Chapter 1

Linear groups

We begin, as we shall end, with the classical groups—those familiar groups of matrices encountered in every branch of mathematics. At the outset, they serve as a library of linear groups, with which to illustrate our theory. Later we shall find that these same groups also serve as the building-blocks for the theory.

Definition 1.1 *A linear group is a closed subgroup of* $\mathbf{GL}(n, \mathbb{R})$.

Remarks:

1. We could equally well say that: *A linear group is a closed subgroup of* $\mathbf{GL}(n, \mathbb{C})$. For as we shall see shortly, $\mathbf{GL}(n, \mathbb{C})$ has an isomorphic image as a closed subgroup of $\mathbf{GL}(2n, \mathbb{R})$; while conversely it is clear that $\mathbf{GL}(n, \mathbb{R})$ can be regarded as a closed subgroup of $\mathbf{GL}(n, \mathbb{C})$.

2. By $\mathbf{GL}(n, k)$ (where $k = \mathbb{R}$ or $\mathbb{C}$) we mean the group of invertible $n \times n$ matrices, ie

$$\boxed{\mathbf{GL}(n, k) = \{T \in \mathbf{M}(n, k) : \det T \neq 0\}}$$

   where $\mathbf{M}(n, k)$ denotes the space of all $n \times n$ real or complex matrices (according as $k = \mathbb{R}$ or $\mathbb{C}$).

3. We define a norm $\|X\|$ on $\mathbf{M}(n, k)$ by

$$\|X\|^2 = \sum_{i,j} |X_{ij}|^2 = \begin{cases} \mathrm{tr}(X'X) & \text{if } k = \mathbb{R}, \\ \mathrm{tr}(X^*X) & \text{if } k = \mathbb{C}, \end{cases}$$

   where as usual X' denotes the transpose and X^* the conjugate transpose. This is just the usual Euclidean norm, if we identify $\mathbf{M}(n, k)$ with k^N, where

$N = n^2$, by taking as the coordinates of the matrix X its n^2 entries X_{ij}. The requirement that a linear group should be closed in $\mathbf{GL}(n, k)$ refers to this metric topology. In other words, if $T(i)$ is a sequence of matrices in G tending towards the matrix $T \in \mathbf{GL}(n, k)$, ie

$$\|T(i) - T\| \to 0,$$

then T must in fact lie in G.

Examples:

1. The *general linear group* $\mathbf{GL}(n, \mathbb{R})$
2. The *special linear group*

$$\boxed{\mathbf{SL}(n, \mathbb{R}) = \{T \in \mathbf{GL}(n, \mathbb{R}) : \det(T) = 1\}}$$

3. The *orthogonal group*

$$\boxed{\mathbf{O}(n) = \{T \in \mathbf{GL}(n, \mathbb{R}) : T'T = I\}}$$

In other words

$$\mathbf{O}(n) = \{T : Q(Tv) = Q(v) \quad \forall v \in \mathbb{R}^n\},$$

where

$$Q(v) = x_1^2 + \ldots + x_n^2 \qquad (v = (x_1, \ldots, x_n) \in \mathbb{R}^n),$$

ie $\mathbf{O}(n)$ is the subgroup of $\mathbf{GL}(n, \mathbb{R})$ leaving the quadratic form Q invariant.

4. The *special orthogonal group*

$$\boxed{\mathbf{SO}(n) = \mathbf{O}(n) \cap \mathbf{SL}(n)}$$

5. The *complex general linear group* $\mathbf{GL}(n, \mathbb{C})$. This calls for some explanation, since $\mathbf{GL}(n, \mathbb{C})$ is not a group of real matrices, as required by Definition 1. However, we can represent each complex matrix $Z \in \mathbf{M}(n, \mathbb{C})$ by a real matrix $\mathbb{R}Z \in \mathbf{M}(2n, \mathbb{R})$ in the following way. (Compare the "realification" of a representation discussed in Part I Chapter 11.)

If we "forget" scalar multiplication by non-reals, the complex vector space $V = \mathbb{C}^n$ becomes a real vector space $\mathbb{R}V$ of twice the dimension, with basis

$$(1, 0, \ldots,), (i, 0, \ldots, 0), (0, 1, \ldots, 0), (0, i, \ldots, 0), \ldots, (0, 0, \ldots, 1), (0, 0, \ldots, i).$$

Moreover each matrix $Z \in \mathbf{M}(n, \mathbb{C})$, ie each linear map

$$Z : V \to V$$

defines a linear map

$$\mathbb{R}Z : \mathbb{R}V \to \mathbb{R}V,$$

ie a matrix $\mathbb{R}Z \in \mathbf{M}(2n, \mathbb{R})$.

Concretely, in passing from Z to $\mathbb{R}Z$ each entry

$$Z_{j,k} = X_{j,k} + iY_{j,k}$$

is replaced by the 2×2 matrix

$$\begin{pmatrix} X_{j,k} & -Y_{j,k} \\ Y_{j,k} & X_{j,k} \end{pmatrix}$$

The map

$$Z \mapsto \mathbb{R}Z : \mathbf{M}(n, \mathbb{C}) \to \mathbf{M}(2n, \mathbb{R})$$

is injective; and it preserves the algebraic structure, ie

- $\mathbb{R}(Z + W) = \mathbb{R}Z + \mathbb{R}W$
- $\mathbb{R}(ZW) = (\mathbb{R}Z)(\mathbb{R}W)$
- $\mathbb{R}(aZ) = a(\mathbb{R}Z) \quad \forall a \in \mathbb{R}$
- $\mathbb{R}I = I$
- $\mathbb{R}(Z^*) = (\mathbb{R}Z)'$.

It follows in particular that $\mathbb{R}Z$ is invertible if and only if Z is; so $\mathbb{R}$ restricts to a map

$$Z \mapsto \mathbb{R}Z : \mathbf{GL}(n, \mathbb{C}) \to \mathbf{GL}(2n, \mathbb{R}).$$

Whenever we speak of $\mathbf{GL}(n, \mathbb{C})$, or more generally of any group G of complex matrices, as a linear group, it is understood that we refer to the image $\mathbb{R}G$ of G under this injection $\mathbb{R}$.

The matrix $X \in \mathbf{GL}(2n, \mathbb{R})$ belongs to $\mathbf{GL}(n, \mathbb{C})$ if is built out of 2×2 matrices of the form

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

This can be expressed more neatly as follows. Let

$$iI \mapsto J = \begin{pmatrix} 0 & -1 & & & \\ 1 & 0 & & & \\ & & 0 & -1 & \\ & & 1 & 0 & \\ & & & & \ddots \end{pmatrix}$$

Since any scalar multiple of the identity commutes with all matrices,

$$X \in \mathbf{M}(n, \mathbb{C}) \implies (iI)X = X(iI).$$

Applying the operator $\mathbb{R}$,

$$X \in \mathbb{R}\mathbf{M}(n, \mathbb{C}) \implies JX = XJ.$$

Converseley, if $JX = XJ$ then it is readily verified that X *is* of the required form. Thus

$$\mathbb{R}\mathbf{M}(n, \mathbb{C}) = \{X \in \mathbf{M}(2n, \mathbb{R}) : JX = XJ\};$$

and in particular

$$\boxed{\mathbf{GL}(n, \mathbb{C}) = \{T \in \mathbf{GL}(2n, \mathbb{R}) : JX = XJ\}}$$

6. The *complex special linear group*

$$\boxed{\mathbf{SL}(n, \mathbb{C}) = \{T \in \mathbf{GL}(n, \mathbb{C}) : \det T = 1\}}$$

Note that the determinant here must be computed in $\mathbf{M}(n, \mathbb{C})$, not in $\mathbf{M}(2n, \mathbb{R})$. Thus

$$T = \begin{pmatrix} i \end{pmatrix} \notin \mathbf{SL}(1, \mathbb{C}),$$

although

$$\mathbb{R}T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbf{SL}(2, \mathbb{R}).$$

7. The *unitary group*

$$\boxed{\mathbf{U}(n) = \{T \in \mathbf{GL}(n, \mathbb{C}) : T^*T = I\}}$$

where T^* denotes the complex transpose of T. In other words

$$\mathbf{U}(n) = \{T : H(Tv) = H(v) \quad \forall v \in \mathbb{C}^n\},$$

where H is the Hermitian form

$$H(v) = |x_1|^2 + \ldots + |x_n|^2 \qquad (v = (x_1, \ldots, x_n) \in \mathbb{C}^n).$$

Since $\mathbb{R}(T^*) = (\mathbb{R}T)'$, a complex matrix is unitary if and only if its real counterpart is orthogonal:

$$\boxed{\mathbf{U}(n) = \mathbf{GL}(n, \mathbb{C}) \cap \mathbf{O}(2n)}$$

8. The *special unitary group*

$$\boxed{\mathbf{SU}(n) = \mathbf{U}(n) \cap \mathbf{SL}(n, \mathbb{C})}$$

9. The *quaternionic general linear group* $\mathbf{GL}(n, \mathbb{H})$. The quaternions

$$q = t + xi + yj + zk \quad (t, x, y, z \in \mathbb{R}),$$

with multiplication defined by

$$\boxed{i^2 = j^2 = k^2 = -1,\ jk = -kj = i,\ ki = -ik = j,\ ij = -ji = k}$$

form a *division algebra* or *skew field*. For the product of any quaternion

$$q = t + xi + yj + zk$$

with its *conjugate*

$$\bar{q} = t - xi - yj - zk$$

gives its *norm-square*

$$\|q\|^2 = \bar{q}q = t^2 + x^2 + y^2 + z^2,$$

so that if $q \neq 0$,

$$q^{-1} = \frac{\bar{q}}{\|q\|^2}.$$

We can equally well regard $\mathbb{H}$ as a 2-dimensional algebra over $\mathbb{C}$, with each quaternion taking the form

$$q = z + wj \quad (z, w \in \mathbb{C}),$$

and multiplication in $\mathbb{H}$ being defined by the rules

$$\boxed{jz = \bar{z}j,\ j^2 = -1}$$

Surprisingly perhaps, the entire apparatus of linear algebra extends almost unchanged to the case of a *non-commutative* scalar field. Thus we may speak of an n-dimensional vector space W over $\mathbb{H}$, of a linear map $t : W \to W$, etc.

