)\, dg, \tag{38.8}$$

where $d_\lambda = \dim W_\lambda$.

Proof. Denote the right side of (38.8) by $\psi_\lambda(x)$. We have

$$\psi_\lambda(g^{-1}xg) = \psi_\lambda(x), \quad \forall\, x, g \in G. \tag{38.9}$$

That is to say, ψ_λ is central, so by Proposition 8.1 it must be a constant multiple of χ_λ. Since $\psi_\lambda(e) = d_\lambda = \chi_\lambda(e)$, we have the identity (38.8).

39. 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

$$\mathrm{SO}(2) \approx S^1, \quad \mathrm{so}(2) = \mathfrak{h} \approx \mathbb{R}, \quad \text{no roots.} \tag{39.1}$$

Moving on to $n = 3$, as shown in §18, there is a 2-fold covering map

$$\mathrm{SU}(2) \longrightarrow \mathrm{SO}(3), \quad \text{hence} \quad \mathrm{so}(3) \approx \mathrm{su}(2), \tag{39.2}$$

we have $\mathfrak{h}$ spanned by X_1, and root vectors $X_\pm = X_2 \mp iX_3$, with X_j as in (18.3). Hence

$$\dim \mathrm{so}(3) = 3, \quad \mathrm{rank}\ \mathrm{so}(3) = 1, \quad \text{there are 2 roots.} \tag{39.3}$$

Next, SO(4) was also studied in §18. As shown there, there is a 2-fold covering map

$$\mathrm{SU}(2) \times \mathrm{SU}(2) \longrightarrow \mathrm{SO}(4), \quad \text{hence} \quad \mathrm{so}(4) \approx \mathrm{su}(2) \oplus \mathrm{su}(2). \tag{39.4}$$

On each factor we have a piece of $\mathfrak{h}$ and a couple of root vectors (in the complexification), so

$$\dim \mathrm{so}(4) = 6, \quad \mathrm{rank}\ \mathrm{so}(4) = 2, \quad \text{there are 4 roots.} \tag{39.5}$$

As a warm-up for studying SO(5), we take a second look at the maximal torus of SO(4) and the roots of so(4). We have

$$\mathbb{T} = \left\{ \begin{pmatrix} R_{\theta_1} & 0 \\ 0 & R_{\theta_2} \end{pmatrix} : \theta_j \in \mathbb{R}/2\pi\mathbb{Z} \right\}, \quad R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{39.6}$$

Equivalently,

$$\mathfrak{h} = \left\{ D_{a,b} = \begin{pmatrix} aJ & 0 \\ 0 & bJ \end{pmatrix} : a, b \in \mathbb{R} \right\}, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{39.7}$$

To get the root spaces $\mathfrak{g}_\alpha$, we decompose the following linear complement to $\mathfrak{h}_\mathbb{C}$ in $\mathrm{so}(4)_\mathbb{C}$,

$$\left\{ A_C = \begin{pmatrix} 0 & C \\ -C^t & 0 \end{pmatrix} : C \in \mathrm{M}(2, \mathbb{C}) \right\}, \tag{39.8}$$

into 4 pieces, each of complex dimension 1, joint eigenvectors for the ad $\mathfrak{h}$ action. A computation gives

$$[D_{a,b}, A_C] = \begin{pmatrix} 0 & aJC - bCJ \\ -bJC^t + aC^tJ & 0 \end{pmatrix}. \tag{39.9}$$

Thus we look for $C_k \in \mathrm{M}(2, \mathbb{C})$ such that

$$aJC_k - bC_kJ = i\alpha_k(a,b)C_k, \tag{39.10}$$

or equivalently,

$$JC_k = i\alpha_k(1,0)C_k, \quad C_kJ = i\alpha_k(0,1)C_k. \tag{39.11}$$

Such matrices can be found by inspection from the formulas

$$J\begin{pmatrix} 1 \\ \pm i \end{pmatrix} = \mp i\begin{pmatrix} 1 \\ \pm i \end{pmatrix}, \quad (1, \pm i)J = \pm i(1, \pm i). \tag{39.12}$$

One obtains
(39.13)

$$C_1 = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix}, \quad C_3 = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}, \quad C_4 = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix},$$

for which (39.10)–(39.11) hold with
(39.14)

$$\alpha_1(a,b) = -(a+b), \quad \alpha_2(a,b) = a+b, \quad \alpha_3(a,b) = -(a-b), \quad \alpha_4(a,b) = a-b.$$

We now tackle the case $n = 5$. A maximal torus of SO(5) is given by

$$\mathbb{T} = \left\{ \begin{pmatrix} R_{\theta_1} & & \\ & R_{\theta_2} & \\ & & 1 \end{pmatrix} : \theta_j \in \mathbb{R}/2\pi\mathbb{Z} \right\}, \tag{39.15}$$

with R_θ as in (39.6). In this case,

$$\mathfrak{h} = \left\{ D_{a,b} = \begin{pmatrix} aJ & & \\ & bJ & \\ & & 0 \end{pmatrix} : a, b \in \mathbb{R} \right\}, \tag{39.16}$$

with J as in (39.7). Parallel to (39.5), we have

$$\dim \mathrm{so}(5) = 10, \quad \mathrm{rank}\ \mathrm{so}(5) = 2, \quad \text{there are 8 roots.} \tag{39.17}$$

Four of the root spaces are spanned by

$$\begin{pmatrix} 0 & C_j & \\ -C_j^t & 0 & \\ & & 0 \end{pmatrix}, \tag{39.18}$$

with C_j as in (39.13). The corresponding roots $\alpha \in \mathfrak{h}' \approx \mathbb{R}^2$, with $\mathfrak{h} \approx \{(a,b) : a,b \in \mathbb{R}\}$ via (39.16), are again given by (39.14).

The other 4 root spaces are 1-dimensional subspaces of

$$\left\{ E_{v,w} = \begin{pmatrix} & & v \\ & & w \\ -v^t & -w^t & 0 \end{pmatrix} : v, w \in \mathbb{R}^2 \right\}. \tag{39.19}$$

A computation gives

$$[D_{a,b}, E_{v,w}] = \begin{pmatrix} & & aJv \\ & & bJw \\ av^tJ & bw^tJ & 0 \end{pmatrix}. \tag{39.20}$$

Referring to (39.12), we take

$$v_j = w_j = \begin{pmatrix} 1 \\ (-1)^{j-1} i \end{pmatrix}, \tag{39.21}$$

to get

$$[D_{a,b}, E_{v_j,0}] = i\beta_j^+(a,b) E_{v_j,0} \tag{39.22}$$

with

$$\beta_j^+(a,b) = (-1)^j a, \tag{39.23}$$

and

$$[D_{a,b}, E_{0,w_j}] = i\beta_j^-(a,b) E_{0,w_j}, \tag{39.24}$$

with

$$\beta_j^-(a,b) = (-1)^j b. \tag{39.25}$$

In summary, the 8 roots of so(5) are α_j $(1 \le j \le 4)$, given by (39.14), and $\beta_j^\pm$ $(1 \le j \le 2)$, given by (39.23) and (39.24). These roots, expanded with respect to the basis dual to $\{D_{1,0}, D_{0,1}\}$ of $\mathfrak{h}$, from (39.16), are depicted in Fig. 39.1.

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 (37.24)–(37.25), 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 37.4. We also have

$$W(\mathrm{SO}(5)) \approx D_4. \tag{39.26}$$

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 $\mathrm{so}(5)_{\mathbb{C}}$ is irreducible. Equivalently,

$$\mathrm{so}(5)_{\mathbb{C}} \ \text{ is simple,} \tag{39.27}$$

as a complex Lie algebra.

Having looked at the roots of 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

$$V_\lambda = \{v \in \mathbb{C}^5 : Hv = i\lambda(H)v,\ \forall\, H \in \mathfrak{h}\}. \tag{39.28}$$

Here H takes the form $D_{a,b}$ of (39.16). We have

$$D_{a,b}\begin{pmatrix}1\\ \pm i\\ 0\\ 0\\ 0\end{pmatrix} = \mp ia\begin{pmatrix}1\\ \pm i\\ 0\\ 0\\ 0\end{pmatrix}, \quad \text{weights } \ \lambda_1^{\mp}(a,b) = \mp a, \tag{39.29}$$

$$D_{a,b}\begin{pmatrix}0\\ 0\\ 1\\ \pm i\\ 0\end{pmatrix} = \mp ib\begin{pmatrix}0\\ 0\\ 1\\ \pm i\\ 0\end{pmatrix}, \quad \text{weights } \ \lambda_2^{\mp}(a,b) = \mp b, \tag{39.30}$$

and

$$D_{a,b}\begin{pmatrix}0\\ 0\\ 0\\ 0\\ 1\end{pmatrix} = 0, \quad \text{weight } \ \lambda_3(a,b) = 0. \tag{39.31}$$

In summary:

$$\begin{aligned} &\lambda_1^+ = a = \beta_2^+, \quad \lambda_1^- = -a = \beta_1^+ \\ &\lambda_2^+ = b = \beta_2^-, \quad \lambda_2^- = -b = \beta_1^-, \quad \lambda_3 = 0. \end{aligned} \tag{39.32}$$

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 §40.

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

$$\begin{aligned} &\mu_1, \mu_2 \text{ highest weights for irreducible representations of } G \\ &\qquad \Longrightarrow \mu_1 + \mu_2 \text{ highest weight for an irreducible representation.} \end{aligned} \tag{39.33}$$

Cf. Proposition 36.2. See Figure 39.2 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 §36. 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 squares in Figure 39.2.

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

$$\mathrm{Spin}(n) \longrightarrow \mathrm{SO}(n). \tag{39.34}$$

This group will be constructed in §42, following prerequisite material on Clifford algebras presented in §41. Section 43 will present spinor representations of Spin(n). For $n = 5$, there will be such a representation with highest weight $\alpha_2/2$.

40. Representations of $\mathrm{SO}(n)$, general n

We present some results on irreducible representations of $\mathrm{SO}(n)$ valid for general n. To start, note that $\mathrm{SO}(n)$ (and more generally $\mathrm{Gl}(n)$) acts on $\Lambda^\ell\mathbb{R}^n$ for each $\ell \in \{0, 1, \dots, n\}$, via

$$\Lambda^\ell(g)\, v_1 \wedge \cdots \wedge v_\ell = gv_1 \wedge \cdots \wedge gv_\ell. \tag{40.1}$$

This extends by complexification to $\Lambda^\ell(g) : \Lambda^\ell\mathbb{C}^n \to \Lambda^\ell\mathbb{C}^n$.

Proposition 40.1. *The representation Λ^ℓ of $\mathrm{SO}(n)$ 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 $\mathrm{SO}(n)$ action, and assume $V \neq 0$. Pick a nonzero $\varphi \in V$, and write

$$\varphi = \sum a_{i_1\cdots i_\ell}\, e_{i_1} \wedge \cdots \wedge e_{i_\ell}, \tag{40.2}$$

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

$$\varphi = e_{i_1} \wedge \cdots \wedge e_{i_\ell}. \tag{40.3}$$

Then as g runs over elements $E^\theta_\sigma \in \mathrm{SO}(n)$ given by

$$E^\theta_\sigma e_j = \theta_j e_{\sigma(j)}, \quad \sigma \in S_n,\ \theta_j = \pm 1, \tag{40.4}$$

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 (40.2) 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

$$a \in \{i_1, \dots, i_\ell\}, \quad a \notin \{j_1, \dots, j_\ell\}, \quad b \notin \{i_1, \dots, i_\ell\} \cup \{j_1, \dots, j_\ell\}. \tag{40.5}$$

Choose $g \in \mathrm{SO}(n)$ so that

$$ge_a = -e_a, \quad ge_b = -e_b, \quad ge_j = e_j \ \text{ otherwise.} \tag{40.6}$$

Then

$$\psi = \varphi + \Lambda^\ell(g)\varphi \tag{40.7}$$

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 40.2. *For $0 \le \ell \le n$, the representations Λ^ℓ of $\mathrm{SO}(n)$ on $\Lambda^\ell \mathbb{C}^n$ and $\Lambda^{n-\ell}$ of $\mathrm{SO}(n)$ on $\Lambda^{n-\ell}\mathbb{C}^n$ are equivalent.*

Proof. We bring in the Hodge star operator

$$* : \Lambda^\ell \mathbb{R}^n \longrightarrow \Lambda^{n-\ell}\mathbb{R}^n, \tag{40.8}$$

defined for $\psi \in \Lambda^\ell \mathbb{R}^n$ by

$$\varphi \wedge *\psi = \langle \varphi, \psi \rangle \, \omega, \quad \forall \, \varphi \in \Lambda^\ell \mathbb{R}^n, \tag{40.9}$$

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

$$* \circ \Lambda^\ell(g) = \Lambda^{n-\ell}(g) \circ * \tag{40.10}$$

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

$$* : \Lambda^\ell \mathbb{C}^n \longrightarrow \Lambda^{n-\ell}\mathbb{C}^n \tag{40.11}$$

by $\mathbb{C}$-linearity, and (40.10) continues to hold. To complete the proof of Proposition 40.2, we note that $*$ in (40.11) is an isomorphism. In fact, a calculation gives

$$** = (-1)^{\ell(n-\ell)} \ \text{ on } \ \Lambda^\ell \mathbb{C}^n. \tag{40.12}$$

Proposition 40.2 finishes the proof of all the statements in Proposition 40.1 about the action of $\mathrm{SO}(n)$ on $\Lambda^\ell \mathbb{C}^n$ for $\ell \ne n/2$. In case $n = 2k$ and $\ell = k$, we have the $\mathrm{SO}(2k)$ action commuting with $* : \Lambda^k \mathbb{C}^{2k} \to \Lambda^k \mathbb{C}^{2k}$. Note that if $1 \le i_1 < \cdots < i_k \le 2k$,

$$*e_{i_1} \wedge \cdots e_{i_k} = \pm e_{j_1} \wedge \cdots \wedge e_{j_k}, \quad \{i_1, \dots, i_k\} \cup \{j_1, \dots, j_k\} = \{1, \dots, 2k\}, \tag{40.13}$$

so $*$ is not a multiple of the identity. According to (40.12),

$$*^2 = (-1)^{k^2} = (-1)^k \ \text{ on } \ \Lambda^k \mathbb{C}^{2k}. \tag{40.14}$$

Hence

$$\begin{aligned} k \text{ even} &\Longrightarrow \mathrm{Spec}\, * = \{\pm 1\} \ \text{ on } \ \Lambda^k \mathbb{C}^{2k}, \\ k \text{ odd} &\Longrightarrow \mathrm{Spec}\, * = \{\pm i\} \ \text{ on } \ \Lambda^k \mathbb{C}^{2k}. \end{aligned} \tag{40.15}$$

(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 (40.14), $*$ 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:

$$\begin{aligned} \Lambda^k\mathbb{C}^{2k} &= \Lambda^k_+\mathbb{C}^{2k} \oplus \Lambda^k_-\mathbb{C}^{2k}, \\ \Lambda^k_\pm\mathbb{C}^{2k} &= \ \pm 1 \ \text{ eigenspace of } * \text{ for } k \text{ even} \\ & \quad\ \ \pm i \ \text{ eigenspace of } * \text{ for } k \text{ odd.} \end{aligned} \tag{40.16}$$

We also have

$$\begin{aligned} g \in \mathrm{O}(n),\ \det g = -1 &\Longrightarrow * \circ \Lambda^\ell(g) = -\Lambda^{n-\ell}(g) \circ *, \quad \forall\, \ell \in \{1,\dots,n\} \\ &\Longrightarrow \Lambda^k(g) : \Lambda^k_+\mathbb{C}^{2k} \overset{\approx}{\to} \Lambda^k_-\mathbb{C}^{2k}, \end{aligned} \tag{40.17}$$

the latter when $n = 2k$.

The following is one complement to Proposition 40.1.

Proposition 40.3. *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 40.1. Say $V \subset \Lambda^k\mathbb{C}^{2k}$ is invariant under the O$(2k)$ action. Take nonzero $\varphi \in V$, represented as in (40.2). If φ has the form (40.3), the argument given before implies $V = \Lambda^k\mathbb{C}^{2k}$. If there are at least two nonzero coefficients in (40.2), 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

$$ge_a = -e_a, \quad ge_j = e_j \ \text{ otherwise.} \tag{40.18}$$

Then, as in (40.7),

$$\psi = \varphi + \Lambda^k(g)\varphi \tag{40.19}$$

has fewer non-vanishing coefficients than φ, but is does have at least one. As in Proposition 40.1, an induction finishes the proof of irreducibility.

Here is another complement to Proposition 40.1.