A quaternionic vector space W defines a complex vector space $\mathbb{C}W$ by "forgetting" scalar multiplication by non-complex quaternions (ie those involving j or k), in just the same way as a complex vector space V defines a real vector space $\mathbb{R}V$. If W has quaternionic dimension n, with basis $\{e_1, e_2, \ldots, e_n\}$, then $\mathbb{C}W$ has complex dimension $2n$, with basis

$$\{e_1, je_1, e_2, je_2, \ldots, e_n, je_n\}.$$

Moreover each matrix $Q \in \mathbf{M}(n, \mathbb{H})$, ie each linear map

$$Q : W \to W$$

defines a linear map

$$\mathbb{C}Q : \mathbb{C}W \to \mathbb{C}W,$$

ie a matrix $\mathbb{C}Q \in \mathbf{M}(2n, \mathbb{C})$.

Concretely, in passing from W to $\mathbb{C}W$ each entry

$$Q_{r,s} = Z_{r,s} + iW_{r,s}$$

is replaced by the 2×2 complex matrix

$$\begin{pmatrix} Z_{r,s} & -W_{r,s} \\ \overline{W_{r,s}} & \overline{Z_{r,s}} \end{pmatrix}$$

The map

$$Q \mapsto \mathbb{C}Q : \mathbf{M}(n, \mathbb{H}) \to \mathbf{M}(2n, \mathbb{C})$$

is injective; and it preserves the algebraic structure, ie

- $\mathbb{C}(Q + Q') = \mathbb{C}Q + \mathbb{C}Q'$
- $\mathbb{C}(QQ') = (\mathbb{C}Q)(\mathbb{C}Q')$
- $\mathbb{C}(aQ) = a(\mathbb{C}Q) \quad \forall a \in \mathbb{C}$
- $\mathbb{C}I = I$
- $\mathbb{C}(Q^*) = (\mathbb{C}Q)^*$.

In this last relation, Q^* denotes the quaternionic matrix with entries

$$(Q^*)_{rs} = \overline{Q_{sr}}.$$

To identify $\mathbf{M}(n, \mathbb{H})$ as a subspace of $\mathbf{M}(2n, \mathbb{C})$, consider the automorphism of $\mathbb{H}$

$$q \mapsto \tilde{q} = jqj^{-1}.$$

In terms of its 2 complex components,

$$q = z + wj \mapsto \tilde{q} = \bar{z} + \bar{w}j.$$

(Note that $\tilde{q} \neq \bar{q}$; indeed, the map $q \mapsto \bar{q}$ is not an automorphism of $\mathbb{H}$, but an anti-automorphism.)

Let J denote the diagonal matrix

$$J = \begin{pmatrix} j & 0 & & \\ 0 & j & & \\ & & \ddots & \end{pmatrix} \in \mathbf{M}(n, \mathbb{H});$$

and consider the map

$$Q \mapsto \tilde{Q} = JQJ^{-1} : \mathbf{M}(n, \mathbb{H}) \to \mathbf{M}(n, \mathbb{H}).$$

We see from above that

$$\mathbb{C}(\tilde{Q}) = \overline{\mathbb{C}Q},$$

where $\bar{X}$ denotes (for $X \in \mathbf{M}(n, \mathbb{C})$) the matrix with entries

$$\bar{X}_{r,s} = \overline{X_{r,s}}.$$

Now

$$J \mapsto \mathbb{C}J = \begin{pmatrix} 0 & -1 & & & \\ 1 & 0 & & & \\ & & 0 & -1 & \\ & & 1 & 0 & \\ & & & & \ddots \end{pmatrix},$$

which we take the liberty of also denoting by J (as we did earlier, when defining the embedding of $\mathbf{GL}(n, \mathbb{C})$ in $\mathbf{GL}(2n, \mathbb{R})$, although there we regarded J as a real matrix rather than a complex one).

Thus

$$\mathbf{M}(n, \mathbb{H}) \subset \{X \in \mathbf{M}(2n, \mathbb{C}) : JXJ^{-1} = \bar{X}\}$$

Conversely, it is readily verified that if $JXJ^{-1} = \bar{X}$, then X is constructed from 2×2 matrices of the form specified above, and so arises from a quaternionic matrix. Hence

$$\boxed{\mathbf{M}(n, \mathbb{H}) = \{X \in \mathbf{M}(2n, \mathbb{C}) : JXJ^{-1} = \bar{X}\}}$$

It follows from the properties of the map $\mathbb{C}$ listed above that if Q is invertible then so is $\mathbb{C}Q$; so $\mathbb{C}$ restricts to a map

$$T \mapsto \mathbb{C}T : \mathbf{GL}(n, \mathbb{H}) \to \mathbf{GL}(2n, \mathbb{C}).$$

In fact, our argument above gives the concrete embedding

$$\boxed{\mathbf{GL}(n, \mathbb{H}) = \{T \in \mathbf{GL}(2n, \mathbb{C}) : JTJ^{-1} = \bar{T}\}}$$

10. The *symplectic group*

$$\boxed{\mathbf{Sp}(n) = \{T \in \mathbf{GL}(n, \mathbb{H}) : T^*T = I\}}$$

Since $(\mathbb{C}T)^* = \mathbb{C}(T^*)$ it follows that

$$T \in \mathbf{Sp}(n) \Longrightarrow (\mathbb{C}T)^*(\mathbb{C}T) = I,$$

and so

$$\boxed{\mathbf{Sp}(n) = \mathbf{GL}(n, \mathbb{H}) \cap \mathbf{U}(2n)}$$

Thus

$$\mathbf{Sp}(n) = \{T \in \mathbf{GL}(2n, \mathbb{C}) : JTJ^{-1} = \bar{T} \ \& \ TT^* = I\}.$$

Since $\bar{T}^{-1} = T'$ from the second relation, the first relation can be re-written

$$T'JT = J.$$

This gives an alternative description of the symplectic group:

$$\boxed{\mathbf{Sp}(n) = \{U \in \mathbf{U}(2n) : U'JU = J\}}$$

In other words, $\mathbf{Sp}(n)$ consists of those unitary matrices leaving invariant the skew-symmetric form

$$J(x, y) = x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3 + \ldots + x_{2n-1}y_{2n} - x_{2n}y_{2n-1}.$$

11. The *Lorentz group*

$$\boxed{\mathbf{O}(1,3) = \{T \in \mathbf{GL}(4,\mathbb{R}) : G(Tv) = G(v) \quad \forall v \in \mathbb{R}^4\}}$$

where G is the space-time metric

$$G(v) = t^2 - x^2 - y^2 - z^2.$$

In matrix terms

$$\mathbf{O}(1,3) = \{T : T'GT = G\},$$

where

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

Exercises

1. Prove that every linear group is *locally compact*. (Hint: Show that every open subset, and every closed subset, of a locally compact space is locally compact.)

2. Show that
$$\mathbf{SO}(2) \cong \mathbf{U}(1), \quad \mathbf{Sp}(1) \cong \mathbf{SU}(2).$$

3. Show that if G and H are linear groups then $G \times H$ is *linearisable*, ie there is a linear group isomorphic to $G \times H$.

4. Show that there is a discrete linear group isomorphic to any finite group G.

5. Show that there is a discrete linear group isomorphic to $\mathbb{Z}^n$.

6. Show that $\mathbb{R}$ is linearisable.

7. Show that $\mathbb{R}^*$ is linearisable. Show also that
$$\mathbb{R}^* \cong \mathbb{R} \times C_2$$
(where C_2 denotes the cyclic group of order 2).

8. Show that $\mathbb{C}$ is linearisable.

9. Show that $\mathbb{C}^*$ is linearisable. Show also that
$$\mathbb{C}^* \ncong \mathbb{C} \times G$$
for any topological group G.

Chapter 2

The Exponential Map

> Napier introduced logarithms to convert difficult multiplication into easy addition. Our motivation is much the same, though we are dealing with matrices rather than numbers. As in the numerical case, it is simpler to start with the exponential function—defined by an everywhere convergent matrix power-series—and derive the logarithmic function as the inverse in a suitable restricted zone.

Proposition 2.1 *For each matrix $X \in \mathbf{M}(n, k)$ (where $k = \mathbb{R}$ or $\mathbb{C}$) the exponential sequence*

$$I + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots$$

converges.

Proof ► In Chapter 1 we defined the norm $\|X\|$ on $\mathbf{M}(n, k)$ by

$$\|X\|^2 = \sum_{i,j} \|X_{ij}\|^2,$$

In other words,

$$\|X\|^2 = \begin{cases} \operatorname{tr}(X'X) & \text{if } k = \mathbb{R} \\ \operatorname{tr}(X^*X) & \text{if } k = \mathbb{C} \end{cases}$$

Lemma 2.1 1. $\|X + Y\| \le \|X\| + \|Y\|$

2. $\|XY\| \le \|X\|\|Y\|$

3. $\|aX\| = |a|\|X\|(a \in k)$

Proof of Lemma ▷ We suppose that $k = \mathbb{C}$; the real case is identical, with X' in place of X^*.