Proposition 40.4. *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

$$V_- = \Lambda^k(g_0)V_+, \tag{40.20}$$

a subspace of $\Lambda^k_-\mathbb{C}^{2k}$, by (40.17). Consider

$$V = V_+ \oplus V_- \subset \Lambda^k\mathbb{C}^{2k}. \tag{40.21}$$

We have

$$g \in \mathrm{O}(2k) \Longrightarrow \Lambda^k(g) : V \to V, \tag{40.22}$$

and hence, by Proposition 40.2, $V = \Lambda^k \mathbb{C}^{2k}$. 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 $\mathrm{SO}(n)$. Assume $n = k$ or $n = 2k+1$. For $1 \le j \le k$, define $R_j(\theta)$ by

$$\begin{aligned} R_j(\theta)e_{2j-1} &= (\cos\theta)e_{2j-1} + (\sin\theta)e_{2j} \\ R_j(\theta)e_{2j} &= -(\sin\theta)e_{2j-1} + (\cos\theta)e_{2j} \\ R_j(\theta)e_i &= e_i \quad \text{for} \quad i \notin \{2j-1, 2j\}. \end{aligned} \tag{40.23}$$

Then we take

$$\mathbb{T} = \{R_1(\theta_1)\cdots R_k(\theta_k) : \theta_j \in \mathbb{R}/2\pi\mathbb{Z}\}. \tag{40.24}$$

The Lie algebra $\mathfrak{h}$ of $\mathbb{T}$ is spanned by $\{J_{2j-1,2j} : 1 \le j \le k\}$, where, for $1 \le i < j \le k$,

$$J_{ij}e_i = e_j, \quad J_{ij}e_j = -e_i, \quad J_{ij}e_m = 0 \quad \text{for} \quad m \notin \{i, j\}. \tag{40.25}$$

Let us also set

$$E_j = J_{2j-1,2j}. \tag{40.26}$$

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 $\{u_1, \dots, u_n\}$, given by

$$\begin{aligned} u_{2j-1} &= \frac{1}{\sqrt{2}}(e_{2j-1} - ie_{2j}), \\ u_{2j} &= \frac{1}{\sqrt{2}}(e_{2j-1} + ie_{2j}), \quad 1 \le j \le k, \\ u_n &= e_n, \quad \text{if} \quad n = 2k+1. \end{aligned} \tag{40.27}$$

We have

$$\begin{aligned} E_j u_{2j-1} &= iu_{2j-1}, \\ E_j u_{2j} &= -iu_{2j}, \\ E_j u_i &= 0, \quad \text{if} \quad i \notin \{2j-1, 2j\}. \end{aligned} \tag{40.28}$$

Since

$$d\Lambda^\ell(E_j)u_{i_1}\wedge\cdots\wedge u_{i_\ell} = E_j u_{i_1}\wedge u_{i_2}\wedge\cdots\wedge u_{i_\ell} + \cdots + u_{i_1}\wedge\cdots\wedge u_{i_{\ell-1}}\wedge E_j u_{i_\ell}, \tag{40.29}$$

we get

$$\begin{aligned} &d\Lambda^\ell(E_j)u_{i_1}\wedge\cdots\wedge u_{i_\ell} \\ &\quad= \begin{cases} i\,u_{i_1}\wedge\cdots\wedge u_{i_\ell}, & \text{if } 2j-1\in\{i_1,\dots,i_\ell\},\ \text{and } 2j\notin\{i_1,\dots,i_\ell\}, \\ -i\,u_{i_1}\wedge\cdots\wedge u_{i_\ell}, & \text{if } 2j\in\{i_1,\dots,i_\ell\},\ \text{and } 2j-1\notin\{i_1,\dots,i_\ell\}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{40.30}$$

Hence we have the following.

Proposition 40.5. *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 (40.30), for the representation Λ^ℓ of $\mathrm{SO}(n)$ on $\Lambda^\ell\mathbb{C}^n$. In particular, this representation has highest weight given by the following k-tuple, if $n = 2k$ or $2k+1$:

$$(1,\dots,1,0,\dots,0)\quad (\ell\ \textit{ones}), \tag{40.31}$$

with highest weight vector

$$u_1\wedge u_3\wedge\cdots\wedge u_{2\ell-1}. \tag{40.32}$$

Recall Proposition 40.2, which then takes care of the cases $n/2<\ell\le n$. Finally, consider the case $n=2k$, $\ell=k$. The calculations (40.27)–(40.30) still apply. We have weight vectors

$$\begin{aligned} \varphi_0 &= u_1\wedge u_3\wedge\cdots\wedge u_{2k-1}, \quad \text{weight } (1,\dots,1,1)\ (k\text{-tuple}), \\ \varphi_1 &= u_1\wedge u_3\wedge\cdots\wedge u_{2k-3}\wedge u_{2k}, \quad \text{weight } (1,\dots,1,-1). \end{aligned} \tag{40.33}$$

These are the two highest weights for the representation Λ^k of $\mathrm{SO}(2k)$ 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 (40.17), 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 $\mathrm{SO}(2k)$ on $\Lambda^k_{-\sigma}\mathbb{C}^{2k}$. We summarize:

Proposition 40.6. *The representations* $\Lambda^k_\pm$ *of* $\mathrm{SO}(2k)$ *on* $\Lambda^k_\pm\mathbb{C}^{2k}$ *have highest weights given by k-tuples*

$$(1,\dots,1,1) \quad \text{and} \quad (1,\dots,1,-1), \tag{40.34}$$

in some order.

We next take a second look at $\Lambda^2\mathbb{C}^n$. This has the following significance. There is an isomorphism

$$\begin{aligned} A : \Lambda^2\mathbb{R}^n &\longrightarrow \mathrm{Skew}(n) = \mathrm{so}(n) \\ A : \Lambda^2\mathbb{C}^n &\longrightarrow \mathrm{so}_{\mathbb{C}}(n), \end{aligned} \tag{40.35}$$

defined by

$$A(u \wedge v)x = \langle u, x\rangle v - \langle v, x\rangle u, \tag{40.36}$$

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, $g \in \mathrm{O}(n)$),

$$\begin{aligned} gA(u \wedge v)g^{-1}x &= \langle u, g^{-1}x\rangle gv - \langle v, g^{-1}x\rangle gu \\ &= \langle gu, x, \rangle gv - \langle gv, x\rangle gu \\ &= A(gu \wedge gv)x. \end{aligned} \tag{40.37}$$

In other words, A intertwines the representations Λ^2 and Ad. We record the consequence:

Proposition 40.7. *The representation* Λ^2 *of* $\mathrm{SO}(n)$ *on* $\Lambda^2\mathbb{C}^n$ *is unitarily equivalent to the adjoint representation of* $\mathrm{SO}(n)$ *on* $\mathrm{so}_{\mathbb{C}}(n)$*.*

Thus the study of Λ^2 has the potential to reveal information about the structure of the Lie algebra $\mathrm{so}(n)$. In particular, the nonzero weights of Λ^2 are the *roots* of $\mathrm{so}(n)$.

In our second look at Λ^2, we relabel the basis (40.27) of $\mathbb{C}^n$ as follows. For convenience, assume $n \ge 4$. Set

$$v_{j\varepsilon} = \frac{1}{\sqrt{2}}(e_{2j-1} - i\varepsilon e_{2j}), \quad 1 \le j \le k, \quad \varepsilon = \pm 1. \tag{40.38}$$

Then $\{v_{j\varepsilon} : 1 \le j \le k,\ \varepsilon = \pm 1\}$ forms a basis of $\mathbb{C}^n$ if $n = 2k$. If $n = 2k+1$, complete the basis by taking

$$v_n = e_n \quad \text{if} \quad n = 2k+1. \tag{40.39}$$

Parallel to (40.28), we have

$$E_j v_{i\varepsilon} = i\varepsilon\delta_{ij} v_{i\varepsilon}, \qquad E_j v_n = 0 \ \text{ if } \ n = 2k+1. \tag{40.40}$$

In the current situation, a basis for $\Lambda^2\mathbb{C}^n$ is given by

$$\{v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2} : 1 \le i_1 < i_2 \le k,\ \varepsilon_1, \varepsilon_2 = \pm 1\} \cup \{v_{i,1} \wedge v_{i,-1} : 1 \le i \le k\}, \tag{40.41}$$

if $n = 2k$, and if $n = 2k+1$ we complete the basis by taking

$$\{v_{i\varepsilon} \wedge v_n : 1 \le i \le k,\ \varepsilon = \pm 1\}. \tag{40.42}$$

Now we calculate $d\Lambda^2(E_j)$ on this basis.
First, we have

$$d\Lambda^2(E_j)(v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2}) = i(\delta_{ji_1}\varepsilon_1 + \delta_{ji_2}\varepsilon_2) v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2}, \tag{40.43}$$

so each $v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2}$ is a weight vector, with weight

$$\begin{matrix} (0, \dots, & \varepsilon_1, & \dots, & \varepsilon_2, & \dots, 0) \\ & \uparrow & & \uparrow & \\ & i_1 & & i_2 & \end{matrix} \qquad (k\text{-tuple}). \tag{40.44}$$

This weight is positive if and only if $\varepsilon_1 = 1$. Next,

$$v_{i,1} \wedge v_{i,-1} = ie_{2i-1} \wedge e_{2i}, \tag{40.45}$$

and

$$d\Lambda^2(E_j)(v_{i,1} \wedge v_{i,-1}) = 0, \tag{40.46}$$

so each $v_{i,1} \wedge v_{i,-1}$ is a weight vector with weight

$$(0, \dots, 0) \qquad (k\text{-tuple}). \tag{40.47}$$

If $n = 2k$, the weights given by (40.44) and (40.47) are all the weights. If $n = 2k+1$, we also have

$$d\Lambda^2(E_j)(v_{i\varepsilon} \wedge v_n) = i\delta_{ij}\varepsilon\, v_{i\varepsilon} \wedge v_n, \tag{40.48}$$

so $v_{i\varepsilon} \wedge v_n$ is a weight vector, with weight

$$\begin{matrix} (0, \dots, & \varepsilon, & \dots, 0) \\ & \uparrow & \\ & i & \end{matrix} \qquad (k\text{-tuple}). \tag{40.49}$$

Taking Proposition 40.7 into account, we have:

Proposition 40.8. *The roots of* so(n) *are given by (40.44) if* $n = 2k$. *If* $n = 2k+1$, *the roots are given by (40.44) and (40.49). The positive roots are given by*

$$\begin{array}{c} (0,\dots,1,\dots,\varepsilon_2,\dots,0) \qquad (k\text{-}tuple), \\ \quad\ \uparrow \qquad\quad \uparrow \\ \quad\ i_1 \qquad\quad i_2 \end{array} \tag{40.50}$$

for $1 \le i_1 < i_2 \le k$, *if* $n = 2k$, *and if* $n = 2k+1$, *also*

$$\begin{array}{c} (0,\dots,1,\dots,0) \qquad (k\text{-}tuple), \\ \uparrow \\ i \end{array} \tag{40.51}$$

for $1 \le i \le k$.

It is useful to record the image under $A : \Lambda^2\mathbb{C}^n \to \mathrm{so}_{\mathbb{C}}(n)$ of the weight vectors given in (40.43), (40.46), and (40.48). The definition (40.36) readily yields

$$A(e_i \wedge e_j) = J_{ij}, \tag{40.52}$$

which is defined in (40.25). Hence $v_{i,1} \wedge v_{i,-1} = ie_{2i-1} \wedge e_{2i} \Rightarrow$

$$A(v_{i,1} \wedge v_{i,-1}) = iJ_{2i-1,2i}, \quad 1 \le i \le k, \tag{40.53}$$

a basis of $\mathfrak{h}_{\mathbb{C}}$, which is expected since the weights (40.47) are zero.

Next $v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2} = (1/2)(e_{2i_1-1} \wedge e_{2i_2-1} - \varepsilon_1\varepsilon_2 e_{2i_1} \wedge e_{2i_2} - i\varepsilon_1 e_{2i_1} \wedge e_{2i_2-1} - i\varepsilon_2 e_{2i_1-1} \wedge e_{2i_2}) \Rightarrow$

(40.54)

$$\begin{aligned} &A(v_{i_1\varepsilon_1} \wedge v_{i_2\varepsilon_2}) = \frac{1}{2}\Big(J_{2i_1-1,2i_2-1} - \varepsilon_1\varepsilon_2 J_{2i_1,2i_2} - i\varepsilon_1 J_{2i_1,2i_2-1} - i\varepsilon_2 J_{2i_1-1,2i_2}\Big), \\ &1 \le i_1 < i_2 \le k, \quad \varepsilon_1, \varepsilon_2 = \pm 1. \end{aligned}$$

These elements span root spaces with roots given by (40.44). In case $n = 2k = 4$, we have

$$\begin{aligned} A(v_{1\varepsilon_1} \wedge v_{2\varepsilon_2}) &= \frac{1}{2}\Big(J_{13} - \varepsilon_1\varepsilon_2 J_{24} - i\varepsilon_1 J_{23} - i\varepsilon_2 J_{14}\Big) \\ &= \frac{1}{2}\begin{pmatrix} & & -1 & i\varepsilon_2 \\ & & i\varepsilon_1 & \varepsilon_1\varepsilon_2 \\ 1 & -i\varepsilon_1 & & \\ -i\varepsilon_2 & -\varepsilon_1\varepsilon_2 & & \end{pmatrix}. \end{aligned} \tag{40.55}$$

Compare the root space calculation (39.8)–(39.13).

If $n = 2k$, the spaces spanned by elements of the form (40.54) give all the root spaces. If $n = 2k+1$, we also have the images under A of (40.48). Then $v_{1\varepsilon} \wedge v_n = (1/\sqrt{2})(e_{2i-1} \wedge e_n - i\varepsilon e_{2i} \wedge e_n) \Rightarrow$

$$A(v_{i\varepsilon} \wedge v_n) = \frac{1}{\sqrt{2}}\Big(J_{2i-1,n} - i\varepsilon J_{2i,n}\Big). \tag{40.56}$$

These elements span root spaces with roots given by (40.49). In case $n = 2k+1 = 5$, we have

$$\begin{aligned} A(v_{1\varepsilon} \wedge v_5) &= \frac{1}{\sqrt{2}}\Big(J_{15} - i\varepsilon J_{25}\Big) \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} & & & & -1 \\ & & & & i\varepsilon \\ & & & & 0 \\ & & & & 0 \\ 1 & -i\varepsilon & 0 & 0 & 0 \end{pmatrix}, \end{aligned} \tag{40.56}$$

with a similar result for $A(v_{2\varepsilon} \wedge v_5)$. Compare the root space calculations (39.19)–(39.21).

Recall from §36 the definition of a dominant integral weight, namely an element $\lambda \in \mathfrak{h}'$ such that

$$2\,\frac{\langle \lambda, \alpha\rangle}{\langle \alpha, \alpha\rangle} \tag{40.57}$$

is a non-negative integer, for each positive root α. Note that $\langle \alpha, \alpha\rangle = 2$ for all roots of the form (40.50) and $\langle \alpha, \alpha\rangle = 1$ for all roots of the form (40.51). We have the following.

Proposition 40.9. *The dominant integral weights for* so$(2k)$ *are given by k-tuples of the form $(d_1, \dots, d_k)$, satisfying*

$$d_1 \geq \cdots \geq d_{k-1} \geq |d_k|, \tag{40.58}$$

where either all the components d_j are integers or they are all (non-integral) half-integers. The dominant integral weights for so$(2k+1)$ *are given by such k-tuples, satisfying*

$$d_1 \geq \cdots \geq d_k \geq 0, \tag{40.59}$$

instead of (40.58).

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

$$(1,0,\dots,0),\ (1,1,0,\dots,0),\dots,(1,\dots,1,1) \ \text{ and } \ (1,\dots,1,-1) \tag{40.60}$$

when $n = 2k$, for $\Lambda^\ell \mathbb{C}^{2k}$, $1 \le \ell \le k-1$, and $\Lambda^k_\pm \mathbb{C}^{2k}$, and they are

$$(1,0,\dots,0),\ (1,1,0,\dots,0),\dots,(1,\dots,1,1) \tag{40.61}$$

when $n = 2k+1$, for $\Lambda^\ell \mathbb{C}^{2k+1}$, $1 \le \ell \le k$.

The dominant integral weights involving half integral d_j are non-negative integral combinations of these plus the k-tuples

$$\Big(\frac{1}{2},\dots,\frac{1}{2},\frac{1}{2}\Big) \ \text{ and } \ \Big(\frac{1}{2},\dots,\frac{1}{2},-\frac{1}{2}\Big), \tag{40.62}$$

for $n = 2k$, and

$$\Big(\frac{1}{2},\dots,\frac{1}{2},\frac{1}{2}\Big), \tag{40.63}$$

for $n = 2k+1$. Constructions of representations of two-fold covers of $\mathrm{SO}(n)$ with these highest weights will be given in §§43–44.