1. We know that

$$\operatorname{tr} Z^* Z = \|Z\|^2 \geq 0$$

for all $Z \in \mathbf{M}(n,k)$. Setting $Z = X + \lambda Y$ (where $\lambda \in \mathbb{R}$),

$$\operatorname{tr} X^* X + \lambda \left(\operatorname{tr} X^* Y + \operatorname{tr} Y^* X\right) + \lambda^2 \operatorname{tr} Y^* Y \geq 0$$

for all $\lambda \in \mathbb{R}$. Hence

$$|\operatorname{tr} X^* Y + \operatorname{tr} Y^* X|^2 \leq 4\|X\|^2 \|Y\|^2,$$

and so

$$|\operatorname{tr} X^* Y + \operatorname{tr} Y^* X| \leq 2\|X\| \|Y\|.$$

We note for future reference that *if X and Y are hermitian*, ie $X^* = X,\ Y^* = Y$, then $\operatorname{tr} X^* Y = \operatorname{tr} XY = \operatorname{tr} YX = \operatorname{tr} Y^* X$; and so

$$X, Y \text{ hermitian} \Longrightarrow \operatorname{tr} XY \leq \|X\| \|Y\|.$$

But now (taking $\lambda = 1$),

$$\begin{aligned} \|X + Y\|^2 &= \|X\|^2 + \operatorname{tr}(X^* Y + Y^* X) + \|Y\|^2 \\ &\leq \|X\|^2 + 2\|X\| \|Y\| + \|Y\|^2; \end{aligned}$$

whence

$$\|X + Y\| \leq \|X\| + \|Y\|.$$

2. We have

$$\begin{aligned} \|XY\|^2 &= \operatorname{tr}(XY)^* XY \\ &= \operatorname{tr} Y^* X^* XY \\ &= \operatorname{tr} X^* XYY^* \\ &= \operatorname{tr} PQ, \end{aligned}$$

where

$$P = X^* X, \quad Q = YY^*.$$

These 2 matrices are hermitian and positive-definite; and

$$\|X\|^2 = \operatorname{tr} P, \quad \|Y\|^2 = \operatorname{tr} Q.$$

Thus it is sufficient to show that for any 2 such matrices

$$\operatorname{tr} PQ \leq \operatorname{tr} P \operatorname{tr} Q.$$

But as we noted in the first part of the proof,

$$\operatorname{tr} PQ \leq \|P\|\|Q\|.$$

It is sufficient therefore to show that

$$\|P\| \leq \operatorname{tr} P.$$

for any positive hermition matrix X. Since

$$\|P\|^2 = \operatorname{tr} P^*P = \operatorname{tr}(P^2),$$

this is equivalent to proving that

$$\operatorname{tr}(P^2) \leq (\operatorname{tr} P)^2.$$

But if the eigenvalues of P are $\lambda_1, \ldots, \lambda_n$, then those of P^2 are $\lambda_1^2, \ldots, \lambda_n^2$, and

$$\operatorname{tr}(P^2) = \lambda_1^2 + \cdots + \lambda_n^2 \leq (\lambda_1 + \cdots + \lambda_n)^2 = (\operatorname{tr} P)^2.$$

since $\lambda_i \geq 0$ for all i.

For a 'tensorial' proof of this result, let S denote the 'skew-symmetrizer'

$$S_{jl}^{ik} = \frac{1}{2}\left(\delta_j^i \delta_l^k - \delta_l^i \delta_j^k\right).$$

under which

$$S(u \otimes v) = \frac{1}{2}(u \otimes v - v \otimes u);$$

and let

$$Z = S(X^* \otimes Y).$$

Then

$$Z^* = (X \otimes Y^*)S,$$

and so, since $S^2 = S$,

$$\begin{aligned}
\operatorname{tr} Z^*Z &= \operatorname{tr}(X \otimes Y^*)S(X^* \otimes Y) \\
&= \operatorname{tr} S(X^* \otimes Y)(X \otimes Y^*) \\
&= \operatorname{tr} S(X^*X \otimes YY^*) \\
&= \frac{1}{2}(\operatorname{tr} X^*X \operatorname{tr} YY^* - \operatorname{tr} X^*XYY^*) \\
&= \frac{1}{2}(\operatorname{tr} X^*X \operatorname{tr} Y^*Y - \operatorname{tr}(XY)^*XY) \\
&= \frac{1}{2}\left(\|X\|^2\|Y\|^2 - \|XY\|^2\right)
\end{aligned}$$

Since $\operatorname{tr} Z^*Z \geq 0$, we conclude that

$$\|XY\| \leq \|X\|\|Y\|.$$

3. We have

$$\begin{aligned}\|aX\|^2 &= \operatorname{tr} X^*\bar{a}aX\\ &= \bar{a}a \operatorname{tr} X^*X\\ &= |a|^2\|X\|^2.\end{aligned}$$

$\lhd$

To show that the exponential series converges for any matrix X, we compare its partial sums with those in the scalar case. By the lemma above,

$$\|X^i/i! + \ldots + X^j/j!\| \leq x^i/i! + \cdots + x^j/j!,$$

where $x = \|X\|$.

It follows that

$$\|X^i/i! + \ldots + X^j/j!\| \to 0$$

as $i, j \mapsto \infty$. Since every Cauchy sequence converges in $\mathbb{R}^N$, this proves the proposition. ◀

Definition 2.1 *For each matrix $X \in \mathbf{M}(n,k)$ (where $k = \mathbb{R}$ or $\mathbb{C}$) we set*

$$\boxed{e^X = I + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots}$$

Examples:

1. $e^0 = I$
2. If X is diagonal, say

$$X = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

then e^X is also diagonal:

$$e^X = \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix}$$

Proposition 2.2 *If $X, Y \in \mathbf{M}(n,k)$ commute, ie*

$$XY = YX,$$

then

$$e^X e^Y = e^{X+Y} = e^Y e^X.$$

Proof ▶ Since $XY = YX$, $(X+Y)^m$ can be expanded by the binomial theorem:

$$(X+Y)^m = X^m + mX^{m-1}Y + \ldots + Y^m.$$

The result follows on summation as in the ordinary (scalar) case

$$e^{x+y} = e^x e^y,$$

all convergences being absolute. ◀

Corollary 2.1 *For all $X \in \mathbf{M}(n,k)$,*

$$\boxed{e^X \in \mathbf{GL}(n,k),}$$

ie e^X is invertible, with

$$\boxed{(e^X)^{-1} = e^{-X}.}$$

Thus the exponential establishes a map

$$\mathbf{M}(n,k) \to \mathbf{GL}(n,k) : X \mapsto e^X.$$

Proposition 2.3 *For $T \in \mathbf{GL}(n,k)$, $X \in \mathbf{M}(n,k)$,*

$$\boxed{Te^XT^{-1} = e^{TXT^{-1}}.}$$

Proof ▶ For each m,

$$(TXT^{-1})^m = TX^mT^{-1}.$$

The result follows on summation. ◀

Proposition 2.4 *If the eigenvalues of X are*

$$\lambda_1, \ldots, \lambda_n$$

then the eigenvalues of e^X are

$$e^{\lambda_1}, \ldots, e^{\lambda_n}.$$

Proof ▶ As we saw in Chapter 1, each matrix $X \in \mathbf{M}(n,\mathbb{C})$ defines a matrix $\mathbb{R}X \in \mathbf{M}(2n,\mathbb{R})$. Conversely, each matrix $X \in \mathbf{M}(n,\mathbb{R})$ defines a matrix $\mathbb{C}X \in \mathbf{M}(n,\mathbb{C})$, namely the matrix with the same entries (now regarded as complex numbers).

Lemma 2.2 *1.* $\mathbb{C}(e^X) = e^{\mathbb{C}X}$ *for all $X \in \mathbf{M}(n,\mathbb{R})$*

2. $\mathbb{R}(e^X) = e^{\mathbb{R}X}$ *for all* $X \in \mathbf{M}(n, C)$

Proof of Lemma $\triangleright$

1. This is immediate, since the matrices arising on each side are identical, the only difference being that in one case they are regarded as real and in the other as complex.

2. This follows from
$$(\mathbb{R}X)^m = \mathbb{R}(X^m),$$
on summation.

$\triangleleft$

Proof of Proposition 4, continued: Since X and $\mathbb{C}X$ have the same eigenvalues, we can suppose X complex, by Part 1 of the Lemma.

- Although a complex matrix cannot in general be diagonalised, it can always be brought to triangular form:
$$TXT^{-1} = Y = \begin{pmatrix} \lambda_1 & a_{12} & \dots & a_{1n} \\ 0 & \lambda_2 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$
(This can be proved by induction on n. Taking an eigenvector of X as first element of a new basis,
$$SXS^{-1} = \begin{pmatrix} \lambda & b_{12} & \dots & b_{1n} \\ 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{pmatrix}$$
The result follows on applying the inductive hypothesis to the $(n-1) \times (n-1)$ matrix T_1.)

- But then
$$Y^m = \begin{pmatrix} \lambda_1^m & c_{12} & \dots & c_{1n} \\ 0 & \lambda_2^m & \dots & c_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^m \end{pmatrix}$$

and so on summation

$$e^Y = \begin{pmatrix} e_1^\lambda & w_{12} & \ldots & w_{1n} \\ 0 & e_2^\lambda & \ldots & w_{2n} \\ \vdots & & \ddots & \vdots \\ \ldots & & & \\ 0 & 0 & \ldots & e_n^\lambda \end{pmatrix}$$

The result now follows, since

$$e^Y = Te^X T^{-1}$$

by Proposition 3; so e^X and e^Y have the same eigenvalues.

◄

Corollary 2.2 *For each $X \in \mathbf{M}(n,k)$,*

$$\boxed{\det e^X = e^{\operatorname{tr} X}.}$$

Proof ► If the eigenvalues of X are $\lambda_1, \ldots, \lambda_n$ then

$$\begin{aligned} \det X = & \quad \lambda_1 \ldots \lambda_n \\ \operatorname{tr} X = & \quad \lambda_1 + \ldots + \lambda_n. \end{aligned}$$

Hence

$$\det e^X = e^{\lambda_1} \cdots e^{\lambda_n} = e^{\lambda_1 + \cdots + \lambda_n} = e^{\operatorname{tr} X}.$$

◄

Proposition 2.5 *1.* $e^{X'} = (e^X)'$ *for all* $X \in \mathbf{M}(n,\mathbb{R})$

2. $e^{X^*} = (e^X)^*$ *for all* $X \in \mathbf{M}(n,\mathbb{C})$

Proof ► For each m,

$$\begin{aligned} (X')^m &= (X^m)', \\ (X^*)^m &= (X^m)^*. \end{aligned}$$

The result follows on summation. ◄

We turn to the analytic properties of the exponential map.

Proposition 2.6 *There exists an open neighbourhood $U \ni 0$ in $\mathbf{M}(n,\mathbb{R})$ which is mapped homeomorphically by the map $X \mapsto e^X$ onto an open neighbourhood $V = e^U \ni I$ in $\mathbf{GL}(n,\mathbb{R})$.*

Lemma 2.3 *Suppose* $X \in \mathbf{M}(n, \mathbb{R})$ *satisfies*

$$\|X\| < 1.$$

Then

1. $T = I + X \in \mathbf{GL}(n, \mathbb{R})$.
2. *The series*
$$\log T = X - \frac{X^2}{2} + \frac{X^3}{3} - \dots$$
is convergent.
3. $e^{\log T} = T$.