We next specify the Weyl group $W(\mathrm{SO}(n))$ for each $n \ge 3$, or more precisely its image under $\overline{\mathcal{W}}$ in $\mathrm{Gl}(\mathfrak{h}')$ (defined by (37.6)–(37.7)), which we will denote $\overline{\mathcal{W}}(\mathrm{SO}(n))$. Recall that for $n = 2k$ the roots of $\mathrm{so}(n)$ are given by (40.44); denote such roots as

$$\varepsilon_1 E'_{i_1} + \varepsilon_2 E'_{i_2} = \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}, \tag{40.64}$$

where $1 \le i_1 < i_2 \le k$, $\varepsilon_1, \varepsilon_2 \in \{\pm 1\}$ and $\{E'_1,\dots,E'_k\}$ is the basis of $\mathfrak{h}'$ dual to the basis $\{E_1,\dots,E_k\}$, specified by (40.26) (which is orthonormal with respect to an Ad-invariant inner product on $\mathrm{so}(n)$). By Proposition 37.2, the following reflections belong to $\overline{\mathcal{W}}(\mathrm{SO}(2k))$:

$$\rho_{i_1\varepsilon_1 i_2\varepsilon_2}(\lambda) = \lambda - \langle \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}, \lambda\rangle \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}, \tag{40.65}$$

since $\langle \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}, \alpha_{i_1\varepsilon_1 i_2\varepsilon_2}\rangle = 2$. Note that

$$\begin{aligned} \rho_{i_1\varepsilon_1 i_2\varepsilon_2} E'_{i_1} &= -\varepsilon_1\varepsilon_2 E'_{i_2}, \\ \rho_{i_1\varepsilon_1 i_2\varepsilon_2} E'_{i_2} &= -\varepsilon_1\varepsilon_2 E'_{i_1}, \\ \rho_{i_1\varepsilon_1 i_2\varepsilon_2} E'_{\ell} &= E'_\ell, \quad \ell \notin \{i_1, i_2\}. \end{aligned} \tag{40.66}$$

Noting that (40.66) is a function of $-\varepsilon_1\varepsilon_2$, we can relabel these reflections as

$$\{R_{i_1 i_2 \varepsilon} : 1 \le i_1 < i_2 \le k,\ \varepsilon = \pm 1\}, \tag{40.67}$$

given by

$$R_{i_1 i_2 \varepsilon} E'_{i_1} = \varepsilon E'_{i_2}, \quad R_{i_1 i_2 \varepsilon} E'_{i_2} = \varepsilon E'_{i_1}, \quad R_{i_1 i_2 \varepsilon} E'_{\ell} = E'_{\ell}, \quad \ell \notin \{i_1, i_2\}. \tag{40.68}$$

By Proposition 37.4, $\overline{\mathcal{W}}(\mathrm{SO}(2k))$ is the group generated by the set of reflections (40.67).

The roots of $\mathrm{SO}(2k+1))$ are given by (40.64) plus

$$\varepsilon E'_i, \quad 1 \leq i \leq k, \quad \varepsilon = \pm 1. \tag{40.69}$$

Thus, in addition to the reflections (40.67), $\overline{\mathcal{W}}(\mathrm{SO}(2k+1))$ contains the set of reflections

$$\{R_i : 1 \leq i \leq k\}, \tag{40.70}$$

given by

$$R_i E'_i = -E'_i, \quad R_i E'_{\ell} = E'_{\ell}, \quad \ell \neq i. \tag{40.71}$$

By Proposition 37.4, $\overline{\mathcal{W}}(\mathrm{SO}(2k+1))$ is the group generated by the set of reflections given in (40.67) and (40.70). It follows that

$$\overline{\mathcal{W}}(\mathrm{SO}(2k+1)) = \{E^{\theta}_{\sigma} : \sigma \in S_k,\ \theta = (\pm 1, \dots, \pm 1)\}, \tag{40.72}$$

where, as in (40.4), $E^{\theta}_{\sigma} \in \mathrm{End}(\mathfrak{h}')$ is defined by

$$E^{\theta}_{\sigma} E'_j = \theta_j E'_{\sigma(j)}. \tag{40.73}$$

Given this, we have

Proposition 40.10. *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*

$$Q^k = \{x \in \mathbb{R}^k : -1 \leq x_i \leq 1,\ \text{for } 1 \leq i \leq k\}. \tag{40.74}$$

Proof. Each transformation given by (40.73) 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

$$\theta_1 E'_{\sigma(1)}, \dots, \theta_k E'_{\sigma(k)} \tag{40.75}$$

for some permutation σ of $\{1, \dots, k\}$ and some $\theta_j \in \{\pm 1\}$, so S is of the form (40.73).

Remark. Inspection of (40.68) shows that

$$\overline{\mathcal{W}}(\mathrm{SO}(2k)) = \{E^{\theta}_{\sigma} : \sigma \in S_k,\ \theta_1 \cdots \theta_k = 1\}. \tag{40.76}$$

41. 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

$$uv + vu = -2Q(u, v) \cdot 1, \quad \forall\, u, v \in V. \tag{41.1}$$

Formally, we construct $\mathcal{C}\ell(V, Q)$ as

$$\mathcal{C}\ell(V, Q) = \otimes^* V/\mathcal{I}, \tag{41.2}$$

where $\otimes^* V$ is the tensor algebra:

$$\otimes^* V = \mathbb{R} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots, \tag{41.3}$$

and

$$\begin{aligned} \mathcal{I} = &\text{ two-sided ideal generated by } \{u \otimes v + v \otimes u + 2Q(u, v)1 : u, v \in V\} \\ = &\text{ two-sided ideal generated by } \{e_j \otimes e_k + e_k \otimes e_j + 2Q(e_j, e_k)1\}, \end{aligned} \tag{41.4}$$

where $\{e_j\}$ is a basis of V. Note that

$$Q = 0 \Longrightarrow \mathcal{C}\ell(V, Q) \approx \Lambda^* V \quad \text{(the exterior algebra).} \tag{41.5}$$

Here is a fundamental property of $\mathcal{C}\ell(V, Q)$.

Proposition 41.1. *Let $\mathcal{A}$ be an associative algebra with unit, and let*

$$M : V \longrightarrow \mathcal{A} \tag{41.6}$$

be a linear map satisfying

$$M(u)M(v) + M(v)M(u) = -2Q(u, v)1, \tag{41.7}$$

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

$$M : \mathcal{C}\ell(V, Q) \longrightarrow \mathcal{A}, \quad M(1) = 1. \tag{41.8}$$

Proof. Given (41.5), there is a homomorphism $\widetilde{M} : \otimes^* V \to \mathcal{A}$ extending M, with $\widetilde{M}(1) = 1$. The relation (41.6) implies $\widetilde{M} = 0$ on $\mathcal{I}$, so it descends to $\otimes^* V/\mathcal{I} \to \mathcal{A}$, giving (41.7).

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 $p,q\ge 0$ such that $p+q=n$, and take $Q(u,v)=\sum_{j\le p}u_jv_j-\sum_{j>p}u_jv_j$, where $u=\sum u_je_j$ and $v=\sum v_je_j$. In such a case, $\mathcal{C}\ell(V,Q)$ is denoted $\mathcal{C}\ell(p,q)$.

We also define the complexification of $\mathcal{C}\ell(V,Q)$:

$$\mathbb{C}\ell(V,Q)=\mathbb{C}\otimes\mathcal{C}\ell(V,Q). \tag{41.9}$$

(We tensor over $\mathbb{R}$.) Note that taking $e_j\mapsto ie_j$ for $p+1\le j\le n$ gives, whenever $p+q=n$,

$$\mathbb{C}\ell(p,q)\approx\mathbb{C}\ell(n,0),\quad\text{which we denote }\ \mathbb{C}\ell(n). \tag{41.10}$$

Use of the anticommutator relations (41.1) 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

$$u=\sum_{i_\nu=0\text{ or }1}a_{i_1\cdots i_n}e_1^{i_1}\cdots e_n^{i_n}, \tag{41.11}$$

or, equivalently, in the form

$$u=\sum_{k=0}^{n}\sum_{j_1<\cdots<j_k}\tilde{a}_{j_1\cdots j_k}e_{j_1}\cdots e_{j_k}. \tag{41.12}$$

(By convention the $k=0$ summand in (41.12) is $\tilde{a}_\emptyset\cdot 1$.) In other words, we see that

$$\{e_{j_1}\cdots e_{j_k}:0\le k\le n,\ j_1<\cdots<j_k\} \tag{41.13}$$

spans $\mathcal{C}\ell(V,Q)$ Again, by convention, the subset of (41.13) for which $k=0$ is $\{1\}$. It is very useful to know that the following is true.

Proposition 41.2. *The set (41.13) 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 (41.13) spans, the assertion is that the dimension of $\mathcal{C}\ell(p,q)$ is 2^n when $p+q=n$. By (41.10), 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 41.2 is that

$$\{e_{j_1}\wedge\cdots\wedge e_{j_k}:0\le k\le n,j_1<\cdots<j_k\}\ \text{ is a basis of }\ \Lambda^*\mathbb{R}^n, \tag{41.14}$$

where $\{e_1,\dots,e_n\}$ is the standard basis of $\mathbb{R}^n$. We will use this in our proof of Proposition 41.1. See Appendix I for a proof of (41.14).

Given that (41.14) is true, we can define a linear map

$$\alpha : \Lambda^*\mathbb{R}^n \longrightarrow \mathcal{C}\ell(n,0) \tag{41.15}$$

by $\alpha(1) = 1$ and

$$\alpha(e_{j_1} \wedge \cdots \wedge e_{j_k}) = e_{j_1} \cdots e_{j_k}, \tag{41.16}$$

when $1 \le j_1 < \cdots < j_k \le n$. The content of Proposition 41.2 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

$$M : \mathcal{C}\ell(n,0) \longrightarrow \operatorname{End}(\Lambda^*\mathbb{R}^n), \tag{41.17}$$

we begin with a linear map

$$M : \mathbb{R}^n \longrightarrow \operatorname{End}(\Lambda^*\mathbb{R}^n), \tag{41.18}$$

defined on the basis $\{e_1, \dots, e_n\}$ as follows. Define

$$\wedge_j : \Lambda^k\mathbb{R}^n \longrightarrow \Lambda^{k+1}\mathbb{R}^n, \quad \iota_j : \Lambda^k\mathbb{R}^n \longrightarrow \Lambda^{k-1}\mathbb{R}^n \tag{41.19}$$

by

$$\wedge_j(e_{j_1} \wedge \cdots \wedge e_{j_k}) = e_j \wedge e_{j_1} \wedge \cdots \wedge e_{j_k}, \tag{41.19}$$

and

$$\begin{aligned} \iota_j(e_{j_1} \wedge \cdots \wedge e_{j_k} &= (-1)^{\ell-1} e_{j_1} \wedge \cdots \wedge \widehat{e_{j_\ell}} \wedge \cdots \wedge e_{j_k} \quad \text{if } j = j_\ell, \\ &\qquad\qquad 0 \qquad\qquad\qquad\qquad \text{if } j \notin \{j_1, \dots, j_k\}. \end{aligned} \tag{41.21}$$

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 (41.14) 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:

$$\begin{aligned} \wedge_j \wedge_k + \wedge_k \wedge_j &= 0, \\ \iota_j \iota_k + \iota_k \iota_j &= 0, \\ \wedge_j \iota_k + \iota_k \wedge_j &= \delta_{jk}. \end{aligned} \tag{41.22}$$

Now we define M in (41.18) by

$$M(e_j) = M_j = \wedge_j - \iota_j. \tag{41.23}$$

From (41.22) we get

$$M_j M_k + M_k M_j = -2\delta_{jk}. \tag{41.24}$$

Hence Proposition 41.1 applies to give the homomorphism of algebras (41.17), with $M(1) = I$, the identity operator.

We can now prove Proposition 41.2. We define a linear map

$$\beta : \mathcal{C}\ell(n,0) \longrightarrow \Lambda^*\mathbb{R}^n, \quad \beta(u) = M(u)1. \tag{41.25}$$

Recalling the map α from (41.15)–(41.16), we have

$$\begin{aligned} \beta \circ \alpha(e_{j_1} \wedge \cdots \wedge e_{j_k}) &= M(e_{j_1} \cdots e_{j_k})1 \\ &= M(e_{j_1}) \cdots M(e_{j_k})1. \end{aligned} \tag{41.26}$$

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

$$j_1 < \cdots < j_k \Longrightarrow M(e_{j_1}) \cdots M(e_{j_k})1 = e_{j_1} \wedge \cdots \wedge e_{j_k}. \tag{41.27}$$

It follows that α and β are inverses, and that each is a linear isomorphism. This proves Proposition 41.2 (granted (41.14)).

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

$$\{a + be_1 : a, b \in \mathbb{R}\}. \tag{41.28}$$

In $\mathcal{C}\ell(1,0),\ e_1^2 = -1$, so

$$\mathcal{C}\ell(1,0) \approx \mathbb{C}, \quad e_2 \leftrightarrow i. \tag{41.29}$$

Meanwhile, in $\mathcal{C}\ell(0,1),\ e_1^2 = 1$, so $\mathcal{C}\ell(0,1)$ is of the form

$$\begin{aligned} &\{\alpha f_+ + \beta f_- : \alpha, \beta \in \mathbb{R}\} \\ &f_\pm = \frac{1 \pm e_1}{2} \Rightarrow f_\pm^2 = f_\pm,\ f_+ f_- = f_- f_+ = 0, \end{aligned} \tag{41.30}$$

and we have

$$\mathcal{C}\ell(0,1) \approx \mathbb{R} \oplus \mathbb{R} \approx C_{\mathbb{R}}(\{+,1\}), \tag{41.31}$$

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

$$\{a + be_1 + ce_2 + de_1e_2 : a,b,c,d \in \mathbb{R}\}. \tag{41.32}$$

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 (3.1)–(3.3). Hence

$$\mathcal{C}\ell(2,0) \approx \mathbb{H} = \{a + bi + cj + dk\}. \tag{41.33}$$

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

$$\begin{aligned}\mathcal{C}\ell(0,2) &\approx \mathrm{M}(2,\mathbb{R}) \\ &= \Big\{aI + b\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + c\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} + d\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} : a,b,c,d \in \mathbb{R}\Big\}.\end{aligned} \tag{41.34}$$

It turns out that also

$$\mathcal{C}\ell(1,1) \approx M(2,\mathbb{R}).$$

We leave this to the reader.

Using (41.31) and (41.34), we find the complexified algebras

$$\mathbb{C}\ell(1) \approx \mathbb{C} \oplus \mathbb{C}, \quad \mathbb{C}\ell(2) \approx \mathrm{M}(2,\mathbb{C}). \tag{41.35}$$

These results are special cases of the following:

Proposition 41.3. *The complex Clifford algebras $\mathbb{C}\ell(n)$ have the properties*

$$\begin{aligned}\mathbb{C}\ell(2k) &\approx \mathrm{M}(2^k,\mathbb{C}), \\ \mathbb{C}\ell(2k+1) &\approx \mathrm{M}(2^k,\mathbb{C}) \oplus \mathrm{M}(2^k,\mathbb{C}).\end{aligned} \tag{41.36}$$

Proposition 41.3 follows inductively from (41.35) and the following result.

Proposition 41.4. *For $n \in \mathbb{N}$, we have isomorphisms of algebras*

$$\mathbb{C}\ell(n+2) \approx \mathbb{C}\ell(n) \otimes \mathbb{C}\ell(2). \tag{41.37}$$

In turn, Proposition 41.4 follows from:

Proposition 41.5. *For $n \in \mathbb{N}$, we have isomorphisms of algebras*

$$\mathcal{C}\ell(n,0) \otimes \mathcal{C}\ell(0,2) \approx \mathcal{C}\ell(0,n+2). \tag{41.38}$$

It remains to prove (41.38). To do this, we construct a homomorphism of algebras

$$M : \mathcal{C}\ell(0, n+2) \longrightarrow \mathcal{C}\ell(n, 0) \otimes \mathcal{C}\ell(0, 2). \tag{41.39}$$

Once it is checked that M is onto, a dimension count guarantees it is an isomorphism.