Proof of Lemma $\triangleright$

1. Explicitly,
$$(I + X)^{-1} = I - X + X^2 - X^3 + \dots.$$
2. Convergence follows as in the scalar case, but with the matrix norm $\|X\|$ in place of the absolute value $|x|$.
3. We know that if $|x| < 1$ then
$$1 + x = e^{\log(1+x)} = 1 + (x - x^2/2 + \dots) + (x - x^2/2 + \dots)^2/2! + \dots.$$
Moreover the convergence on the right is absolute; and the identity therefore holds for any matrix X satisfying $\|X\| < 1$.

$\triangleleft$

Proof of Proposition ► Let

$$V = \{T \in \mathbf{GL}(n, \mathbb{R}) : \|T - I\| < 1\};$$

and let

$$U = \log V = \{\log T : T \in V\}.$$

Then it follows from Part 3 of the Lemma that the maps

$$T \mapsto \log T : V \to U \text{ and } X \mapsto e^X : U \to V$$

are mutually inverse. Since e^X and $\log T$ are continuous (as in the scalar case) and V is open, it follows that U is also open. ◄

Remark: We shall call the region U (and also occasionally its image V) the *logarithmic zone*, since for $T \in U$ the exponential has the well-defined inverse $\log T$:

$$\begin{aligned} e^{\log T} &= T \qquad \text{for all } T \in U \\ \log e^X &= X \qquad \text{for all } X \in V \end{aligned}$$

Proposition 2.7 *For each $X \in \mathbf{M}(n, \mathbb{R})$ the map*

$$t \mapsto e^{tX} : \mathbb{R} \to \mathbf{GL}(n, \mathbb{R})$$

is a continuous homomorphism; and every continuous homomorphism $\mathbb{R} \to \mathbf{GL}(n, \mathbb{R})$ is of this form.

Proof ► Since sX and tX commute, it follows from Proposition 3 that

$$e^{sX} e^{tX} = e^{(s+t)X}.$$

Hence the map $t \to e^{tX}$ is a homomorphism, which is clearly continuous.

Conversely, suppose

$$T : \mathbb{R} \to \mathbf{GL}(n, \mathbb{R})$$

is a continuous homomorphism. For sufficiently small t, say

$$t \in J = [-c, c],$$

$T(t)$ must lie in the logarithmic zone; and we can therefore set

$$X(t) = \log T(t)$$

for $t \in J$.

Since $T : \mathbb{R} \to \mathbf{GL}(n, \mathbb{R})$ is a homomorphism,

$$T(s)T(t) = T(s+t).$$

We want to convert this to the additive form

$$X(s) + X(t) = X(s+t)$$

by taking logarithms. To this end, note first that $T(s)$ and $T(t)$ commute:

$$T(s)T(t) = T(s+t) = T(t)T(s) \qquad \text{for all } s, t.$$

It follows that $X(s)$ and $X(t)$, as power-series in $T(s)$ and $T(t)$, also commute:

$$X(s)X(t) = X(t)X(s) \qquad \text{for all } s,t \in J.$$

So by Proposition 2, if s, t and $s+t$ all lie in J,

$$e^{X(s)+X(t)} = e^{X(s)}e^{X(t)} = T(s)T(t) = T(s+t) = e^{X(s+t)}.$$

Now if s and t are small enough, say

$$s,t \in J' = [-c', c'],$$

then not only $X(s)$, $X(t)$ and $X(s+t)$, but also $X(s)+X(t)$, will lie in the logarithmic zone V. In that case, on taking logarithms in the last relation,

$$X(s) + X(t) = X(s+t) \qquad \text{for all } s,t \in J'.$$

We want to deduce from this that

$$X(t) = tX$$

for some X. Note first that, on replacing t by $c't$ (and X by $c'^{-1}X$), we can suppose that $c' = 1$, ie that the last relation holds for

$$s,t \in I = [-1,1].$$

We have to show that

$$X(t) = tX(1) \qquad \text{for all } t \in I.$$

Suppose first that s is a positive integer. By repeated application of the basic identity,

$$sX(\frac{1}{s}) = X(\frac{1}{s}) + \ldots + X(\frac{1}{s}) = X(\frac{1}{s} + \ldots + \frac{1}{s}) = X(1),$$

ie

$$X(\frac{1}{s}) = \frac{1}{s}X(1).$$

Now suppose $0 \le r \le s$. Then

$$X(\frac{r}{s}) = X(\frac{1}{s}) + \ldots + X(\frac{1}{s}) = rX(\frac{1}{s}) = \frac{r}{s}X(1).$$

We have thus established the result for rationals $t = r/s \in [0,1]$. Since the rationals lie densely among the reals, it follows by continuity that

$$X(t) = tX(1) \qquad \text{for all } t \in [0,1].$$

The extension to the negative interval [-1,0] is immediate, since

$$X(0) + X(0) = X(0) \Longrightarrow X(0) = 0,$$

and so

$$X(t) + X(-t) = X(0) \Longrightarrow X(-t) = -X(t).$$

Returning to the given homomorphism,

$$T(t) = e^{X(t)} = e^{tX} \qquad \text{for all } t \in J'.$$

We can extend the range by taking powers:

$$T(nt) = T(t)^n = (e^{tX})^n = e^{ntX} \qquad \text{for all } n \in N, t \in J'.$$

Hence

$$T(t) = e^{tX} \qquad \text{for all } t \in \mathbb{R}.$$

Finally, the uniqueness of X follows on taking logarithms for sufficiently small t. ◄

Summary: The exponential map sets up a 1–1 correspondence—more precisely, a homeomorphism—between a neighbourhood U of 0 in $\mathbf{M}(n,\mathbb{R})$ and a neighbourhood V of I in $\mathbf{GL}(n,\mathbb{R})$. The inverse logarithmic function projects this 'polar region' V back into $\mathbf{M}(n,\mathbb{R})$—which we can identify with the tangent-space to $\mathbf{GL}(n,\mathbb{R})$ at I. The picture is reminiscent of Mercator's projection of the globe onto the pages of an atlas, or Riemann's projection of the sphere onto a plane.

Exercises

In Exercises 01–15 calculate e^X for the given matrix X:

1. $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ 2. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 3. $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ 4. $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ 5. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

6. $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ 7. $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ 8. $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ 9. $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ 10. $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

11. $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ 12. $\begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix}$ 13. $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$ 14. $\begin{pmatrix} a & -b \\ b & -a \end{pmatrix}$ 15. $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$

In Exercises 16–25 determine whether or not the given matrix is of the form e^X for some $X \in \mathbf{M}(2, \mathbb{R})$.

16. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 17. $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ 18. $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ 19. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ 20. $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

21. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 22. $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ 23. $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ 24. $\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ 25. $\begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}$

Chapter 3

[

The Lie Algebra of a Linear Group I]The Lie Algebra of a Linear Group I: The Underlying Space

> Suppose $G \subset \mathbf{GL}(n, \mathbb{R})$ is a linear group. The rays in $\mathbf{M}(n, \mathbb{R})$ corresponding to 1-parameter subgroups trapped inside G fill a vector subspace $\mathcal{L}G$. This correspondence between closed subgroups of $\mathbf{GL}(n, \mathbb{R})$ and certain subspaces of $\mathbf{M}(n, \mathbb{R})$ is the foundation of Lie theory.

Definition 3.1 *Suppose $G \subset \mathbf{GL}(n, \mathbb{R})$ is a linear group. Then we set*

$$\boxed{\mathcal{L}G = \{X \in \mathbf{M}(n, \mathbb{R}) : e^{tX} \in G \quad \textit{for all } t \in \mathbb{R}\}.}$$

Remark: In the case of the classical groups $\mathbf{GL}(n, \mathbb{R})$, $\mathbf{SO}(n, \mathbb{R})$, etc, considered in Chapter 1, it is customary to denote the corresponding space by the same letters in lower case, eg

$$\begin{aligned} \mathbf{o}(n) &= \mathcal{L}\mathbf{O}(n), \\ \mathbf{sl}(n, \mathbb{R}) &= \mathcal{L}\mathbf{SL}(n, \mathbb{R}), \\ \mathbf{sp}(n) &= \mathcal{L}\mathbf{Sp}(n). \end{aligned}$$

Proposition 3.1 *If $G \subset \mathbf{GL}(n, \mathbb{R})$ is a linear group then $\mathcal{L}G$ is a vector subspace of $\mathbf{M}(n, \mathbb{R})$.*

Proof ► We have to show that

1. $X \in \mathcal{L}G, a \in \mathbb{R} \Longrightarrow aX \in \mathcal{L}G$,

2. $X, Y \in \mathcal{L}G \Longrightarrow X + Y \in \mathcal{L}G$.

The first result follows at once from the definition of $\mathcal{L}G$. The second is a consequence of the following result.

Lemma 3.1 *Suppose $T(t) \in G$ for $0 < t \leq d$; and suppose*

$$T(t) = I + tX + o(t)$$

for some $X \in \mathbf{M}(n, \mathbb{R})$, ie

$$(T(t) - I)/t \to X \text{ as } t \to 0.$$

Then

$$X \in \mathcal{L}G.$$

Proof of Lemma ▷ We must show that

$$e^{tX} \in G \qquad \text{for all } t.$$

The argument can be divided into 5 steps.

1. Recall the formula for the scalar exponential as a limit:

$$(1 + \frac{x}{m})^m \to e^x \qquad \text{as } m \to \infty.$$

This is most simply proved by taking the logarithm of each side. On the left

$$\begin{aligned} \log\left((1 + \frac{x}{m})^m\right) &= m \log(1 + \frac{x}{m}) \\ &= m(\frac{x}{m} + o(\frac{1}{m})) \\ &= x + o(1). \end{aligned}$$

In other words

$$\log\left((1 + \frac{x}{m})^m\right) \to x \qquad \text{as } m \to \infty.$$

The result follows on taking the exponentials of both sides.