To produce (41.39), we start with a linear map

$$M : \mathbb{R}^{n+2} \longrightarrow \mathcal{C}\ell(n, 0) \otimes \mathcal{C}\ell(0, 2), \tag{41.40}$$

defined by

$$\begin{aligned} Me_j &= M_j = e_j \otimes e_{n+1}e_{n+2}, \quad 1 \le j \le n, \\ Me_j &= M_j = 1 \otimes e_j, \qquad j = n+1, n+2. \end{aligned} \tag{41.41}$$

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 (41.40) to (41.39), we need to establish the anticommutation relations

$$M_j M_k + M_k M_j = 2\delta_{jk}, \quad 1 \le j, k \le n+2. \tag{41.42}$$

To get this for $1 \le j, k \le n$, we use the computations

$$\begin{aligned} &(e_{n+1}e_{n+2})^2 = -e_{n+1}^2 e_{n+2}^2 = -1, \\ &(e_j \otimes e_{n+1}e_{n+2})(e_k \otimes e_{n+1}e_{n+2}) = e_j e_k \otimes (e_{n+1}e_{n+2})^2 = -e_j e_k \otimes 1, \end{aligned} \tag{41.43}$$

which yield

$$\begin{aligned} 1 \le j, k \le n \Rightarrow M_j M_k + M_k M_j &= -(e_j e_k \otimes 1 + e_k e_j \otimes 1) \\ &= 2\delta_{jk}, \end{aligned} \tag{41.44}$$

as desired. Next we have

(41.45)

$$\begin{aligned} 1 \le j \le n \Longrightarrow M_j M_{n+1} &+ M_{n+1} M_j \\ &= (e_j \otimes e_{n+1}e_{n+2})(1 \otimes e_{n+1}) + (1 \otimes e_{n+1})(e_j \otimes e_{n+1}e_{n+2}) \\ &= e_j \otimes e_{n+1}e_{n+2}e_{n+1} + e_j \otimes e_{n+1}e_{n+1}e_{n+2} \\ &= 0, \end{aligned}$$

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,

$$M_{n+1} M_{n+1} = (1 \otimes e_{n+1})(1 \otimes e_{n+1}) = 1 \otimes e_{n+1}^2 = 1, \tag{41.46}$$

and similarly $M_{n+2}M_{n+2} = 1$. Finally,
(41.47)

$$\begin{aligned} M_{n+1}M_{n+2} + M_{n+2}M_{n+1} &= (1 \otimes e_{n+1})(1 \otimes e_{n+2}) + (1 \otimes e_{n+2})(1 \otimes e_{n+1}) \\ &= 1 \otimes (e_{n+1}e_{n+2} + e_{n+2}e_{n+1}) \\ &= 0. \end{aligned}$$

This establishes (41.42). Hence, by Proposition 41.1, M extends to the algebra homomorphism (41.39) (with $M1 = I$). It is routine to verify that the elements on the right side of (41.41) generate $\mathcal{C}\ell(n,0) \otimes \mathcal{C}\ell(0,2)$, so M in (41.39) is onto, hence an isomorphism. This completes the proof of Proposition 41.5, hence Propositions 41.3–41.4.

Remark. The following companions to (41.38),

$$\begin{aligned} \mathcal{C}\ell(0,n) \otimes \mathcal{C}\ell(2,0) &\approx \mathcal{C}\ell(n+2,0), \\ \mathcal{C}\ell(p,q) \otimes \mathcal{C}\ell(1,1) &\approx \mathcal{C}\ell(p+1,q+1), \end{aligned} \tag{41.48}$$

have essentially the same proof. From (41.38) and (41.48) it follows that

$$\mathcal{C}\ell(n+8,0) \approx \mathcal{C}\ell(n,0) \otimes \mathcal{C}\ell(0,2) \otimes \mathcal{C}\ell(2,0) \otimes \mathcal{C}\ell(0,2) \otimes \mathcal{C}\ell(2,0). \tag{41.49}$$

Meanwhile,

$$\mathcal{C}\ell(0,2) \otimes \mathcal{C}\ell(2,0) \approx \mathrm{M}(2,\mathbb{R}) \otimes \mathbb{H}. \tag{41.50}$$

This, together with the isomorphism

$$\mathbb{H} \otimes \mathbb{H} \approx \mathrm{M}(4,\mathbb{R}), \tag{41.51}$$

leads to

$$\mathcal{C}\ell(n+8,0) \approx \mathcal{C}\ell(n,0) \otimes \mathrm{M}(16,\mathbb{R}). \tag{41.52}$$

42. 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 wil not deal with this here; cf. [T1], [LM] 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 $x \in V$,

$$\begin{aligned} \tau(v)x &= vxv \\ &= -xvv - 2\langle v, x\rangle v \\ &= x - 2\langle v, x\rangle v. \end{aligned} \tag{42.1}$$

Hence $\tau(v) : V \to V$ is reflection across the hyperplane $(v)^\perp$. With this in mind, we set

$$\mathrm{Pin}(n) = \{v_1 \cdots v_k \in \mathcal{C}\ell(n,0) : k \in \mathbb{N}, v_j \in \mathbb{R}^n,\ \langle v_j, v_j \rangle = 1\}, \tag{42.2}$$

and define

$$\tau : \mathrm{Pin}(n) \longrightarrow \mathrm{O}(n) \tag{42.3}$$

by

$$\begin{aligned} \tau(v_1 \cdots v_k)x &= v_1 \cdots v_k x v_k \cdots v_1 \\ &= \tau(v_1) \cdots \tau(v_k)x, \end{aligned} \tag{42.4}$$

so $\tau(v_1 \cdots v_k)$ is a product of k reflections of the form (42.1).

We need to show that (42.4) is well defined, independently of the representation of an element of Pin(n) as a particular product. The following takes care of this.

Lemma 42.1. *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. To see this, write each v_j as a linear combination of $\{e_1, \dots, e_n\}$, and present $v_1 \cdots v_k$ in the form (41.11). If k is odd, then each summand in (41.11) for $u = v_1 \cdots v_k$ will have an odd number of factors. 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.

Note that $\det \tau(v_1 \cdots v_k) = (-1)^k$. Hence, if we set

$$\text{Spin}(n) = \{v_1 \cdots v_k : k \in 2\mathbb{N},\ v_j \in \mathbb{R}^n,\ \langle v_j, v_j \rangle = 1\}, \tag{42.5}$$

we have

$$\tau : \text{Spin}(n) \longrightarrow \text{SO}(n). \tag{42.6}$$

Note that $\text{Pin}(n)$ and $\text{Spin}(n)$ 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 42.2. *The maps (42.3) and (42.6) are surjective, and*

$$\textit{Ker}\ \tau = \{\pm 1\}. \tag{42.7}$$

First we discuss the surjectivity. For $v \in \mathbb{R}^n$, $\tau(v)$ in (42.1) is a reflection, and each reflection on $\mathbb{R}^n$ has this form. Hence the surjectivity follows from:

Proposition 42.3. *Each $A \in \text{O}(n)$ is a product of reflections.*

Proof. We make use of the basic linear algebra result that given such A, there exist $Q \in \text{O}(n)$ such that $Q^{-1}AQ$ has the block diagonal form

$$Q^{-1}AQ = \begin{pmatrix} R_{\theta_1} & & & & & & & & \\ & \ddots & & & & & & & \\ & & R_{\theta_k} & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix}. \tag{42.8}$$

Since $Q\tau Q^{-1}$ is a reflection whenever τ is a reflection, it suffices to show that the right side of (42.8) is a product of reflections. It suffices to consider the separate blocks. In particular, we show that each rotation

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{42.9}$$

is a product of two reflections in $\mathbb{R}^2$, of the form

$$\rho(v)x = x - 2\langle v, x \rangle v, \quad x, v \in \mathbb{R}^2,\ |v| = 1. \tag{42.10}$$

In fact, it is readily verified that

$$v \in \mathbb{R}^2,\ |v| = 1 \Longrightarrow \rho(R_{\theta/2}v)\rho(v) = R_\theta. \tag{42.11}$$

As far as the diagonal entries ± 1 represented as reflections, this is obvious, so the surjectivity assertion of Proposition 42.2 is proven.

Our next task is to establish (42.7). To tackle this, suppose $v_j \in \mathbb{R}^n$ are unit vectors such that

$$\tau(u) = I, \quad u = v_1 \cdots v_k. \tag{42.12}$$

Since each $\tau(v_j)$ has determinant -1, k is even in (42.12). So $(v_k \cdots v_1)(v_1 \cdots v_k) = 1$. Referring to (42.4), we have

$$\begin{aligned} &\tau(u)x = x, \quad \forall\, x \in \mathbb{R}^n \\ &\Leftrightarrow ux = xu, \quad \forall\, x \in \mathbb{R}^n, \\ &\Leftrightarrow xux = -|x|^2 u, \quad \forall\, x \in \mathbb{R}^n \\ &\Leftrightarrow u = -e_j u e_j, \end{aligned} \tag{42.13}$$

for the standard orthonormal basis $\{e_j\}$ of $\mathbb{R}^n$. Now, using Proposition 41.2, set

$$u = \sum_{i_\nu = 0 \text{ or } 1} a_{i_1 \cdots i_n} e_1^{i_1} \cdots e_n^{i_n}. \tag{42.14}$$

We have, with $i_1 + \cdots + i_n = 2k$,

$$\begin{aligned} &- e_j(e_1^{i_1} \cdots e_j^{i_j} \cdots e_n^{i_n})e_j \\ &= (-1)^{2k - i_j + 1} e_1^{i_1} \cdots e_j^{i_j + 2} \cdots e_n^{i_n} \\ &= (-1)^{2k - i_j + 2} e_1^{i_1} \cdots e_j^{i_j} \cdots e_n^{i_n}. \end{aligned} \tag{42.15}$$

Hence, for u as in (42.14), if (42.13) holds, then

$$u = \sum_{i_\nu = 0 \text{ or } 1} (-1)^{i_j} a_{i_1 \cdots i_n} e_1^{i_1} \cdots e_n^{i_n}, \quad \forall\, j. \tag{42.16}$$

Given Proposition 41.2, we deduce that, for u of the form (42.14),

$$\tau(u) = I \Longrightarrow a_{i_1 \cdots i_n} = 0, \ \text{ except for } a_{0 \cdots 0}, \tag{42.17}$$

which gives (42.7), and completes the proof of Proposition 42.1.

The next result is an important complement to Proposition 42.1, to the effect that (42.6) presents $\mathrm{Spin}(n)$ as a *connected* double cover of $\mathrm{SO}(n)$.

Proposition 42.4. *For each $n \geq 2$, $\mathrm{Spin}(n)$ is connected.*

Proof. Since we know $\mathrm{SO}(n)$ is connected, it suffices to show that there is a continuous path in $\mathrm{Spin}(n)$ from 1 to -1. Set

$$\gamma(t) = e_1 \cdot \big((\cos t)e_1 + (\sin t)e_2\big), \quad 0 \leq t \leq \pi. \tag{42.18}$$

We have $\gamma : [0, \pi] \to \mathrm{Spin}(n)$, and

$$\gamma(0) = -1, \quad \gamma(\pi) = 1, \tag{42.19}$$

so Proposition 42.4 is proven.

We examine the Lie algebra $\mathrm{spin}(n)$ of $\mathrm{Spin}(n)$, i.e., the tangent space to $\mathrm{Spin}(n)$ at 1. The Lie algebra $\mathrm{so}(n)$ of $\mathrm{SO}(n)$ is spanned by elements

$$J_{jk} = -E_{jk} + E_{kj}, \quad j < k, \tag{42.20}$$

where $E_{jk} \in \mathrm{M}(n, \mathbb{R})$ is defined by

$$E_{jk} e_\ell = \delta_{k\ell} e_j. \tag{42.21}$$

The element J_{jk} generates the group $R_{jk}(t) = e^{tJ_{jk}}$ of rotations in the $e_j - e_k$ plane, given by

$$\begin{aligned} R_{jk}(t)e_j &= (\cos t)e_j + (\sin t)e_k \\ R_{jk}(t)e_k &= -(\sin t)e_j + (\cos t)e_k \\ R_{jk}(t)e_\ell &= e_\ell, \qquad \ell \notin \{j, k\}. \end{aligned} \tag{42.22}$$

Comparing (42.1) with (42.10)–(42.11), we see that

$$R_{jk}(t) = \tau\big(R_{jk}(t/2)e_j \cdot e_j\big). \tag{42.23}$$

Since $R'_{jk}(0) = J_{jk}$, this gives

$$d\tau(e_j e_k) = -d\tau(e_k e_j) = -2J_{jk}. \tag{42.24}$$

Hence

$$\mathrm{spin}(n) = \mathrm{Span}\,\{e_j e_k : j < k\} \subset \mathcal{C}\ell(n, 0). \tag{42.25}$$

The one parameter group in $\mathrm{Spin}(n)$ generated by $e_j e_k$ is

$$\mathrm{Exp}(te_j e_k) = -R_{jk}(t)e_j \cdot e_j = -\big((\cos t)e_j + (\sin t)e_k\big) \cdot e_j. \tag{42.26}$$

Note also that for $j < k$

$$(e_j e_k)^2 = -1 \Longrightarrow e^{te_je_k} = \cos t + (\sin t)e_j e_k \tag{42.27}$$
$$= -(\cos t)e_j^2 + (\sin t)e_j e_k$$
$$= \operatorname{Exp}(te_j e_k).$$

Either by a calculation or by applying analogues of reasoning done in Appendix E, we see that the Lie bracket on spin(n) is given by

$$[e_j e_k, e_\ell e_m] = e_j e_k e_\ell e_m - e_\ell e_m e_j e_k, \tag{42.28}$$

and

$$d\tau([e_j e_k, e_\ell e_m]) = [-2J_{jk}, -2J_{\ell m}]. \tag{42.29}$$

The space

$$\mathfrak{h} = \operatorname{Span}\{E_j : 1 \le j \le k\} \subset \mathrm{so}(n), \quad E_j = J_{2j-1,2j}, \tag{42.30}$$

is the Lie algebra of a maximal torus of SO(n), when $n = 2k$ or $n = 2k + 1$. The preimage under $d\tau$ is

$$\tilde{\mathfrak{h}} = \operatorname{Span}\{e_{2j-1}e_{2j} : 1 \le j \le k\} \subset \mathrm{spin}(n). \tag{42.31}$$

43. 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

$$(u, v) = \langle u, v\rangle + i\langle u, Jv\rangle. \tag{43.1}$$

We set

$$S = S(V, \langle \ , \ \rangle, J) = \Lambda^*_{\mathbb{C}}\mathcal{V} = \bigoplus_{j=0}^{k} \Lambda^j_{\mathbb{C}}\mathcal{V}. \tag{43.2}$$

The inner product (43.1) 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

$$\{v_{j_1} \wedge \cdots \wedge v_{j_\ell} : j_1 < \cdots < j_\ell\} \ \text{ is an orthonormal basis of } \ \Lambda^\ell_{\mathbb{C}}\mathcal{V}. \tag{43.3}$$

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

$$S = \Lambda^*_{\mathbb{C}}\mathbb{C}^k. \tag{43.4}$$

In order to define a representation of $\mathcal{C}\ell(n, 0)$ on S, we produce an $\mathbb{R}$-linear map

$$M : \mathbb{R}^n \longrightarrow \mathrm{End}_{\mathbb{C}}(\Lambda^*_{\mathbb{C}}\mathcal{V}), \tag{43.5}$$

in the form

$$M(v) = \wedge_v - j_v, \tag{43.6}$$

where

$$\wedge_v : \Lambda^\ell_{\mathbb{C}}\mathcal{V} \longrightarrow \Lambda^{\ell+1}_{\mathbb{C}}\mathcal{V}, \quad \wedge_v\varphi = v \wedge \varphi, \tag{43.7}$$

with v interpreted as an element of $\mathcal{V}$, and

$$j_v : \Lambda^{\ell+1}_{\mathbb{C}} \longrightarrow \Lambda^{\ell}_{\mathbb{C}}\mathcal{V}, \quad j_v\psi = (\wedge_v)^*\psi, \tag{43.8}$$

that is,

$$(v \wedge \varphi, \psi) = (\varphi, j_v\psi), \quad \varphi \in \Lambda^{\ell}_{\mathbb{C}}\mathcal{V},\ \psi \in \Lambda^{\ell+1}_{\mathbb{C}}\mathcal{V}. \tag{43.9}$$

We claim that, for $u, v \in \mathbb{R}^n$,

$$M(u)M(v) + M(v)M(u) = -2\langle u, v\rangle I. \tag{43.10}$$

It suffices to show that

$$M(v)^2 = -\langle v, v\rangle I, \tag{43.11}$$

and insert $u \pm v$ into this identity. To prove (43.11), 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 $j_1 < \cdots < j_{\ell+1}$,

$$\begin{aligned} j_v(v_{j_1} \wedge \cdots \wedge v_{j_{\ell+1}}) = v_{j_2}\wedge \cdots \wedge v_{j_{\ell+1}} &\quad \text{if } j_1 = 1, \\ 0 &\quad \text{if } j_1 > 1. \end{aligned} \tag{43.12}$$

Since $\wedge_v^2 = 0$ and (hence) $j_v^2 = 0$, we get

$$M(v)^2 = -(\wedge_v j_v + j_v\wedge_v) = -\langle v, v\rangle I, \tag{43.13}$$

the last identity via (43.12).

Now Proposition 41.1 implies M extends to a homomorphism of algebras

$$M : \mathcal{C}\ell(2k, 0) \longrightarrow \mathrm{End}_{\mathbb{C}}(\Lambda^*_{\mathbb{C}}\mathbb{C}^k), \tag{43.14}$$

which in turn extends to a $\mathbb{C}$-linear algebra homomorphism

$$M : \mathbb{C}\ell(2k) \longrightarrow \mathrm{End}_{\mathbb{C}}(\Lambda^*_{\mathbb{C}}\mathbb{C}^k). \tag{43.15}$$

The following is fundamental.

Proposition 43.1. *In (43.14), M is an isomorphism of algebras.*

Proof. Note that $\dim_{\mathbb{C}} \Lambda^*_{\mathbb{C}}\mathbb{C}^k = 2^k$; hence

$$\dim \mathrm{End}_{\mathbb{C}}(\Lambda^*_{\mathbb{C}}\mathbb{C}^k) = 2^{2k} = \dim_{\mathbb{C}} \mathbb{C}\ell(2k). \tag{43.16}$$

Thus it suffices to prove M is injective. Clearly $M(v) \neq 0$ for nonzero $v \in \mathbb{R}^n$, and $\mathrm{Ker}\, M$ must be a two-sided ideal in $\mathbb{C}\ell(2k)$. Recall from (43.35) that $\mathbb{C}\ell(2k) \approx \mathrm{M}(2^k, \mathbb{C})$. It is a fact that, for each $m \in \mathbb{N}$,

$$\mathrm{M}(m, \mathbb{C}) \ \text{ has no proper two-sided ideals;} \tag{43.17}$$

i.e., $\mathrm{M}(m, \mathbb{C})$ is *simple*. See Appendix J for a proof. This finishes the proof of Proposition 43.1.

The algebra homomorphism M in (43.15) restricts to $\mathrm{Pin}(2k) \subset \mathcal{C}\ell(2k, 0)$, yielding a group homomorphism

$$D_{1/2} : \mathrm{Pin}(2k) \longrightarrow \mathrm{Gl}(\Lambda^*_{\mathbb{C}}\mathbb{C}^k), \tag{43.18}$$

i.e., a representation of $\mathrm{Pin}(2k)$ on $\Lambda^*_{\mathbb{C}}\mathbb{C}^k$. Since the linear span of $\mathrm{Pin}(2k)$ (over $\mathbb{R}$) is $\mathcal{C}\ell(2k, 0)$, we have from Proposition 43.1 that:

Corollary 43.2. *The representation* $D_{1/2}$ *of* $\mathrm{Pin}(2k)$ *on* $S = \Lambda^*_{\mathbb{C}}\mathbb{C}^k$ *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 $\mathrm{Spec}\, M(v) \subset \{\pm i\}$, and hence $M(v)$ is unitary.

The restriction of $D_{1/2}$ to $\mathrm{Spin}(2k)$ is not irreducible. In fact, the spaces

$$\bigoplus_{j \text{ even}} \Lambda^*_{\mathbb{C}}\mathbb{C}^k = S_+(2k), \quad \bigoplus_{j \text{ odd}} \Lambda^*_{\mathbb{C}}\mathbb{C}^k = S_-(2k) \tag{43.19}$$

are invariant under the action of $D_{1/2}$ restricted to $\mathrm{Spin}(2k)$, and more generally under the action of M restricted to $\mathcal{C}\ell^0(2k)$. We have

$$M : \mathbb{C}\ell^0(2k) \longrightarrow \mathrm{End}_{\mathbb{C}}(S_+(2k)) \oplus \mathrm{End}_{\mathbb{C}}(S_-(2k)). \tag{43.20}$$

Note that

$$v \in M,\ v \neq 0 \Longrightarrow M(v) : S_\pm(2k) \to S_\mp(2k), \tag{43.21}$$

while left multiplication by v takes $\mathbb{C}\ell^0(2k)$ to $\mathbb{C}\ell^1(2k)$, and these maps are all isomorphisms. In particular,

$$\dim_{\mathbb{C}} S_\pm(2k) = 2^{2k-1} = \dim_{\mathbb{C}} \mathbb{C}\ell^0(2k). \tag{43.22}$$

We already know from Proposition 43.1 that M in (43.20) is injective, so it is an isomorphism. We deduce the following.

Corollary 4.3. *The representation* $D_{1/2}$ *restricted to* $\mathrm{Spin}(2k)$ *splits into two factors:*

$$D^\pm_{1/2} : \mathrm{Spin}(2k) \longrightarrow \mathrm{Gl}(S_\pm(2k)), \tag{43.23}$$

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

$$\mathbb{R}^{2k-1} \longrightarrow \mathcal{C}\ell^0(2k, 0), \quad v \mapsto ve_{2k}, \quad v \in \mathbb{R}^{2k-1} \tag{43.24}$$

gives rise, via Proposition 41.1, to a homomorphism of algebras.

$$\mathcal{C}\ell(2k-1, 0) \longrightarrow \mathcal{C}\ell^0(2k, 0). \tag{43.25}$$

This map is not zero, hence, since $\mathcal{C}\ell(2k-1, 0)$ is simple, it is injective, and thus, since the dimensions match, it is an isomorphism. The inclusion $\mathrm{Spin}(2k-1) \subset \mathcal{C}\ell(2k-1, 0)$ gives an inclusion

$$\mathrm{Spin}(2k-1) \hookrightarrow \mathrm{Spin}(2k), \tag{43.26}$$

and restricting $D^+_{1/2}$ gives a representation

$$D^+_{1/2} : \mathrm{Spin}(2k-1) \longrightarrow \mathrm{Gl}(S_+(2k)). \tag{43.26}$$

There is also a representation $D^-_{1/2}$ of $\mathrm{Spin}(2k-1)$ on $S_-(2k))$, but these two are intertwined by the isomorphism $M(e_{2k}) : S_+(2k) \to S_-(2k)$.

44. Weight spaces for the spinor representations

In §43 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 (42.31), will be

$$\Big\{\frac{1}{2}e_2e_1, \frac{1}{2}e_4e_3, \dots\Big\} = \Big\{\frac{1}{2}e_{2j}e_{2j-1} : 1 \le j \le k\Big\}, \tag{44.1}$$

given $n = 2k$ or $2k + 1$. Recall that this maps via $d\tau$ to the ordered basis

$$\{J_{2j-1,2j} : 1 \le j \le k\} \tag{44.2}$$

for the Lie algebra of a maximal torus of $\mathrm{SO}(n)$, described in (42.20); cf. (42.24). We mention that some sources choose the ordered basis $\{(1/2)e_{2j-1}e_{2j} : 1 \le j \le k\}$, which changes signs and hence the order of the weights.

We first treat the case $n = 2k$. To get started, note that

$$\begin{aligned} &\gamma_{ij}(t) = e^{t(e_ie_j)} = (\cos t)1 + (\sin t)e_ie_j \\ &\Longrightarrow D^\pm_{1/2}(\gamma_{ij}(t))\varphi = (\cos t)\varphi + (\sin t)M(e_i)M(e_j)\varphi \\ &\Longrightarrow dD^\pm_{1/2}(e_ie_j)\varphi = M(e_i)M(e_j)\varphi = M_iM_j\varphi, \end{aligned} \tag{44.3}$$

for $1 \le i, j \le 2k,\ \varphi \in \Lambda^*\mathbb{C}^k$, and, as in (43.6), $M(v) = \wedge_v - j_v,\ v \in \mathbb{R}^{2k}$, and we introduce simplified notation

$$M_i = M(e_i) = \wedge_{e_i} - j_{e_i} = \wedge_i - j_i. \tag{44.4}$$

Hence

$$dD^\pm_{1/2}(e_{2j}e_{2j-1})\varphi = M_{2j}M_{2j-1}\varphi = (\wedge_{2j} - j_{2j})(\wedge_{2j-1} - j_{2j-1})\varphi. \tag{44.5}$$

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

$$\begin{aligned} dD^\pm_{1/2}(e_{2j}e_{2j-1})\varphi &= i(j_{2j}\wedge_{2j} - \wedge_{2j}\, j_{2j})\varphi \\ &= -i(1 - 2j_{2j}\wedge_{2j})\varphi, \end{aligned} \tag{44.6}$$

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

$$\{v_j : 1 \le j \le k\}, \quad v_1 = e_{2j} = ie_{2j-1}. \tag{44.7}$$

We have

$$dD^{\pm}_{1/2}(e_{2j}e_{2j-1}) = iQ_j \tag{44.8}$$

where, for $1 \le i_1 < \cdots < i_\ell \le k$,

$$\begin{aligned} Q_j \; v_{i_1} \wedge \cdots \wedge v_{i_\ell} = \;& v_{i_1} \wedge \cdots \wedge v_{i_\ell} \quad \text{if} \;\; j \notin \{i_1, \dots, i_\ell\} \\ & - v_{i_1} \wedge \cdots \wedge v_{i_\ell} \quad \text{if} \;\; j \in \{i_1, \dots, i_\ell\}. \end{aligned} \tag{44.9}$$

In particular,

$$\frac{1}{2}\, dD^{+}_{1/2}(e_{2j}e_{2j-1})1 = \frac{1}{2}, \quad \forall\, j \in \{1, \dots, k\}, \tag{44.10}$$

and, for $1 \le i, j \le k$,

$$\begin{aligned} \frac{1}{2}\, dD^{-}_{1/2}(e_{2j}e_{2j-1})v_i = \;& \frac{1}{2} v_i, \quad j \ne i, \\ & -\frac{1}{2} v_i, \quad j = i. \end{aligned} \tag{44.11}$$

These calculations prove the following:

Proposition 44.1. *The representation $D^{+}_{1/2}$ of* Spin$(2k)$ *on $S_+(2k)$ has highest weight vector $1 \in \Lambda^0 \mathbb{C}^k$, and its highest weight is given by the k-tuple*

$$\Big(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2}\Big). \tag{44.12}$$

The representation $D^{-}_{1/2}$ of Spin$(2k)$ *on $S_-(2k)$ has highest weight vector $v_k \in \Lambda^1 \mathbb{C}^k$, and its highest weight is given by the k-tuple*

$$\Big(\frac{1}{2}, \dots, \frac{1}{2}, -\frac{1}{2}\Big). \tag{44.13}$$

Remark. If the ordering on $\tilde{\mathfrak{h}}'$ is determined by the choice of ordered basis $\{(1/2)e_{2j-1}e_{2j} : 1 \le j \le k\}$ instead of (44.1), the weight vectors stay the same, but the signs of the weights all change, and hence the identity of the highest weights changes. Therefore the statement of Proposition 44.1 gets modified, in ways the reader is invited to sort out.

We turn to the case $n = 2k-1$. Then we replace the basis (44.1) by

$$\Big\{\frac{1}{2}e_{2j}e_{2j-1} : 1 \le j \le k-1\Big\}. \tag{44.14}$$

Recalling the description (43.24)–(43.26) 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 (44.3), for $1 \le i, j \le 2k-1$:

$$\begin{aligned} \gamma_{ij}(t) &= e^{t(e_i e_{2k} e_j e_{2k})} \\ &= e^{t(e_i e_j)} \\ &= (\cos t)1 + (\sin t)e_i e_j, \end{aligned} \tag{44.15}$$

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 (44.3), and the calculations (44.4)–(44.10) need essentially no further changes. We have

$$dD^+_{1/2}(e_{2j}e_{2j-1}) = iQ_j, \quad 1 \le j \le k-1, \tag{44.16}$$

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 (44.5). We thus have the following counterpart to Proposition 44.1.

Proposition 44.2. *The representation $D^+_{1/2}$ of $\mathrm{Spin}(2k-1)$ on $S_+(2k)$ has highest weight vector $1 \in \Lambda^0\mathbb{C}^k$, and its highest weight is given by the $(k-1)$-tuple*

$$\Big(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2}\Big). \tag{44.17}$$

A. Flows and vector fields

Let $U \subset \mathbb{R}^n$ be open. A vector field on U is a smooth map

$$X : U \longrightarrow \mathbb{R}^n. \tag{A.1}$$

Consider the corresponding ODE

$$\frac{dy}{dt} = X(y), \quad y(0) = x, \tag{A.2}$$

with $x \in U$. A curve $y(t)$ solving (A.2) is called an integral curve of the vector field X. It is also called an *orbit*. For fixed t, write

$$y = y(t, x) = \mathcal{F}_X^t(x). \tag{A.3}$$

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:

$$\mathcal{L}_X f(x) = \lim_{h \to 0} h^{-1} \big[f(\mathcal{F}_X^h x) - f(x) \big] = \frac{d}{dt} f(\mathcal{F}_X^t x)\big|_{t=0}. \tag{A.4}$$

We also use the common notation

$$\mathcal{L}_X f(x) = Xf, \tag{A.5}$$

that is, we apply X to f as a first order differential operator.

Note that, if we apply the chain rule to (A.4) and use (A.2), we have

$$\mathcal{L}_X f(x) = X(x) \cdot \nabla f(x) = \sum a_j(x) \frac{\partial f}{\partial x_j}, \tag{A.6}$$

if $X = \sum a_j(x) e_j$, with $\{e_j\}$ the standard basis of $\mathbb{R}^n$. In particular, using the notation (A.5), we have

$$a_j(x) = X x_j. \tag{A.7}$$

In the notation (A.5),

$$X = \sum a_j(x) \frac{\partial}{\partial x_j}. \tag{A.8}$$

We note that X is a *derivation*, i.e., a map on $C^\infty(U)$, linear over $\mathbb{R}$, satisfying

$$X(fg) = (Xf)g + f(Xg). \tag{A.9}$$

Conversely, any derivation on $C^\infty(U)$ defines a vector field, i.e., has the form (A.8), as we now show.

Proposition A.1. *If X is a derivation on $C^\infty(U)$, then X has the form (A.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.

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.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 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 identify a point $x' \in \mathbb{R}^{n-1}$ near the origin with $x' \in M$. We can define a map

$$\mathcal{F} : M \times (-t_0, t_0) \longrightarrow U \tag{A.10}$$

by

$$\mathcal{F}(x', t) = \mathcal{F}_X^t(x'). \tag{A.11}$$

This is C^∞ and has surjective derivative, so by the Inverse Function Theorem is a local diffeomorphism. This defines a new coordinate system near p, in whch the flow generated by X has the form

$$\mathcal{F}_X^s(x', t) = (x', t + s). \tag{A.12}$$

If we denote the new coordinates by $(u_1, \dots, u_n)$, we see that the following result is established.

Theorem A.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*

$$X = \frac{\partial}{\partial u_n}. \tag{A.13}$$

B. Lie brackets

If $F : V \to W$ is a diffeomorphism between two open domains in $\mathbb{R}^n$, or between two smooth manifolds, and Y is a vector field on W, we define a vector field $F_\# Y$ on V so that

$$\mathcal{F}^t_{F_\# Y} = F^{-1} \circ \mathcal{F}^t_Y \circ F, \tag{B.1}$$

or equivalently, by the chain rule,

$$F_\# Y(x) = \big(DF^{-1}\big)\big(F(x)\big)Y\big(F(x)\big). \tag{B.2}$$

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:

$$\mathcal{L}_X Y = \lim_{h\to 0}\ h^{-1}\big(\mathcal{F}^h_\# Y - Y\big) = \frac{d}{dt}\mathcal{F}^t_\# Y\big|_{t=0}, \tag{B.3}$$

as a vector field on U.

Another natural construction is the operator-theoretic bracket:

$$[X, Y] = XY - YX, \tag{B.4}$$

where the vector fields X and Y are regarded as first order differential operators on $C^\infty(U)$. One verifies that (B.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 B.1. *If X and Y are smooth vector fields, then*

$$\mathcal{L}_X Y = [X, Y]. \tag{B.5}$$

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 (B.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 (B.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.

Corollary B.2. *If X and Y are smooth vector fields on U, then*

$$\frac{d}{dt}\mathcal{F}^t_{X\#}Y = \mathcal{F}^t_{X\#}[X,Y] \tag{B.6}$$

for all t.