2. It is evident from this proof that we can replace $1 + x/m$ by any function $a(m)$ satisfying

$$a(m) = 1 + \frac{x}{m} + o(\frac{1}{m});$$

for any such function

$$a(m)^m \to e^x.$$

3. Both this last result and its proof carry over to the matrix case. If
$$A = I + \frac{X}{m} + o(\frac{1}{m}),$$
with $X \in \mathbf{M}(n, \mathbb{R})$, then
$$A(m)^m \to e^X$$

4. Applying this with
$$A(m) = T(\frac{t}{m})$$
we deduce that
$$T(\frac{t}{m})^m \to e^{tX}.$$

5. Since $T(t/m) \in G$ for sufficiently large m, and since G is a group,
$$T(\frac{t}{m})^m \in G.$$
Hence, since G is closed,
$$e^{tX} \in G.$$

◁

Remark: In geometric language this result shows that $\mathcal{L}G$ can be regarded as *the tangent-space to G at $T = I$.*

Proof of Proposition 1 (completion). Suppose $X, Y \in \mathcal{L}G$. Then
$$e^{tX} e^{tY} \in G \quad \text{for all } t.$$

But
$$\begin{aligned} e^{tX} e^{tY} &= (I + tX)(I + tY) + o(t) \\ &= I + t(X + Y) + o(t). \end{aligned}$$

Hence by the Lemma
$$X + Y \in \mathcal{L}G.$$

◀

Examples:

1. *The General Linear Group:* Since $e^{tX} \in \mathbf{GL}(n, \mathbb{R})$ for all $X \in \mathbf{M}(n, \mathbb{R})$, by Proposition 2.1,
$$\boxed{\mathbf{gl}(n, \mathbb{R}) = \mathbf{M}(n, \mathbb{R}).}$$

2. *The Special Linear Group:* $\mathbf{sl}(n,\mathbb{R})$ consists of all *trace-free matrices*, ie

$$\boxed{\mathbf{sl}(n,\mathbb{R}) = \{X \in \mathbf{M}(n,\mathbb{R}) : \operatorname{tr} X = 0\}.}$$

By definition

$$\begin{aligned} X \in \mathbf{sl}(n,\mathbb{R}) &\implies e^{tX} \in \mathbf{SL}(n,\mathbb{R}) \quad \text{for all } t \\ &\implies \det(e^{tX}) = 1 \quad \text{for all } t. \end{aligned}$$

To the first order in t,

$$e^{tX} = I + tX + o(t).$$

Hence

$$\begin{aligned} \det(e^{tX}) &= \det(I + tX) + o(t) \\ &= 1 + t \operatorname{tr} X + o(t); \end{aligned}$$

for on expanding the determinant the off-diagonal terms in tX will only appear in second or higher order terms. Hence

$$X \in \mathbf{sl}(n,\mathbb{R}) \implies \operatorname{tr} X = 0.$$

Conversely,

$$\begin{aligned} \operatorname{tr} X = 0 &\implies \det e^{tX} = e^{t \operatorname{tr} X} = 1 \\ &\implies e^{tX} \in \mathbf{SL}(n,\mathbb{R}). \end{aligned}$$

Remarks:

(a) A linear group G is said to be *algebraic* if it can be defined by polynomial conditions $F_k(T)$ on the matrix entries T_{ij}, say

$$G = \{T \in \mathbf{GL}(n,\mathbb{R}) : F_k(T) = 0,\ k = 1, \dots, N\}.$$

Not every linear group is algebraic; but all the ones we meet will be.

The technique above—working to the first order in t—is the recommended way of determining the Lie algebra $\mathcal{L}G$ of an ‘unknown’ linear group G. From above,

$$\begin{aligned} X \in \mathcal{L}G &\implies e^{tX} \in G \\ &\implies F_k(I + tX + \cdots) = 0 \\ &\implies F_k(I + tX) = O(t^2) \end{aligned}$$

Thus

$$\mathcal{L}G \subset \{X : F_k(I + X) = O(t^2)\}.$$

In theory it only gives us a necesssary condition for X to lie in $\mathcal{L}G$; and it is easy to think of artificial cases where the condition is not in fact sufficient. For example the equation

$$\operatorname{tr}(I - T)'(I - T) = 0$$

defines the trivial group $\{I\}$; but it is satisfied by e^{tX} to the first order in t for all X.

In practice the condition is usually sufficient. However it must be verified—having determined $\mathcal{L}G$ (as we hope) in this way, we must then show that e^X does in fact lie in G. (Since the condition on X is linear this will automatically ensure that $e^t X \in G$ for all $t \in \mathbb{R}$.)

(b) An alternative way of describing the technique is to say that since each defining condition is satisfied by e^{tX} identically in t, the differential of this condition must vanish at $t = 0$, eg

$$\mathbf{sl}(n, \mathbb{R}) \equiv \{X : \frac{d}{dt} \det e^{tX} = 0 \text{ at } t = 0\}.$$

(c) To summarise: Given a linear group G,

i. Find all X for which $I + tX$ satisfies the equations for G to the first order in t; and then

ii. Verify that $e^X \in G$ for these X.

3. *The Orthogonal Group:* $\mathbf{o}(n)$ consists of all *skew-symmetric matrices*, ie

$$\boxed{\mathbf{o}(n) = \{X : X' + X = 0\}.}$$

For

$$(I + tX)'(I + tX) = I + t(X' + X) + o(t).$$

Hence

$$X \in \mathbf{o}(n) \Longrightarrow X' + X = 0.$$

Conversely,

$$\begin{aligned} X' + X = 0 \quad &\Longrightarrow \quad X' = -X \\ &\Longrightarrow \quad (e^X)' = e^{X'} = e^{-X} = (e^X)^{-1} \\ &\Longrightarrow \quad e^X \in \mathbf{O}(n). \end{aligned}$$

4. *The Special Orthogonal Group:*

$$\boxed{\mathbf{so}(n) = \mathbf{o}(n).}$$

This follows from the trivial

Lemma 3.2 *If G is an intersection of linear groups, say*

$$G = \cap G_i,$$

then

$$\mathcal{L}G = \cap \mathcal{L}G_i.$$

Applying this to

$$\mathbf{SO}(n) = \mathbf{O}(n) \cap \mathbf{SL}(n, \mathbb{R})$$

we deduce that

$$\mathbf{so}(n) = \mathbf{o}(n) \cap \mathbf{sl}(n, \mathbb{R}).$$

But

$$X' + X = 0 \Longrightarrow \operatorname{tr} X = 0,$$

ie

$$\mathbf{o}(n) \subset \mathbf{sl}(n, \mathbb{R}).$$

Hence

$$\mathbf{so}(n) = \mathbf{o}(n) \cap \mathbf{sl}(n, \mathbb{R}) = \mathbf{o}(n).$$

The reason *why* $\mathbf{SO}(n)$ and $\mathbf{O}(n)$ have the same Lie algebra is that they coincide in the neighbourhood of I. In effect, $\mathbf{O}(n)$ has an extra ‘piece’ far from I, where $\det T = -1$. Since Lie theory only deals with what happens in the neighbourhood of I, it cannot distinguish between groups that coincide there.

Technically, as we shall see in Chapter 4, 2 linear groups having the same *connected component of I* (like $\mathbf{SO}(n)$ and $\mathbf{O}(n)$) will have the same Lie algebra.

5. *The Complex General Linear Group:*

$$\boxed{\mathbf{gl}(n, \mathbb{C}) = \mathbf{M}(n, \mathbb{C}).}$$

For from Example 1.5,

$$\begin{aligned} X \in \mathbf{M}(n, \mathbb{C}) &\implies XJ = JX \\ &\implies X^m J = JX^m \\ &\implies e^{tX} J = J e^{tX} \\ &\implies e^{tX} \in \mathbf{GL}(n, \mathbb{R}) \\ &\implies X \in \mathbf{gl}(n, \mathbb{R}); \end{aligned}$$

while conversely

$$\begin{aligned} e^{tX} \in \mathbf{GL}(n, \mathbb{C}) \quad \text{for all } t &\implies e^{tX} J = J e^{tX} \\ &\implies XJ = JX \\ &\implies X \in \mathbf{M}(n, \mathbb{C}), \end{aligned}$$

on equating coefficients of t.

6. *The Complex Special Linear Group:*

$$\boxed{\mathbf{sl}(n, \mathbb{C}) = \{X \in \mathbf{M}(n, \mathbb{C}) : \operatorname{tr} X = 0\}.}$$

Note that $\operatorname{tr} X$ here denotes the trace of X as a complex matrix. The result follows exactly as in Example 2.

7. *The Unitary Group:* $\mathbf{u}(n)$ consists of all *skew-hermitian matrices*, ie

$$\boxed{\mathbf{u}(n) = \{X \in \mathbf{M}(n, \mathbb{C}) : X^* + X = 0\}.}$$

For

$$(I + tX)^*(I + tX) = I + t(X^* + X) + o(t).$$

Hence

$$X \in \mathbf{u}(n) \implies X^* + X = 0.$$

Conversely,

$$\begin{aligned} X^* + X = 0 &\implies X^* = -X \\ &\implies (e^{tX})^* = e^{tX^*} = e^{-tX} = (e^{tX})^{-1} \\ &\implies e^{tX} \in \mathbf{U}(n) \quad \text{for all } t \\ &\implies X \in \mathbf{u}(n). \end{aligned}$$

8. *The Special Unitary Group:*

$$\boxed{\mathbf{su}(n) = \{X \in \mathbf{M}(n, \mathbb{C}) : X^* + X = 0, \operatorname{tr} X = 0\}.}$$

This follows at once from the Lemma above:

$$\mathbf{SU}(n) = \mathbf{U}(n) \cap \mathbf{SL}(n, \mathbb{C}) \Longrightarrow \mathbf{su}(n) = \mathbf{u}(n) \cap \mathbf{sl}(n, \mathbb{C}).$$

Notice that $\mathbf{su}(n)$ is *not* the same as $\mathbf{u}(n)$; a skew-hermitian matrix is not necessarily trace-free.