Proof. Since locally $\mathcal{F}^{t+s}_X = \mathcal{F}^s_X\mathcal{F}^t_X$, we have the same identity for $\mathcal{F}^{t+s}_{X\#}$, which yields (B.6) upon taking the s-derivative.

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

$$[X,Y] = 0. \tag{B.7}$$

By (B.6), this implies $\mathcal{F}^t_{Y\#}X = X$ 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

$$Y = \frac{\partial}{\partial u}. \tag{B.8}$$

In this coordinate system, write $X = a(u,v)\partial/\partial u + b(u,v)\partial/\partial v$. The condition (B.7) implies $\partial a/\partial u = 0 = \partial b/\partial u$, so in fact we have

$$X = a(v)\frac{\partial}{\partial u} + b(v)\frac{\partial}{\partial v}. \tag{B.9}$$

Integral curves of (B.9) satisfy

$$u' = a(v), \quad v' = b(v) \tag{B.10}$$

and can be found explicitly in terms of integrals; one has

$$\int b(v)^{-1}\,dv = t + C_1, \tag{B.11}$$

and then

$$u = \int a(v(t))\,dt + C_2. \tag{B.12}$$

More generally than (B.7), we can suppose that, for some constant c,

$$[X,Y] = cX, \tag{B.13}$$

which by (B.6) is the same as

$$\mathcal{F}^t_{Y\#}X = e^{-ct}X. \tag{B.14}$$

An example would be

$$X = f(x,y)\frac{\partial}{\partial x} + g(x,y)\frac{\partial}{\partial y}, \tag{B.15}$$

where f and g satisfy "homogeneity" conditions of the form

$$f(r^a x, r^b y) = r^{a-c} f(x,y), \quad g(r^a x, r^b y) = r^{b-c} g(x,y), \tag{B.16}$$

for $r > 0$; in such a case one can take explicitly

$$\mathcal{F}^t_Y(x,y) = (e^{at}x, e^{bt}y). \tag{B.17}$$

Now, if one again has (B.8) in a local coordinate system (u,v), then X must have the form

$$X = e^{cu}\Big[a(v)\frac{\partial}{\partial u} + b(v)\frac{\partial}{\partial v}\Big] \tag{B.18}$$

which can be explicitly integrated, since

$$u' = e^{cu}a(v),\ v' = e^{cu}b(v) \Longrightarrow \frac{du}{dv} = \frac{a(v)}{b(v)}. \tag{B.19}$$

The hypothesis (B.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.

C. Frobenius' theorem

Let $G : U \to V$ be a diffeomorphism. Recall from §B the action on vector fields:

$$G_\# Y(x) = DG(y)^{-1} Y(y), \quad y = G(x). \tag{C.1}$$

As noted there, an alternative characterization of $G_\# Y$ is given in terms of the flow it generates. One has

$$\mathcal{F}^t_Y \circ G = G \circ \mathcal{F}^t_{G_\# Y}. \tag{C.2}$$

The proof of this is a direct consequence of the chain rule. As a special case, we have the following

Proposition C.1. *If $G_\# Y = Y$, then $\mathcal{F}^t_Y \circ G = G \circ \mathcal{F}^t_Y$.*

From this, we derive the following condition for a pair of flows to commute. Let X and Y be vector fields on U.

Proposition C.2. *If X and Y commute as differential operators, i.e.,*

$$[X, Y] = 0, \tag{C.3}$$

then locally $\mathcal{F}^s_X$ and $\mathcal{F}^t_Y$ commute, i.e., for any $p_0 \in U$, there exists $\delta > 0$ such that, for $|s|, |t| < \delta$,

$$\mathcal{F}^s_X \mathcal{F}^t_Y p_0 = \mathcal{F}^t_Y \mathcal{F}^s_X p_0. \tag{C.4}$$

Proof. By Proposition C.1, it suffices to show that $\mathcal{F}^s_{X\#} Y = Y$. Clearly this holds at $s = 0$. But by (B.6), we have

$$\frac{d}{ds} \mathcal{F}^s_{X\#} Y = \mathcal{F}^s_{X\#} [X, Y],$$

which vanishes if (C.3) holds. This finishes the proof.

We have stated that, given (C.3), then (C.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 (C.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 (C.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 C.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

$$F(t_1, \dots, t_k) = \mathcal{F}^{t_1}_{X_1} \cdots \mathcal{F}^{t_k}_{X_k} x_0. \tag{C.5}$$

Clearly $(\partial/\partial t_1)F = X_1(F)$. Similarly, since $\mathcal{F}^{t_j}_{X_j}$ all commute, we can put any $\mathcal{F}^{t_j}_{X_j}$ 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.

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 $b^{j\ell}_m(x)$ such that

$$[X_j, X_\ell] = \sum_{m=1}^{k} b^{j\ell}_m(x) X_m. \tag{C.6}$$

The following is Frobenius' Theorem.

Theorem C.4. *If $(X_1, \dots, X_k)$ are C^∞ vector fields on U, linearly independent at each point, and the involutivity condition (C.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

$$\text{span } (X_1, \dots, X_k) = \text{ span } \Big(\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_k}\Big). \tag{C.7}$$

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

$$Y_j = X_j - (X_j u_1)\frac{\partial}{\partial u_1} \quad \text{for} \quad 1 \le j \le k, \quad Y_{k+1} = \frac{\partial}{\partial u_1}. \tag{C.8}$$

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

$$Z = Y_{k+1} - \sum_{\ell=1}^{k} \big(Y_{k+1}y_\ell\big)\frac{\partial}{\partial y_\ell} = \sum_{\ell>k} (Y_{k+1}y_\ell)\frac{\partial}{\partial y_\ell}. \tag{C.9}$$

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 (C.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 C.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.

Second proof. Let $X_1, \dots, X_k$ be C^∞ vector fields, linearly independent at each point, and satisfying the condition (C.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

$$\Phi(t_1, \dots, t_k, x) = \mathcal{F}^{t_1}_{X_1} \cdots \mathcal{F}^{t_k}_{X_k} x. \tag{C.10}$$

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 $\Phi(0, \dots, t_j, \dots, 0, x) = \mathcal{F}^{t_j}_{X_j} x$, we have

$$\frac{\partial}{\partial t_j}\Phi(0, \dots, 0, x) = X_j(x), \quad x \in \mathcal{O}. \tag{C.11}$$

To establish the claim, it suffices to show that $\mathcal{F}^t_{X_j\#} X_\ell$ is a linear combination with coefficients in $C^\infty(U)$ of $X_1, \dots, X_k$. This is accomplished by the following:

Lemma C.5. *Suppose* $[Y, X_j] = \sum_\ell \lambda_{j\ell}(x) X_\ell$, *with smooth coefficients* $\lambda_{j\ell}(x)$. *Then* $\mathcal{F}^t_{Y\#} X_j$ *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 $\Lambda(t) = \Lambda(t, x) = (\lambda_{j\ell}(\mathcal{F}^t_Y x))$. Now let $A(t) = A(t, x)$ be the unique solution to the ODE

$$\frac{d}{dt}A(t) = \Lambda(t)A(t), \quad A(0) = I. \tag{C.12}$$

Write $A = (\alpha_{j\ell})$. We claim that

$$\mathcal{F}^t_{Y\#}X_j = \sum_\ell \alpha_{j\ell}(t,x)X_\ell. \tag{C.13}$$

This formula will prove the lemma. Indeed, we have

$$\begin{aligned}\frac{d}{dt}(\mathcal{F}^t_Y)_\# X_j &= (\mathcal{F}^t_Y)_\#[Y, X_j] \\ &= (\mathcal{F}^t_Y)_\# \sum_\ell \lambda_{j\ell}X_\ell \\ &= \sum_\ell (\lambda_{j\ell} \circ \mathcal{F}^t_Y)(\mathcal{F}^t_{Y\#}X_\ell).\end{aligned}$$

Uniqueness of the solution to (C.12) gives (C.13), and we are done.

This completes the second proof of Frobenius' Theorem.

D. 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

$$\frac{dy}{dt} = F(y), \quad y(0) = x. \tag{D.1}$$

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 (D.1) with respect to x suggests that $W = D_x y(t, x)$ satisfies an ODE called the *linearization* of (D.1):

$$\frac{dW}{dt} = DF(y)W, \quad W(0) = I. \tag{D.2}$$

In other words, $w(t, x) = D_x y(t, x)w_0$ satisfies

$$\frac{dw}{dt} = DF(y)w, \quad w(0) = w_0. \tag{D.3}$$

To justify this, we want to compare $w(t)$ and

$$z(t) = y_1(t) - y(t) = y(t, x + w_0) - y(t, x). \tag{D.4}$$

It would be convenient to show that z satisfies an ODE similar to (D.3). Indeed, $z(t)$ satisfies

$$\frac{dz}{dt} = F(y_1) - F(y) = \Phi(y_1, y)z, \quad z(0) = w_0, \tag{D.5}$$

where

$$\Phi(y_1, y) = \int_0^1 DF\big(\tau y_1 + (1 - \tau)y\big)\, d\tau. \tag{D.6}$$

If we assume

$$\|DF(u)\| \le M \text{ for } u \in U, \tag{D.7}$$

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

$$\|y_1(t) - y(t)\| \le e^{|t|M} \|w_0\|. \tag{D.8}$$

This establishes that $y(t,x)$ is *Lipschitz* in x.

To continue, since $\Phi(y,y) = DF(y)$, we rewrite (D.5) as

$$\frac{dz}{dt} = \Phi(y+z,y)z = DF(y)z + R(y,z), \quad w(0) = w_0. \tag{D.9}$$

where

$$F \in C^1(U) \Longrightarrow \|R(y,z)\| = o(\|z\|) = o(\|w_0\|). \tag{D.10}$$

Now comparing the ODE (6.9) with (6.3), we have

$$\frac{d}{dt}(z-w) = DF(y)(z-w) + R(y,z), \quad (z-w)(0) = 0. \tag{D.11}$$

Then Duhamel's principle yields

$$z(t) - w(t) = \int_0^t S(t,s)R\big(y(s), z(s)\big)\ ds, \tag{D.12}$$

so by the bound $\|S(t,s)\| \le e^{|t-s|M}$ and (6.10) we have

$$z(t) - w(t) = o(\|w_0\|). \tag{D.13}$$

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 (D.2). We state our first result.

Proposition D.1. *If $F \in C^1(U)$, and if solutions to (D.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 (D.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 (D.2) for $W(t,x)$ and $W(t,x+w_0)$, in the spirit of the analysis of (D.11).

If F possesses further smoothness, we can obtain higher differentiability of $y(t,x)$ in x by the following trick. *Couple* (6.1) and (6.2), to get an ODE for (y,W) :

$$\frac{dy}{dt} = F(y), \quad \frac{dW}{dt} = DF(y)W \tag{D.14}$$

with initial condition

$$y(0) = x, \quad W(0) = I. \tag{D.15}$$

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 D.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

$$\frac{dy}{dt} = F(\tau, y), \quad y(0) = x \tag{D.16}$$

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

$$\frac{dy}{dt} = F(z, y), \ \frac{dz}{dt} = 0; \quad y(0) = x, \ z(0) = \tau. \tag{D.17}$$

Thus we get smoothness of $y(t, \tau, x)$ in (τ, x). Furthermore, $v(t, \tau, x) = \partial_\tau y(t, \tau, x)$ satisfies

$$\frac{dv}{dt} = D_y F(\tau, y) v + F_\tau(\tau, y), \quad v(0, \tau, x) = 0. \tag{D.18}$$

Note that (2.15) (with t replaced by s and other notational changes) is a special case of this.

E. Lie algebras of matrix groups

Here we present some results on the Lie algebra of a Lie subgroup

$$G \subset \mathrm{Gl}(n, \mathbb{R}). \tag{E.1}$$

These results are special cases of more general and systematic results established in §§12–14. The advantage of the presentation here is that it can be read right after §2.

As a linear space, the Lie algebra of G can be identified as

$$\mathfrak{g} = T_I G. \tag{E.2}$$

One fundamental result is that $\mathrm{Exp} : \mathrm{M}(n, \mathbb{R}) \to \mathrm{Gl}(n, \mathbb{R})$, given by $\mathrm{Exp}(A) = e^A$, satisfies the following:

Proposition E.1. *For $A \in \mathrm{M}(n, \mathbb{R})$, we have*

$$A \in \mathfrak{g} \Longleftrightarrow e^{tA} \in G, \quad \forall\, t \in \mathbb{R}. \tag{E.3}$$

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 2.1 that this follows by inspection for G of the form (1.7)–(1.10), in which case $\mathfrak{g}$ is given by (2.10). We will postpone the proof of Proposition E.1 in the general case until later in this section.

Using (E.3), we establish the following:

Proposition E.2. *With $[A, B] = AB - BA$, we have*

$$A, B \in \mathfrak{g} \Longrightarrow [A, B] \in \mathfrak{g}. \tag{E.4}$$

Proof. Given $g \in G,\ A \in \mathfrak{g}$,

$$g^{-1}e^{tA}g = e^{tg^{-1}Ag}, \quad \forall\, t, \tag{E.5}$$

and the left side of (6.5) belongs to G, so by (E.3) we have

$$g^{-1}Ag \in \mathfrak{g}, \quad \forall\, g \in G,\ A \in \mathfrak{g}. \tag{E.6}$$

Setting $g = e^{tB},\ B \in \mathfrak{g}$, we have

$$e^{-tB}Ae^{tB} \in \mathfrak{g}, \quad \forall\, A, B \in \mathfrak{g}. \tag{E.7}$$

Applying d/dt at $t = 0$ gives (E.4).

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

$$\begin{aligned} e^{tA}e^{sB} &= \Bigl(I + tA + \frac{t^2}{2}A^2 + O(t^3)\Bigr)\Bigl(I + sB + \frac{s^2}{2}B^2 + O(s^3)\Bigr) \\ &= I + tA + sB + stAB + \frac{t^2}{2}A^2 + \frac{s^2}{2}B^2 + O(|(s,t)|^3), \end{aligned} \tag{E.8}$$

and similarly

$$e^{sB}e^{tA} = I + tA + sB + stBA + \frac{t^2}{2}A^2 + \frac{s^2}{2}B^2 + O(|(s,t)|^3), \tag{E.9}$$

Hence

$$e^{tA}e^{sB} = e^{sB}e^{tA} + st[A, B] + O(|(s,t)|^3). \tag{E.10}$$

We apply these calculations to show how the Lie algebra structure is preserved under representations of G. Thus, assume we have a smooth homomorphism

$$\pi : G \longrightarrow \mathrm{Gl}(n, \mathbb{R}). \tag{E.11}$$

(It is shown in Proposition 11.10 that every continuous homomorphism of G into $\mathrm{Gl}(n, \mathbb{R})$ is actually smooth.) Let us set

$$\sigma = D\pi(I) : \mathfrak{g} \longrightarrow \mathrm{M}(n, \mathbb{R}), \quad \text{so} \quad \sigma(A) = \frac{d}{ds}\pi(e^{sA})\big|_{s=0}, \tag{E.12}$$

for $A \in \mathfrak{g}$. Note that for such A,

$$\begin{aligned} \frac{d}{dt}\pi(e^{tA}) &= \frac{d}{ds}\pi(e^{(s+t)A})\big|_{s=0} \\ &= \frac{d}{ds}\pi(e^{sA})\pi(e^{tA})\big|_{s=0} \\ &= \sigma(A)\pi(e^{tA}), \end{aligned} \tag{E.13}$$

and since $\gamma(t) = \pi(e^{tA})$ satisfies $\gamma(0) = I$, this gives

$$\pi(e^{tA}) = e^{t\sigma(A)}. \tag{E.14}$$

We are ready to prove:

Proposition E.3. *For π, σ as in (E.11)–(E.12), $A, B \in \mathfrak{g}$, we have*

$$\sigma([A,B]) = [\sigma(A), \sigma(B)] = \sigma(A)\sigma(B) - \sigma(B)\sigma(A). \tag{E.15}$$

Proof. Setting $s = t$ in (E.10), we have

$$e^{tA}e^{tB}e^{-tA}e^{-tB} = I + t^2[A,B] + O(t^3). \tag{E.16}$$

Applying π, and noting that

$$\pi(e^{tA}) = e^{t\sigma(A)}, \quad \pi(e^{tB}) = e^{t\sigma(B)}, \tag{E.17}$$

we have

$$\pi(e^{tA}e^{tB}e^{-tA}e^{-tB}) = \pi(I + t^2[A,B] + O(t^3)) \tag{E.18}$$

equal to

$$e^{t\sigma(A)}e^{t\sigma(B)}e^{-t\sigma(A)}e^{-t\sigma(B)} = I + t^2[\sigma(A), \sigma(B)] + O(t^3), \tag{E.19}$$

the last identity holding by (E.16), with A, B replaced by $\sigma(A), \sigma(B)$. Comparing the right sides of (E.18) and (E.19), we see that

$$D\pi(I)[A,B] = [\sigma(A), \sigma(B)], \tag{E.20}$$

which gives (E.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

$$X_A(g) = gA, \tag{E.21}$$

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

$$L_h g = hg, \tag{E.22}$$

then we have

$$X_A(hg) = DL_h(g) X_A(g). \tag{E.23}$$

We now have the following simple result:

Proposition E.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

$$DL_g(I) : T_I G \longrightarrow T_g G, \tag{E.24}$$

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

$$\mathcal{F}_A^t g = g e^{tA}, \tag{E.25}$$

as is readily checked:

$$\begin{aligned} \frac{d}{dt}\mathcal{F}_A^t g\big|_{t=0} &= X_A(g) \quad \text{(by definition)} \\ &= gA, \end{aligned} \tag{E.26}$$

which coincides with $(d/dt)ge^{tA}|_{t=0}$. With these observations, we can give a short

Proof of Proposition E.1 (the "⇒" part). The flow $\mathcal{F}_A^t$ generated by X_A preserves each smooth submanifold of $\mathrm{Gl}(n,\mathbb{R})$ to which X_A is tangent. If $A \in \mathfrak{g}$, then G has this property, and $e^{tA} = \mathcal{F}_A^t I$.

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 $A \in \mathrm{M}(n,\mathbb{R})$,

$$\begin{aligned} X_A u(g) &= \frac{d}{dt} u(ge^{tA})\big|_{t=0} \\ &= Du(g) \cdot gA, \end{aligned} \tag{E.27}$$

where the "dot product" gives the pairing between $Du(g) \in \mathcal{L}(\mathrm{M}(n,\mathbb{R}),\mathbb{R}) = \mathrm{M}(n,\mathbb{R})'$ and $gA \in \mathrm{M}(n,\mathbb{R})$. Recall the Lie bracket of vector fields is given by

$$[X_A, X_B] = X_A X_B - X_B X_A. \tag{E.28}$$

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 §12.

Proposition E.5. *Given $A, B \in \mathrm{M}(n,\mathbb{R})$, we have*

$$[X_A, X_B] = X_{[A,B]}. \tag{E.29}$$

Proof. To begin, we have

$$X_A X_B u(g) = \frac{\partial^2}{\partial s \partial t} u(ge^{tA}e^{sB})\big|_{s,t=0}, \tag{E.30}$$

and hence

$$(X_A X_B - X_B X_A)u(g) = \frac{\partial^2}{\partial s \partial t}\Big[u(ge^{tA}e^{sB}) - u(ge^{sB}e^{tA})\Big]\Big|_{s,t=0}. \tag{E.31}$$

Recalling (E.10), we see that

$$\begin{aligned} u(ge^{tA}e^{sB}) &= u(ge^{sB}e^{tA} + stg[A,B] + O(|(s,t)|^3)) \\ &= u(ge^{sB}e^{tA}) + stDu(ge^{sB}e^{tA}) \cdot g[A,B] + O(|(s,t)|^3). \end{aligned} \tag{E.32}$$

Applying $(\partial^2/\partial s\partial t)|_{s,t=0}$, we obtain

$$\begin{aligned} [X_A, X_B]u(g) &= Du(g) \cdot g[A,B] \\ &= [X_A, X_B]u(g), \end{aligned} \tag{E.33}$$

the last identity holding by (E.27). This proves (E.29).

F. 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 §17. We can also form the space

$$\mathcal{P}(\mathfrak{g}) = \{P : \mathfrak{g}' \to \mathbb{C},\ \text{polynomial}\}. \tag{F.1}$$

There is a natural linear map

$$\beta : \mathcal{P}(\mathfrak{g}) \longrightarrow \mathfrak{U}(\mathfrak{g}), \tag{F.2}$$

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

$$\beta(X_1^{\alpha_1} \cdots X_n^{\alpha_n}) = (X_1 \otimes \cdots \otimes X_1) \otimes \cdots \otimes (X_n \otimes \cdots \otimes X_n), \quad \text{mod } J, \tag{F.3}$$

where on the right side of (F.3) we have α_j factors of $X_j \otimes \cdots \otimes X_j$. The Poincaré-Birkhoff-Witt theorem is the following:

Theorem F.1. *The map β in (F.2) is a linear isomorphism.*

To prove that β is surjective, we note that $\mathfrak{U}(\mathfrak{g})$ is spanned by monomials

$$X_{j_1} \otimes X_{j_2} \otimes \cdots \otimes X_{j_k} \quad (\text{mod } J). \tag{F.4}$$

The assertion that β is surjective is equivalent to the assertion that $\mathfrak{U}(\mathfrak{g})$ is actually spanned by monomials of the form (F.4) satisfying

$$j_1 \le j_2 \le \cdots \le j_k. \tag{F.5}$$

To see this, consider for example a monomial of the form (F.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 (F.5) and the second of which has lower order. A finite iteration rewrites each monomial (F.4) as a linear combination of monomials satisfying this order criterion (mod J), showing that β is surjective.

To complete the proof of Theorem F.1, it remains to show that β is injective. To do this, we bring in another linear map:

$$\alpha : \mathfrak{U}(\mathfrak{g}) \longrightarrow \mathcal{D}_L(G), \tag{F.6}$$

where $\mathcal{D}_L(G)$ is the space of left-invariant differential operators on G. This is defined by

$$\alpha(X_{j_1}\otimes\cdots\otimes X_{j_k}) = X_{j_1}\cdots X_{j_k}, \tag{F.7}$$

where the right side is the product of first order differential operators $X_{j_1},\dots,X_{j_k}$. Now

$$\alpha(X\otimes Y - Y\otimes X - [X,Y]) = XY - YX - [X,Y] = 0, \tag{F.8}$$

so α annihilates J, and hence (F.6) is well defined. Furthermore, we can compose α and β to get $\gamma = \alpha\circ\beta : \mathcal{P}(\mathfrak{g}) \to \mathcal{D}_L(G)$:

$$\begin{array}{ccc} \mathcal{P}(\mathfrak{g}) & \xrightarrow{\beta} & \mathfrak{U}(\mathfrak{g}) \\ & \gamma \searrow & \downarrow \alpha \\ & & \mathcal{D}_L(G) \end{array} \tag{F.9}$$

The formula looks tautological:

$$\gamma(X_1^{\alpha_1}\cdots X_n^{\alpha_n}) = X_1^{\alpha_1}\cdots X_n^{\alpha_n}, \tag{F.10}$$

but note that $X_1^{\alpha_1}\cdots X_n^{\alpha_n}$ on the left side of (F.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 (F.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 F.1, it suffices to prove:

Lemma F.2. *The map γ in (F.9) is injective.*

Proof. The injectivity of γ is equivalent to the following assertion. Assume

$$\sum_{|\alpha|\le k} C_\alpha X_1^{\alpha_1}\cdots X_n^{\alpha_n} = 0 \quad\text{in}\quad \mathcal{D}_L(G). \tag{F.11}$$

Then we assert that

$$C_\alpha = 0, \quad \forall\,\alpha. \tag{F.12}$$

To see this, use coordinates $(x_1,\dots,x_n)$ on $\mathfrak{g}$,

$$X = x_1X_1 + \cdots x_nX_n, \tag{F.13}$$

and use the map $\mathrm{Exp} : \mathfrak{g} \to G$, 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

$$X_j = \sum_{\ell=1}^{n} A_{j\ell}(x)\frac{\partial}{\partial x_\ell}. \tag{F.14}$$

Note that

$$A_{j\ell}(0) = \delta_{j\ell}. \tag{F.15}$$

The hypothesis (F.11) implies that

$$\sum_{|\alpha|\le k} C_\alpha \Big(\sum_\ell A_{1\ell}(x)\frac{\partial}{\partial x_\ell}\Big)^{\alpha_1} \cdots \Big(\sum_\ell A_{n\ell}(x)\frac{\partial}{\partial x_\ell}\Big)^{\alpha_n} = 0. \tag{F.16}$$

Now the left side of (F.16) is a differential operator of order k:

$$\sum_{|\alpha|\le k} \widetilde{C}_\alpha(x)\partial_1^{\alpha_1}\cdots\partial_n^{\alpha_n}, \tag{F.17}$$

(with $\partial_j = \partial/\partial x_j$), and from (F.14)–(F.15) we obtain

$$\widetilde{C}_\alpha(0) = C_\alpha, \quad \forall\, |\alpha| = k. \tag{F.18}$$

Now if (F.14) holds, then $\widetilde{C}_\alpha \equiv 0$ for all α, so we deduce that $C_\alpha = 0$ whenever $|\alpha| = k$. Recording this in (F.11) then gives

$$\sum_{|\alpha|\le k-1} C_\alpha X_1^{\alpha_1}\cdots X_n^{\alpha_n} = 0 \ \text{ in } \ \mathcal{D}_L(G), \tag{F.19}$$

and iterating this argument finishes the proof of Lemma F.2.

Remark. It is also the case that α and γ in (F.6) and (F.9) are linear isomorphisms. This follows from the results proven above plus the result that

$$\alpha \ \text{ is surjective.} \tag{F.20}$$

For a proof of (F.20), see [T1], 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.

G. Analytic continuation from U(n) to Gl($n,\mathbb{C}$), another approach

The following result proved useful in the analysis of the irreducible representations of U(n) in §19.

Theorem G.1. *If π is a representation of* U(n) *on a finite dimensional complex vector space V, then π extends to a holomorphic representation of* Gl($n,\mathbb{C}$) *on V.*

This was proven in §22. The proof given there extended $d\pi$, $\mathbb{C}$-linearly, from the Lie algebra u(n) of U(n) to the Lie algebra M($n,\mathbb{C}$) of Gl($n,\mathbb{C}$). Then it was shown that this Lie algebra representation arose from a representation of 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 G.1, deriving it from Proposition G.2, which is of independent interest.

To set up Proposition G.2, we define the representation $T^{p.q}$ of U(n) on $T^{p,q}(\mathbb{C}^n) = (\otimes^p\mathbb{C}^n) \otimes (\otimes^q\mathbb{C}^n)$ by

$$
\begin{aligned}
&T^{p,q}(g)v_1 \otimes \cdots \otimes v_p \otimes w_1 \otimes \cdots \otimes w_q \\
&= gv_1 \otimes \cdots \otimes gv_p \otimes (g^{-1})^t w_1 \otimes \cdots \otimes (g^{-1})^t w_q.
\end{aligned} \tag{G.1}
$$

Note that

$$
g \in \mathrm{U}(n) \Longleftrightarrow (g^{-1})^t = \overline{g}. \tag{G.2}
$$

Next, we define the representation T_K of U(n) on $T_K(\mathbb{C}^n) = \oplus_{p+q\le K} T^{p,q}(\mathbb{C}^n)$ by

$$
T_K(g)\Big(\bigoplus_{p+q\le K} v_{pq}\Big) = \bigoplus_{p+q\le K} T^{p,q}(g)v_{pq}. \tag{G.3}
$$

The following result is closely related to the "easy" proof of the Peter-Weyl theorem for compact matrix groups given in §7.

Proposition G.2. *If π is a finite dimensional representation of* U(n) *on V, then there exists $K < \infty$ such that π is contained in T_K.*

The content of Proposition G.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

$$
\pi(g) = J^{-1}T_K(g)J, \tag{G.4}
$$

for all $g \in$ U(n). Given this result, Theorem G.1 has the following simple proof. The formula (G.1) is clearly well defined for $g \in$ Gl($n,\mathbb{C}$), holomorphic in g, and this formula together with (G.3) provides an explicit extension of T_K from $\mathrm{U}(n)$ to $\mathrm{Gl}(n,\mathbb{C})$. In turn, (G.4) extends π from $\mathrm{U}(n)$ to $\mathrm{Gl}(n,\mathbb{C})$.

To prove Proposition G.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 (G.1)–(G.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 G.2.

H. 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 G.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 11.8, and is apparent for the standard examples. Thus we have

$$\rho : G \longrightarrow \mathrm{U}(n) \subset \mathrm{Gl}(n, \mathbb{C}), \tag{H.1}$$

taking G isomorphically onto its image G^ρ, with

$$d\rho : \mathfrak{g} \longrightarrow \mathrm{u}(n) \subset \mathrm{M}(n, \mathbb{C}), \tag{H.2}$$

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 $\mathfrak{g}^\rho_{\mathbb{C}} \subset \mathrm{M}(n, \mathbb{C})$. 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 G.1.

Theorem H.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*

$$\pi^\rho \circ \rho(g) = \pi(g), \quad \forall\, g \in G. \tag{H.3}$$

The proof will parallel that of Theorem G.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

$$\begin{aligned} &T^{p,q}_\rho(g) v_1 \otimes \cdots \otimes v_p \otimes w_1 \otimes \cdots \otimes w_q \\ &= \rho(g) v_1 \otimes \cdots \otimes \rho(g) v_p \otimes \rho(g^{-1})^t w_1 \otimes \cdots \otimes \rho(g^{-1})^t w_q. \end{aligned} \tag{H.4}$$

Note that

$$\rho(g) \in \mathrm{U}(n) \Longrightarrow \rho(g^{-1})^t = \overline{\rho(g)}. \tag{H.5}$$

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

$$T_{K,\rho}(g) \Big(\bigoplus_{p+q \le K} v_{pq} \Big) = \bigoplus_{p+q \le K} T^{p,q}_\rho(g) v_{pq}. \tag{H.6}$$

Then we have:

Proposition H.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 G.2.

The content of Proposition H.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

$$\pi(g) = J^{-1} T_{K,\rho}(g) J, \tag{H.7}$$

for all $g \in G$. Note that, with T_K as in (G.3)–(G.4), we have

$$\pi(g) = J^{-1} T_K(\rho(g)) J, \quad \forall\, g \in G. \tag{H.8}$$

Theorem H.1 follows easily from this. As noted in Appendix G, T_K extends from $\mathrm{U}(n)$ to $\mathrm{Gl}(n, \mathbb{C})$, holomorphically, and we obtain (H.3) with

$$\pi^\rho(\tilde{g}) = J^{-1} T_K(\tilde{g}) J, \quad \tilde{g} \in G^\rho_{\mathbb{C}}. \tag{H.9}$$

We next establish uniqueness:

Proposition H.3. *If σ is another faithful unitary representation of G, on $\mathbb{C}^m$, we have a natural holomorphic isomorphism*

$$G^\rho_{\mathbb{C}} \approx G^\sigma_{\mathbb{C}}. \tag{H.10}$$

Proof. Applying Theorem H.1 to $\pi = \sigma$, we have a holomorphic representation σ^ρ of $G^\rho_{\mathbb{C}}$ on $\mathbb{C}^m$, i.e.,

$$\sigma^\rho : G^\rho_{\mathbb{C}} \longrightarrow \mathrm{Gl}(m, \mathbb{C}), \tag{H.11}$$

such that

$$\sigma^\rho \circ \rho(g) = \sigma(g), \quad \forall\, g \in G. \tag{H.12}$$

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

$$\sigma^\rho : G^\rho_{\mathbb{C}} \longrightarrow G^\sigma_{\mathbb{C}}, \tag{H.13}$$

with $d\sigma^\rho : \mathfrak{g}^\rho_{\mathbb{C}} \to \mathfrak{g}^\sigma_{\mathbb{C}}$, isomorphically. Interchanging the roles of ρ and σ, we have a holomorphic homomorphism

$$\rho^\sigma : G^\sigma_{\mathbb{C}} \longrightarrow G^\rho_{\mathbb{C}}. \tag{H.14}$$

It readily follows that the maps in (H.13) and (H.14) are inverse to each other, so we have (H.10).

In light of this uniqueness, we choose any such $G^\rho_{\mathbb{C}}$ as constructed above, denote it $G_{\mathbb{C}}$, and call it "the complexification" of G.

I. 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

$$\Lambda^* V = \bigotimes{}^* V / \mathcal{I}, \tag{I.1}$$

where $\otimes^* V = \mathbb{R} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots$ is the tensor algebra and

$$\mathcal{I} \text{ is the 2-sided ideal generated by } \{u \otimes v + v \otimes u : u, v \in V\}. \tag{I.2}$$

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

$$u, v \in V \Longrightarrow u \wedge v = -v \wedge u. \tag{I.3}$$

We see that

$$\Lambda^* V = \bigoplus_{k=0}^{n} \Lambda^k V, \tag{I.4}$$

where $\Lambda^0 V = \mathbb{R}$ (or $\mathbb{C}$) and $\Lambda^k V$ is spanned by

$$\{e_{j_1} \wedge \cdots \wedge e_{j_k} : 1 \le j_1 < \cdots < j_k \le n\}. \tag{I.5}$$

Our goal in this appendix is to prove the following result, which was used in §41; cf. (41.14).