9. *The Symplectic Group:*

$$\boxed{\mathbf{sp}(n) = \{X \in \mathbf{M}(2n, \mathbb{C}) : X^* + X = 0, XJ = JX\}.}$$

For

$$\mathbf{Sp}(n) = \mathbf{Sp}(n, \mathbb{C}) \cap \mathbf{U}(2n),$$

where

$$\mathbf{Sp}(n, \mathbb{C}) = \{T \in \mathbf{GL}(n, \mathbb{C}) : XJ = JX\}.$$

Hence

$$\mathbf{sp}(n) = \mathbf{sp}(n, \mathbb{C}) \cap \mathbf{u}(2n),$$

by the Lemma above. The result follows, since

$$\mathbf{sp}(n, \mathbb{C}) = \{X \in \mathbf{M}(2n, \mathbb{C}) : XJ = JX\}$$

just as in Example 5.

Definition 3.2 *The dimension of a linear group G is the dimension of the real vector space $\mathcal{L}G$:*

$$\boxed{\dim G = \dim_{\mathbb{R}} \mathcal{L}G.}$$

Examples:

1. $\dim \mathbf{GL}(2, \mathbb{R}) = 4$.

2. $\dim \mathbf{SL}(2, \mathbb{R}) = 3$. For the general matrix in $\mathbf{sl}(2, \mathbb{R})$ is of the form

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$$

giving a vector space of dimension 3.

3. $\dim \mathbf{O}(n) = 3$. For the general matrix in $\mathbf{o}(3)$ is of the form

$$\begin{pmatrix} 0 & -a & b \\ a & 0 & -c \\ -b & c & 0 \end{pmatrix}$$

4. $\dim \mathbf{SO}(3) = 3$. For $\mathbf{so}(3) = \mathbf{o}(3)$.
5. $\dim \mathbf{U}(2) = 4$. For the general matrix in $\mathbf{u}(2)$ is of the form
$$\begin{pmatrix} ia & -b+ic \\ b+ic & id \end{pmatrix}$$
6. $\dim \mathbf{SU}(2) = 3$. For the general matrix in $\mathbf{su}(2)$ is of the form
$$\begin{pmatrix} ia & -b+ic \\ b+ic & -ia \end{pmatrix}$$
7. $\dim \mathbf{Sp}(1) = 3$. For the general *complex* matrix in $\mathbf{sp}(1, \mathbb{C})$ is
$$\begin{pmatrix} z & -w \\ \bar{w} & \bar{z} \end{pmatrix}$$
If this is also in $\mathbf{u}(2)$ then z must be pure imaginary, say
$$\begin{pmatrix} ia & -b-ic \\ b-ic & ia \end{pmatrix}$$

In Chapter 2 we defined the *logarithmic zone* in $\mathbf{GL}(n, \mathbb{R})$: an open neighbourhood of I mapped homeomorphically onto an open set in $\mathbf{M}(n, \mathbb{R})$. Our next result shows that *every linear group has a logarithmic zone*.

Proposition 3.2 *Suppose G is a linear group. Then there exists an open neighbourhood $W \ni 0$ in $\mathcal{L}G$ which is mapped homeomorphically by e^X onto an open neighbourhood $e^W \ni I$ in G.*

Proof ► This result is rather remarkable. It asserts that there exists a $\delta > 0$ such that
$$\|X\| \le \delta, \; e^X \in G \Longrightarrow X \in \mathcal{L}G.$$

Suppose this is not so. Then we can find a sequence $X_i \in \mathbf{M}(n, \mathbb{R})$ such that
$$X_i \to 0, \; e^{X_i} \in G, \; X_i \notin \mathcal{L}G.$$

Let us resolve each X_i into components along and perpendicular to $\mathcal{L}G$ (where perpendicularity is taken with respect to the inner product associated to the quadratic form $\|X\|^2$):
$$X_i = Y_i + Z_i \qquad (Y_i \in \mathcal{L}G, \; Z_i \perp \mathcal{L}G).$$

Notice that $\|Y_i\|, \|Z_i\| \leq \|X_i\|$.

Consider the set of matrices

$$\frac{Z_i}{\|Z_i\|}$$

on the unit ball in $\mathbf{M}(n, \mathbb{R})$. Since this ball is compact, we can find a convergent subsequence. Taking this subsequence in place of the original sequence, we may assume that

$$\frac{Z_i}{\|Z_i\|} \to Z$$

Since $\mathbb{Z}_i \perp \mathcal{L}G$ it follows that $Z \perp \mathcal{L}G$.

Now consider the sequence

$$\begin{aligned} T_i &= e^{-Y_i} e^{X_i} \\ &= \left(I - Y_i + \frac{Y_i^2}{2!} - \cdots\right)\left(I + (Y_i + Z + i) + \frac{(Y_i + Z_i)^2}{2!} + \cdots\right) \\ &= I + Z_i + O(\|X_i\| \|Z_i\|), \end{aligned}$$

since each remaining term will contain Z_i and will be of degree ≥ 2 in Y_i, Z_i. Let $\|Z_i\| = t_i$. Then

$$Z_i = t_i(Z + E_i),$$

where $E_i \to 0$. Thus

$$T_i = I + t_i Z + o(t_i).$$

From the Lemma to Proposition 1 above, this implies that

$$Z \in \mathcal{L}G.$$

But $Z \perp \mathcal{L}G$. So our original hypothesis is untenable; we *can* find a $\delta > 0$ such that if $\|X\| \leq \delta$ then

$$e^X \in G \iff X \in \mathcal{L}G.$$

◄

Corollary 3.1 *If G is a linear group then*

$$\boxed{\mathcal{L}G = 0 \iff G \text{ is discrete.}}$$

Proposition 3.3 *The connected component of I in a linear group G is a normal open subgroup G_0, generated by the exponentials $e^X (X \in \mathcal{L}G)$:*

$$\boxed{G_0 = \{e^{X_1} e^{X_2} \ldots e^{X_r} : X_1, X_2, \ldots, X_r \in \mathcal{L}G\}.}$$

Proof ▶ The set of all such matrices is clearly closed under multiplication and inversion, and so forms a subgroup G_0.

The path

$$t \mapsto e^{tX_1} \ldots e^{tX_r} \quad (0 \le t \le 1)$$

connects I to

$$e^{X_1} \ldots e^{X_r} \ni G_0.$$

Hence G_0 is connected.

By Proposition 2, $e^{\mathcal{L}G}$ is a neighbourhood of $0 \in G$. Hence $ge^{\mathcal{L}G}$ is a neighbourhood of g for each $g \in G$. But

$$ge^{\mathcal{L}G} \subset G_0 \text{ if } g \in G_0.$$

Hence G_0 is open. Recall that this implies G_0 is also closed. (For each coset of G_0 is open. Hence any union of cosets is open. Hence the complement of G_0, which is the union of all cosets apart from G_0 itself, is open, ie G_0 is closed.)

Since G_0 is also connected, it must be the connected component of I in G.

Finally, if $g \in G$ then gG_0g^{-1} is also connected. Hence

$$gG_0g^{-1} \subset G_0,$$

ie G_0 is normal in G. ◀

Remarks:

1. Note that by this Proposition, if G is a linear group then

$$G \text{ connected} \implies G \text{ arcwise connected}.$$

2. This Proposition is usually applied in reverse, ie we first determine (by some other means) the connected component G_0 of I in G. The Proposition then shows that each element of G_0 is expressible as a product of exponentials of elements of $\mathcal{L}G$.

   The following result—which really belongs to homotopy theory—is often useful in determining G_0.

Lemma 3.3 *Suppose the compact linear group G acts transitively on the compact space X; and suppose $x \in X$. Let $S = S(x)$ denote the stabiliser subgroup of x, ie*

$$s = \{T \in G : Tx = x\}.$$

Then

$$S \textit{ and } X \textit{ connected} \implies G \textit{ connected}.$$

Remark: Those familiar with homotopy will recognise in this the 0-dimensional part of the infinite exact homotopy sequence

$$0 \to \pi_0(S) \to \pi_0(G) \to \pi_0(X) \to \pi_1(S) \to \dots,$$

where π_i denotes the ith homotopy group) of the fibre space (G, X, S) with total space G, base space X, fibre S and projection

$$g \mapsto gx : G \to X.$$

Although we shall assume no knowledge of homotopy theory, it is interesting to note that the 1-dimensional part of this sequence will play a similar role in Chapter 7, in showing that

$$S \text{ and } X \text{ simply connected } \implies G \text{ simply connected}.$$

Proof of Lemma ▷ Suppose $g \in G_0$. Then gS is connected; and so $gS \subset G_0$. Hence

$$G_0 = G_0 S.$$

Since G_0 is closed (and so compact), so is

$$gG_0 x \subset X.$$

On the other hand since G_0 is open, $G - G_0$ is closed, and so therefore is

$$(G - G_0)x = X - G_0 x.$$

Thus $G_0 x$ is both open and closed; and so

$$G_0 x = X,$$

since X is connected. But since $G_0 = G_0 S$ this implies that

$$G_0 = G,$$

ie G is connected. ◁

Examples:

1. $\mathbf{SO}(n)$ *is connected for all* n. For consider the action

$$(T, v) \to Tv$$

of $\mathbf{SO}(n)$ on the $(n-1)$-sphere

$$S^{n-1} = \{v \in \mathbb{R}^n : |v| = 1\}.$$

This action is transitive; and the stabiliser subgroup of

$$e_n = (0, 0, \ldots, 1)$$

can be identified with $\mathbf{SO}(n-1)$. Applying the Lemma,

$$\mathbf{SO}(n-1) \text{ and } S^{n-1} \text{ connected } \Longrightarrow \mathbf{SO}(n) \text{ connected}.$$

Thus by induction, starting from $\mathbf{SO}(1)$ and S^1, we deduce that $\mathbf{SO}(n)$ is connected for all n.

2. $\mathbf{SO}(n)$ *is the connected component of* I *in* $\mathbf{O}(n)$. For

$$T \in \mathbf{O}(n) \Longrightarrow T'T = I \Longrightarrow (\det T)^2 = 1 \Longrightarrow \det T = \pm 1.$$

Thus no path in $\mathbf{O}(n)$ can enter or leave $\mathbf{SO}(n)$, since this would entail a sudden jump in $\det T$ from $+1$ to -1, or vice versa.