Proposition I.1. *The set (I.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

$$e_1 \wedge \cdots \wedge e_n \ne 0, \tag{I.6}$$

or, equivalently,

$$e_1 \wedge \cdots \wedge e_n \ne -e_1 \wedge \cdots \wedge e_n. \tag{I.7}$$

Note from (I.3) that if $\sigma \in S_n$, i.e., σ is a permutation of $\{1, \dots, n\}$, then

$$e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} = (\operatorname{sgn} \sigma)\, \sigma_1 \wedge \cdots \wedge e_n. \tag{I.8}$$

The content of (I.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 $\mathbb{R}^n \approx V$:

$$R(\sigma)f(x_1,\dots,x_n) = f(x_{\sigma(1)},\dots,x_{\sigma(n)}), \tag{I.9}$$

and note that $R(\sigma)$ leaves invariant

$$W = \operatorname{Span} P(x), \quad P(x) = \prod_{1\le j<k\le n} (x_j - x_k). \tag{I.10}$$

Furthermore,

$$R(\sigma)P(x) = (\operatorname{sgn}\sigma)\, P(x), \tag{I.11}$$

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 (I.5) with $k = \ell$, suppose

$$\sum_{1\le j_1<\cdots<j_\ell\le n} a_{j_1\cdots j_\ell}\, e_{j_1} \wedge \cdots \wedge e_{j_\ell} = 0. \tag{I.12}$$

Then, for each $m \in \{1,\dots,n\}$, wedge this with e_m on the left to get

$$\sum_{m\notin\{j_1,\dots,j_\ell\}} a_{j_1\cdots j_\ell}\, e_m \wedge e_{j_1} \wedge \cdots \wedge e_{j_\ell} = 0. \tag{I.13}$$

One can reorder $(m, j_1, \dots, j_\ell)$ to express (I.13) as a linear combination of monomials of the form (I.5) with $k = \ell + 1$. The inductive hypothesis yields

$$a_{j_1\cdots j_\ell} = 0, \tag{I.14}$$

for all multi-indices $(j_1,\dots,j_\ell)$ not containing m, for each m, hence (I.14) holds for all multi-indices $(j_1,\dots,j_\ell)$. This completes the inductive argument.

J. Simplicity of $\mathrm{M}(n, \mathbb{F})$

The following result (in the case $\mathbb{F} = \mathbb{C}$) was useful in the proof of Proposition 43.1.

Proposition J.1. *If $\mathbb{F}$ is a field, then the associative algebra $\mathrm{M}(v, \mathbb{F})$ 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 $\mathbb{F}^n$,

$$E_{jk} e_\ell = \delta_{k\ell} e_j. \tag{J.1}$$

A calculation gives

$$E_{j\ell} A E_{mk} = a_{\ell m} E_{jk}, \tag{J.2}$$

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 $\mathcal{I} = \mathrm{M}(n, \mathbb{F})$.

K. Two-step nilpotent Lie algebras

Every 2-step nilpotent Lie algebra $\mathfrak{n}$ has the form

$$\mathfrak{n} = V \oplus \mathfrak{z}, \tag{K.1}$$

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

$$A : V \times V \longrightarrow \mathfrak{z}. \tag{K.2}$$

Namely,

$$[(X_1, Z_1), (X_2, Z_2)] = (0, A(X_1, X_2)), \quad X_j \in V,\ Z_j \in \mathfrak{z}. \tag{K.3}$$

A structure equivalent to (K.2) is $A : \Lambda^2 V \to \mathfrak{z}$; another equivalent structure is

$$A' : \mathfrak{z}' \longrightarrow (\Lambda^2 V)' \approx \Lambda^2 V'. \tag{K.4}$$

Inner products on $\mathfrak{z}$ and on V produce isomorphisms $\mathfrak{z}' \approx \mathfrak{z}$ and $\Lambda^2 V' \approx \text{Skew}(V)$, the space of skew-adjoint linear operators on V, and hence the structure (K.4) is equivalent to

$$j : \mathfrak{z} \longrightarrow \text{Skew}(V), \tag{K.5}$$

related to A by

$$\langle j(Z)X, Y \rangle = \langle A(X, Y), Z \rangle, \tag{K.6}$$

for $X, Y \in V,\ Z \in \mathfrak{z}$, where the left side of (K.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

$$A(X, Y) = 0 \ \ \forall\, Y \in V \Longrightarrow X = 0, \tag{K.7}$$

or equivalently, if we have inner products on V and $\mathfrak{z}$ and use (K.5) to define the Lie algebra structure,

$$\bigcap_{Z \in \mathfrak{z}} \ker j(Z) = 0 \subset V. \tag{K.8}$$

Example 1. The Heisenberg Lie algebra $\mathcal{H}^n$ has the form (K.1)–(K.3) with

$$V = T^*\mathbb{R}^n \approx \mathbb{R}^{2n}, \quad \mathfrak{z} = \mathbb{R}, \tag{K.9}$$

and A the symplectic form on V (specified below). For $X_j = (x_j, y_j)^t \in V$, we have

$$A(X_1, X_2) = x_1 \cdot y_2 - x_2 \cdot y_1, \quad \langle X_1, X_2 \rangle = x_1 \cdot x_2 + y_1 \cdot y_2, \tag{K.10}$$

and hence

$$j(Z) = ZJ, \quad J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}, \quad Z \in \mathbb{R}. \tag{K.11}$$

Generalizing Example 1, one says $\mathfrak{n}$ is of Heisenberg type if it is defined by the structures (K.1) and (K.5), with

$$j(Z)^2 = -|Z|^2 I, \quad \forall\, Z \in \mathfrak{z}. \tag{K.12}$$

This is equivalent to requiring

$$j(Z)j(W) + j(W)j(Z) = -2\langle Z, W \rangle\, I, \quad \forall\, Z, W \in \mathfrak{z}. \tag{K.13}$$

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

$$j(Z) = \sigma_D(Z) \tag{K.14}$$

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 (K.1)–(K.3). The set of Lie algebras so produced is hence parametrized by $\mathcal{L}(\Lambda^2 V, \mathfrak{z})$. The condition (K.7) that the center of $\mathfrak{n}_A$ be exactly $\mathfrak{z}$ is that A belong to

$$\mathcal{N}^0(V, \mathfrak{z}) = \{A \in \mathcal{L}(\Lambda^2 V, \mathfrak{z}) : A(X, Y) = 0\ \forall\, Y \in V \Rightarrow X = 0\}. \tag{K.15}$$

We examine when A_1 and $A_2 \in \mathcal{N}^0(V, \mathfrak{z})$ yield isomorphic Lie algebras:

$$T : \mathfrak{n}_{A_1} \xrightarrow{\approx} \mathfrak{n}_{A_2}. \tag{K.16}$$

Since T must preserve the common center $\mathfrak{z}$, we see that T must have the form

$$T(X, Z) = (QX, RX + SZ), \quad Q \in \mathrm{Gl}(V),\ S \in \mathrm{Gl}(\mathfrak{z}),\ R \in \mathcal{L}(V, \mathfrak{z}). \tag{K.17}$$

The condition that such T be a Lie algebra isomorphism is

$$A_2(QX, QY) = SA_1(X, Y), \quad \forall\, X, Y \in V, \tag{K.18}$$

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 (K.1) and either (K.2) or (K.5). Suppose a group G has representations π on V and ρ on $\mathfrak{z}$. If

$$A(\pi(g)X, \pi(g)Y) = \rho(g)A(X, Y), \quad \forall\, X, Y \in V, \quad g \in G, \tag{K.19}$$

it is clear that

$$(X, Z) \mapsto (\pi(g)X, \rho(g)Z) \tag{K.20}$$

yields a Lie algebra automorphism of $\mathfrak{n} = V \oplus \mathfrak{z}$. (This specializes (K.18) to $A_1 = A_2 = A$.) If we assume V and $\mathfrak{z}$ have inner products, via which we pass from (K.2) to (K.5), and the operators $\pi(g)$ and $\rho(g)$ preserve these inner products, then the hypothesis

$$j(\rho(g)Z) = \pi(g)j(Z)\pi(g)^{-1}, \quad \forall\, g \in G,\ Z \in \mathfrak{z} \tag{K.21}$$

readily yields (K.19), displaying the action of G as a group of automorphisms of $\mathfrak{n}$ in (K.20). If we do not assume $\pi(g)$ and $\rho(g)$ preserve these inner products, replace (K.21) by

$$j(\rho(g)^t Z) = \pi(g)^t j(Z)\pi(g), \quad \forall g \in G,\ Z \in \mathfrak{z}. \tag{K.22}$$

Example 2. Dilations. If we take $G = \mathbb{R}$ and $\pi(t)X = e^t X$, $\rho(t)Z = e^{2t}Z$ in (K.20), it is clear that (K.19) holds. Thus each two-step nilpotent Lie algebra $\mathfrak{n} = V \oplus \mathfrak{z}$ has the group of dilations

$$\delta(t)(X, Z) = (e^t X, e^{2t} Z) \tag{K.23}$$

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 (K.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

$$j = d\pi : \mathfrak{z} \longrightarrow \operatorname{Skew}(V), \tag{K.24}$$

to define a 2-step nilpotent Lie algebra $\mathfrak{n}$. We have (K.21) with

$$\rho(g)Z = (\operatorname{Ad} g)Z, \tag{K.25}$$

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:

$$G = \mathrm{SO}(p, q), \quad V = \mathbb{R}^{p,q}. \tag{K.26}$$

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 $\rho = \operatorname{Ad}$, as in (K.25). We take $V = \mathbb{R}^n$ and let π be the standard representation of $\mathrm{SO}(n)$ on $\mathbb{R}^n$. Then $\mathfrak{g} \approx \operatorname{Skew}(\mathbb{R}^n)$, and we have

$$\pi(g)X = gX, \quad \rho(g)Z = gZg^t, \quad X \in \mathbb{R}^n,\ Z \in \operatorname{Skew}(\mathbb{R}^n), \tag{K.27}$$

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 (K.23), which in this setting is tautological:

$$j(Z) = Z, \quad Z \in \mathfrak{g} \approx \operatorname{Skew}(\mathbb{R}^n). \tag{K.28}$$

Example 4 specializes to yield $\mathrm{SO}(n)$ acting on $\mathfrak{n}$ as a group of automorphisms. We claim this enlarges to

$$\mathrm{Gl}(n,\mathbb{R}) \longrightarrow \mathrm{Aut}\,\mathfrak{n}, \tag{K.29}$$

given by (K.20), where $\pi(g)$ and $\rho(g)$ are again defined by (K.27), for $g \in \mathrm{Gl}(n,\mathbb{R})$. To verify (K.22), note that

$$j(\rho(g)^t Z) = g^t Z g \quad \text{and} \quad \pi(g)^t j(Z)\pi(g) = g^t Z g, \tag{K.30}$$

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 (K.29) contains both the $\mathrm{SO}(n)$ action and the group $\delta(t)$ of dilations from Example 2.

L. More on quaternions and Sp(n)

With $\alpha_\nu, \beta_\nu \in \mathbb{C}$, set

$$\xi = \alpha_1 + j\alpha_2, \quad \eta = \beta_1 + j\beta_2. \tag{L.1}$$

We have a $\mathbb{C}$-linear isomorphism

$$\kappa : \mathbb{H} \longrightarrow \mathbb{C}^2, \quad \kappa(\xi) = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \tag{L.2}$$

where $\mathbb{C}$ acts on $\mathbb{H}$ on the *right.* We set

$$\operatorname{Co} \xi = \alpha_1, \quad \operatorname{Sp} \xi = \alpha_2. \tag{L.3}$$

Note that

$$\begin{aligned} \overline{\eta}\xi &= (\overline{\beta}_1 - \overline{\beta}_2 j)(\alpha_1 + j\alpha_2) \\ &= \alpha_1\overline{\beta}_1 + \alpha_2\overline{\beta}_2 - j(\alpha_1\beta_2 - \alpha_2\beta_1) \end{aligned} \tag{L.4}$$

Hence

$$\begin{aligned} \operatorname{Co}(\overline{\eta}\xi) &= \alpha_1\overline{\beta}_1 + \alpha_2\overline{\beta}_2 \\ &= ((\kappa(\xi), \kappa(\eta))), \end{aligned} \tag{L.5}$$

where $((\ ,\ ))$ denotes the standard Hermitian inner product on $\mathbb{C}^2$, and

$$\begin{aligned} \operatorname{Sp}(\overline{\eta}\xi) &= -(\alpha_1\beta_2 - \alpha_2\beta_1) \\ &= -\sigma(\kappa(\xi), \kappa(\eta)), \end{aligned} \tag{L.6}$$

where $\sigma(\ ,\ )$ is the standard ($\mathbb{C}$-bilinear) symplectic form on $\mathbb{C}^2$.

Moving on to $\mathbb{H}^n$, we have in analogy to (L.2) a $\mathbb{C}$-linear isomorphism

$$\kappa : \mathbb{H}^n \longrightarrow \mathbb{C}^{2n}, \tag{L.7}$$

where $\mathbb{C}$ acts on $\mathbb{H}^n$ on the right, and the $\mathbb{H}$-valued inner product $\langle\ ,\ \rangle$, defined in (3.30), satisfies

$$\begin{aligned} \operatorname{Co}\langle \xi, \eta \rangle &= ((\kappa(\xi), \kappa(\eta))), \\ \operatorname{Sp}\langle \xi, \eta \rangle &= -\sigma(\kappa(\xi), \kappa(\eta)), \end{aligned} \tag{L.8}$$

where $((\ ,\ ))$ is the Hermitian inner product on $\mathbb{C}^{2n}$ and $\sigma(\ ,\ )$ the symplectic form. It follows that κ produces an isomorphism

$$\operatorname{Sp}(n) \approx U(2n) \cap \operatorname{Sp}(2n, \mathbb{C}). \tag{L.9}$$

References

[AI] D. Altschuler and C. Itzykson, Remarks on integration over Lie algebras, Ann. Inst. Henri Poincaré 54 (1991), 1–8.

[BI] P. Bleher and A. Its (eds.), Random Matrix Models and Their Applications, MSRI Publications #40, Cambridge Univ. Press, 2001.

[D] P. Deift, Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach, Courant Lecture Notes #3, AMS, Providence, R.I., 1998.

[DE] P. Diaconis and S. Evans, Linear functionals of eigenvalues of random matrices, Trans. AMS 353 (2001), 2615–2633.

[DS] P. Diaconis and M. Shahshahani, On the eigenvalues of random matrices, Trans. AMS 353 (2001), 2615–2633.

[DK] J.J. Duistermaat and J. Kolk, Lie Groups, Springer-Verlag, New York, 2000.

[EM] 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.

[FH] W. Fulton and J. Harris, Representation Theory, a First Course, Springer-Verlag, New York, 1991.

[GW] R. Goodman and N. Wallach, Representations and invariants of the classical groups, Cambridge Univ. Press, 1998.

[HS] M. Hausner and J. Schwartz, Lie Groups; Lie Algebras, Gordon and Breach, London, 1968.

[Hel] S. Helgason, Groups and Geometric Analysis, Academic Press, New York, 1984.

[How] R. Howe, Remarks on classical invariant theory, Trans. AMS 313 (1989), 539–570.

[IZ] C. Itzykson and J. Zuber, The planar approximation, II, J. Math. Phys. 21 (1980), 411–421.

[LM] H. B. Lawson and M.-L. Michelson, Spin Geometry, Princeton Univ. Press, Princeton, N.J., 1989.

[M] M. Mehta, Random Matrices, Academic Press, New York, 1991.

[Sak] S. Sakai, C^*-algebras and W^*-algebras, Springer-Verlag, New York, 1971.

[Si] B. Simon, Representations of finite and compact groups, AMS, Providence, R.I., 1996.

[Str] R. Strichartz, The Campbell-Hausdorff formula and DuHamel's principle, Preprint.

[T1] M. Taylor, Noncommutative Harmonic Analysis, AMS, Providence, R.I., 1986.

[T2] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer-Verlag, New York, 1996.

[V] V.S. Varadarajan, Lie groups, Lie algebras, and their representations, Prentice-

Hall, Englewood Cliffs, N.J., 1974.

[Wal] N. Wallach, Harmonic Analysis on Homogeneous Spaces, Marcel Dekker, New York, 1973.

[Z] D. Zelobenko, Compact Lie Groups and their Representations, Transl. Math. Monogr., Vol. 40, AMS, Providence, R.I., 1973.