3. $\mathbf{U}(n)$ *is connected for all* n. For consider the action of $\mathbf{U}(n)$ on

$$S^{2n-1} = \{v \in C^n : |v| = 1\}.$$

The stabiliser subgroup of

$$e_n = (0, 0, \ldots, 1)$$

can be identified with $\mathbf{U}(n-1)$; and so we deduce by induction, as in Example 1, that $\mathbf{U}(n)$ is connected for all n.

4. $\mathbf{SU}(n)$ *is connected for all* n. Again this follows from the action of $\mathbf{SU}(n)$ on S^{2n-1}.

5. $\mathbf{Sp}(n)$ *is connected for all* n. This follows in the same way from the action of $\mathbf{Sp}(n)$ on

$$S^{4n-1} = \{v \in \mathbb{H}^n : |v| = 1\}.$$

6. $\mathbf{SL}(n, \mathbb{R})$ *is connected for all* n. For suppose $T \in \mathbf{SL}(n, \mathbb{R})$. Then $T'T$ is a positive-definite matrix, and so has a positive-definite square-root, Q say:

$$T'T = Q^2.$$

(To construct Q, diagonalise the quadratic form $v'Tv$ and take the square-root of each diagonal element.) Now set

$$O = TQ^{-1}.$$

Then

$$O'O = I,$$

ie $O \in \mathbf{O}(n)$. In fact, since $\det T = 1$ and $\det Q > 0$,

$$O \in \mathbf{SO}(n).$$

Thus

$$T = OQ,$$

where $O \in \mathbf{SO}(n)$ and $Q \in P$, the space of positive-definite matrices. Thus

$$\mathbf{SL}(n, \mathbb{R}) = \mathbf{SO}(n)P.$$

Now P is connected; in fact it is convex:

$$A, B \in P \Longrightarrow tA + (1-t)B \in P \quad \text{for all } t \in [0, 1].$$

Since $\mathbf{SO}(n)$ is also connected, so too is

$$\mathbf{SL}(n, \mathbb{R}) = \mathbf{SO}(n)P.$$

Summary: To each linear group $G \subset \mathbf{GL}(n, \mathbb{R})$ there corresponds a vector subspace $\mathcal{L}G \subset \mathbf{M}(n, \mathbb{R})$. Two linear groups G and H correspond to the same subspace if and only if they have the same connected component of the identity: $\mathcal{L}G = \mathcal{L}H \iff G_0 = H_0$. In other words, there is a 1–1 correspondence between *connected* linear groups G and the subspaces $\mathcal{L}G$.

The exponential map and the inverse logarithmic map define a homeomorphism (but *not* a homomorphism) between a neighbourhood U of I in G and a neighbourhood V of 0 in $\mathcal{L}G$. With a little artistic licence we may say that the logarithmic projection turns subgroups into subspaces.

Exercises

In Exercises 01–10 determine the dimension of the given group

1. $\mathbf{GL}(n,\mathbb{R})$ 2. $\mathbf{SL}(n,\mathbb{R})$ 3. $\mathbf{O}(n)$ 4. $\mathbf{SO}(n)$ 5. $\mathbf{GL}(n,\mathbb{C})$

6. $\mathbf{SL}(n,\mathbb{C})$ 7. $\mathbf{U}(n)$ 8. $\mathbf{SU}(n)$ 9. $\mathbf{Sp}(n)$ 10. $\mathbf{O}(1,3)$

In Exercises 11–15 determine the connected component of I, and the number of components, in each of the following groups.

11. $\mathbf{GL}(n,\mathbb{R})$ 12. $\mathbf{GL}(n,\mathbb{C})$ 13. $\mathbf{SL}(n,\mathbb{C})$ 14. $\mathbf{O}(1,1)$ 15. $\mathbf{O}(1,3)$

Chapter 4

[

The Lie Algebra of a Linear Group II]The Lie Algebra of a Linear Group II: The Lie Product

> The subspace $\mathcal{L}G$ corresponding to a linear group G is closed under the Lie product $[X, Y] = XY - YX$, and thus consitutes a Lie algebra. Algebraically, the Lie product reflects the non-commutativity of G—an abelian group has trivial Lie algebra. Geometrically, the Lie product measures the curvature of G.

Definition 4.1 *For $X, Y \in \mathbf{M}(n, \mathbb{R})$ we set*

$$\boxed{[X, Y] = XY - YX.}$$

The matrix $[X, Y]$ is called the Lie product of X and Y.

Proposition 4.1 *If G is a linear group then*

$$\boxed{X, Y \in \mathcal{L}G \Longrightarrow [X, Y] \in \mathcal{L}G.}$$

Proof ► Our argument is very similar to that used in the proof of Proposition 3.1 to show that

$$X, Y \in \mathcal{L}G \Longrightarrow X + Y \in \mathcal{L}G.$$

Suppose

$$X, Y \in \mathcal{L}G.$$

Then $e^{tX}, e^{tY} \in G$ for all $t \in \mathbb{R}$; and so

$$e^{tX} e^{tY} e^{-tX} e^{-tY} \in G.$$

But

$$\begin{aligned}
& e^{tX}e^{tY}e^{-tX}e^{-tY} \\
= \quad & (I + tX + \frac{t^2X^2}{2})(I + tY + \frac{t^2Y^2}{2})(I - tX + \frac{t^2X^2}{2})(I - tY + \frac{t^2Y^2}{2}) + o(t^2) \\
= \quad & I + t^2[X,Y] + o(t^2) \\
= \quad & I + s[X,Y] + o(s),
\end{aligned}$$

with $s = t^2$. Hence by the Lemma to Proposition 3.1,

$$[X,Y] \in \mathcal{L}G.$$

◄

Proposition 4.2 *The Lie product $[X,Y]$ defines a skew-symmetric bilinear map $\mathcal{L}G \times \mathcal{L}G \to \mathcal{L}G$, ie*

1. $[aX,Y] = a[X,Y]$
2. $[X_1 + X_2, Y] = [X_1, Y] + [X_2, Y]$
3. $[X, Y_1 + Y_2] = [X, Y_1] + [X, Y_2]$
4. $[Y,X] = -[X,Y]$

In addition it satisfies Jacobi's identity

$$\boxed{[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0}$$

Proof ► All is clear except for (J), and that is a matter for straightforward verification:

$$\begin{aligned}
& [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] \\
= \quad & X(YZ - ZY) - (YZ - ZY)X + Y(ZX - XZ) \\
& \quad -(ZX - XZ)Y + Z(XY - YX) - (XY - YX)Z \\
= \quad & 0,
\end{aligned}$$

the 12 terms cancelling in pairs. ◄

Definition 4.2 *A Lie algebra $\mathcal{L}$ over k (where $k = \mathbb{R}$ or $\mathbb{C}$) is a finite-dimensional vector space $\mathcal{L}$ over k, together with a skew-symmetric bilinear map $\mathcal{L} \times \mathcal{L} \to \mathcal{L}$, which we denote by $[X,Y]$, satisfying Jacobi's identity (J).*

If G is a linear group then the real Lie algebra defined on $\mathcal{L}G$ by the Lie product is called the Lie algebra of G.

Remarks:

1. Henceforth $\mathcal{L}G$ (and similarly $\mathbf{gl}(n, \mathbb{R})$, etc) will denote the Lie algebra of G, ie the space $\mathcal{L}G$ together with the Lie product on this space.

2. Note that the Lie algebra of a linear group is always *real*, even if G is complex, ie $G \subset \mathbf{GL}(n, C)$. The point of introducing complex Lie algebras will only become apparent in Chapter7, when we consider complex representations of linear groups.

3. It follows at once from the skew-symmetry of the Lie product that
$$[X, X] = 0 \quad \forall X \in \mathcal{L}.$$

4. In defining the Lie product in a Lie algebra $\mathcal{L}$ with basis $e_1, \ldots, e_m$ it is only necessary to give the m(m-1)/2 products
$$[e_i, e_j] \quad (i < j)$$
since these determine the general product $[X, Y]$ by skew-symmetry and linearity.

5. With the same notation, we can express each Lie product $[e_i, e_j]$ of basis elements as a linear combination of these elements:
$$[e_i, e_j] = \sum_k c_{ij}^k e_k.$$
The m^3 scalars
$$c_{ij}^k \quad (1 \leq i, j, k \leq m)$$
are called the *structure constants* of the Lie algebra. In theory we could define a Lie algebra by giving its structure constants; in practice this is rarely a sensible approach.

Examples:

1. The space $\mathbf{so}(3)$ consists of all skew-symmetric 3×3 matrices. As basis we might choose
$$U = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

It is readily verified that

$$[U, V] = UV - VU = W,\ [U, W] = UW - WU = -V,\ [V, W] = VW - WV = U.$$

Thus we can write

$$\mathbf{so}(3) = \langle U, V, W : [U, V] = W,\ [V, W] = U,\ [W, U] = V \rangle.$$

2. The space $\mathbf{su}(2)$ consists of all skew-hermitian 2×2 complex matrices with trace 0. As basis we might choose

$$A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

It is readily verified that

$$[A, B] = 2C,\ [A, C] = -2B,\ [B, C] = 2A.$$

Thus

$$\mathbf{su}(2) = \langle A, B, C : [A, B] = 2C,\ [B, C] = 2A,\ [C, A] = 2B \rangle.$$

Notice that the Lie algebras $\mathbf{so}(3)$ and $\mathbf{su}(2)$ are isomorphic,

$$\mathbf{so}(3) = \mathbf{su}(2),$$

under the correspondence

$$2U \longleftrightarrow A,\ 2V \longleftrightarrow B,\ 2W \longleftrightarrow C.$$

Intuitively, we can see how this isomorphism arises. The covering homomorphism $\Theta : \mathbf{SU}(2) \to \mathbf{SO}(3)$ establishes a *local isomorphism* between $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$. If one remains close to I these 2 groups look exactly the same. But Lie theory is a *local* theory—the Lie algebra $\mathcal{L}G$ depends only on the structure of G *near to* I. (In particular, it depends only on the connected component of I in G, whence $\mathbf{so}(3) = \mathbf{o}(3)$.) Thus *covering groups have the same Lie algebra*. We shall return—more rigorously—to this very important point in Chapter 8.

But we note now that the groups $\mathbf{SO}(3)$ and $\mathbf{SU}(2)$ are certainly not isomorphic, since one has trivial, and the other non-trivial, centre:

$$Z\mathbf{SO}(3) = \{I, -I\}, Z\mathbf{SU}(2) = \{I\}.$$

3. The space $\mathbf{sl}(2,\mathbb{R})$ consists of all real 2×2 matrices with trace 0. As basis we might choose

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

It is readily verified that

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

Thus

$$\mathbf{sl}(2,\mathbb{R}) = \langle H,E,F : [H,E] = 2E,\ [H,F] = -2F,\ [E,F] = H\rangle.$$

This Lie algebra is not isomorphic to $\mathbf{so}(3) = \mathbf{su}(2)$. For the subspace $\langle U, V\rangle$ is closed under the Lie product. (It corresponds, as we shall see in Chapter 6, to the subgroup of lower triangular matrices.) But it is readily verified that no 2-dimensional subspace of $\mathbf{so}(3)$ is closed under the Lie product.

4. As an exercise, let us determine all the Lie algebras of dimension ≤ 3. This will allow us to introduce informally some concepts which we shall define formally later.

(a) The Lie algebra $\mathcal{L}G$ is said to be *abelian* if the Lie product is trivial:

$$[X,Y] = 0 \qquad \text{for all } X,Y \in \mathcal{L}.$$

We shall look at abelian Lie algebras in the next Chapter. Evidently there is just 1 such algebra in each dimension.

(b) The *derived algebra* $\mathcal{L}'$ of a Lie algebra $\mathcal{L}$ is the subspace spanned by all Lie products:

$$\mathcal{L}' = \{[X,Y] :\ X,Y \in \mathcal{L}\}.$$

This is an *ideal* in $\mathcal{L}$:

$$X \in \mathcal{L},\ Y \in \mathcal{L}' \Longrightarrow [X,Y] \in \mathcal{L}'.$$

Evidently $\mathcal{L}' = 0$ if and only if $\mathcal{L}$ is abelian.

(c) The *centre* of the Lie algebra $\mathcal{L}$,

$$Z\mathcal{L} = \{Z \in \mathcal{L} :\ [X,Z] = 0 \text{ for all } X \in \mathcal{L}\},$$

is also an ideal in $\mathcal{L}$. Evidently $\mathcal{L}$ is abelian if and only if $Z\mathcal{L} = \mathcal{L}$.

(d) Suppose $X \in \mathcal{L}$. We denote by $\operatorname{ad} X : \mathcal{L} \to \mathcal{L}$ the linear map defined by

$$\operatorname{ad} X : Y \mapsto [X, Y].$$

Jacobi's identity can be re-stated in the form

$$\operatorname{ad}[X, Y] = \operatorname{ad} X \operatorname{ad} Y - \operatorname{ad} Y \operatorname{ad} X.$$

To see this, apply both sides to Z. On the left we have

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

On the right we have

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

The two are equal by Jacobi's identity.

Dimension 1 Since $[X, X] = 0$, if $\dim \mathcal{L} = 1$ then $\mathcal{L}$ is abelian.

Dimension 2 Either $\mathcal{L}$ is abelian, or $\dim \mathcal{L}' = 1$, since

$$[aX + bY, cX + dY] = (ad - bc)[X, Y].$$

Suppose $\dim \mathcal{L}' = 1$, say $\mathcal{L}' = \langle X \rangle$. Then we can find a Y such that $[X, Y] = X$. Thus

$$\mathcal{L} = \langle X, Y : [X, Y] = X \rangle.$$

So there are just 2 Lie algebras of dimension 2.

Dimension 3 We have $\dim \mathcal{L}' = 0, 1, 2$ or 3.

If $\dim \mathcal{L}' = 0$ then $\mathcal{L}$ is abelian.

If $\dim \mathcal{L}' = 1$, say $\mathcal{L}' = \langle X \rangle$, then it is not hard to show that the centre $Z\mathcal{L}$ must also be of dimension 1; and

$$\mathcal{L} = \mathcal{L}_1 \oplus Z\mathcal{L},$$

where $\mathcal{L}_1$ is a 2-dimensional Lie algebra. Thus there is just 1 Lie algebra in this category:

$$\mathcal{L} = \langle X, Y, Z : [X, Y] = X,\ [X, Z] = [Y, Z] = 0 \rangle.$$

If $\dim \mathcal{L}' = 2$, say $\mathcal{L}' = \langle X, Y \rangle$, then $\mathcal{L}'$ must be abelian. For suppose not. Then we can find $X, Y \in \mathcal{L}'$ such that $[X, Y] = X$. Suppose

$$[Y, Z] = aX + bY,\ [Z, X] = cX + dY.$$

Then Jacobi's identity gives

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = bX - cX + cX + dY.$$

Thus $b = d = 0$. But then $\mathcal{L}' = \langle X \rangle$, contrary to hypothesis.

Thus $\mathcal{L}' = \langle X, Y \rangle$ is abelian. Take any $Z \in \mathcal{L} \setminus \mathcal{L}'$. Since $\operatorname{ad} Z(\mathcal{L}) \subset \mathcal{L}'$, $\operatorname{ad} Z$ defines a map

$$A_Z : \mathcal{L}' \to \mathcal{L}'.$$

If we take $Z' = aX + bY + cZ$ in place of Z then

$$A_{Z'} = cA_Z.$$

Thus the map A_Z is defined up to a scalar multiple by $\mathcal{L}$. Conversely, it is readily verified that any linear map $A_Z : \mathcal{L}' \to \mathcal{L}'$ defines a Lie algebra, since all 3 terms in Jacobi's identity for X, Y, Z vanish.

There are 2 families of solutions, according as the eigenvalues of A_Z are real, or complex conjugates. Note that A_Z is surjective, since $\operatorname{ad}_Z(L) = \mathcal{L}'$. Hence the eigenvalues of A_Z are not both 0 (ie A_Z is not idempotent).

If the eigenvalues are complex conjugates $\rho e^{i\theta}$ then we can make $\rho = 1$ by taking $\frac{1}{\rho}Z$ in place of Z. This gives the family of Lie algebras

$$\mathcal{L}(\theta) = \langle X, Y, Z : [X,Y] = 0,\ [Z,X] = \cos\theta X - \sin\theta Y,\ [Z,Y] = \sin\theta X + \cos\theta Y \rangle.$$

Similary if the eigenvalues are real we can take them as $1, \rho$. and we obtain the family

$$\mathcal{L}(\rho) = \langle X, Y, Z : [X,Y] = 0,\ [Z,X] = X,\ [Z,Y] = \rho Y \rangle.$$

We come now to the case of greatest interest to us, where

$$\mathcal{L}' = \mathcal{L}.$$

A Lie algebra with this property is said to be *semisimple*. (We shall give a different definition of semisimplicity later, but we shall find in the end that it is equivalent to the above.) Perhaps the most important Theorem in this Part is that *every representation of a semisimple Lie algebra is semisimple*. But that is a long way ahead.

Since

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

it follows that

$$\mathrm{tr}\ \mathrm{ad}[X, Y] = 0.$$

Thus

$$X \in \mathcal{L}' \Longrightarrow \mathrm{tr}\ \mathrm{ad}\, X = 0.$$

In particular if $\mathcal{L}' = \mathcal{L}$ then $\mathrm{tr}\ \mathrm{ad}\, X = 0$ for all $X \in \mathcal{L}$. Hence the eigenvectors of $\mathrm{ad}\, X$ are $0, \pm\lambda$. On the other hand, the characteristic equation is real. So the eigenvectors are either of the form $0, \pm\rho$ or $0, \pm\rho i$ with ρ real.

Before going further, let us dispose of one possibility: that the eigenvalues of $\mathrm{ad}\, X$ might always be $0, 0, 0$. Recall that this is the case if and only if $\mathrm{ad}\, X$ is nilpotent. (This follows from the Cayley-Hamilton Theorem, that a linear transformation—or square matrix—satisfies its own characteristic equation.)

A Lie algebra $\mathcal{L}$ is said to be *nilpotent* if $\mathrm{ad}\, X$ is nilpotent for all $X \in \mathcal{L}$. As we shall see later (Engel's Theorem) a nilpotent Lie algebra $\mathcal{L}$ cannot have $\mathcal{L}' = \mathcal{L}$, ie $\mathcal{L}$ cannot be both nilpotent and semisimple.

This is very easy to establish in the present case, where $\dim \mathcal{L} = 3$. Note first that

$$\mathcal{L} = \langle X, Y, Z \rangle \Longrightarrow \mathcal{L}' = \langle [Y, Z],\ [Z, X],\ [X, Y] \rangle.$$

It follows that if $\mathcal{L}' = L$ then

$$[X, Y] = 0 \Longleftrightarrow X, Y \text{ linearly dependent}.$$

Now suppose $\mathrm{ad}\, X$ is nilpotent for some $X \neq 0$. Then $\mathrm{ad}\, X \neq 0$, since otherwise $[X, Y] = 0$ for all $Y \in \mathcal{L}$. Thus we can find $Y \in \mathcal{L}$ such that

$$Z = \mathrm{ad}\, X(Y) \neq 0 \text{ but } \mathrm{ad}\, X(Z) = [X, Z] = 0.$$

This implies, as we have seen, that $Z = \rho X$. Thus

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

So $\mathrm{ad}\, Y$ has eigenvalue $-\rho \neq 0$, and is not nilpotent.

Thus there exists an $X \in \mathcal{L}$ with $\mathrm{ad}\, X$ not nilpotent, with the 2 possibilites outlined above.

(a) For some $X \in \mathcal{L}$, $\mathrm{ad}\, X$ has eigenvalues $0, \pm\rho$ where $\rho > 0$. Taking $\frac{1}{\rho}X$ in place of X, we may suppose that $\rho = 1$. Taking the eigenvectors of X as a basis for $\mathcal{L}$, we get

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

Suppose

$$[Y, Z] = aX + bY + cZ.$$

Jacobi's identity yields

$$[X, [Y, Z]]+[Y, [Z, X]]+[Z, [X, Y]] = bY - cZ + [Y, Z] + [Z, Y] = 0.$$

Thus $b = c = 0$, ie

$$[Y, Z] = aX.$$

Dividing Y by a, we get

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

So we have just 1 Lie algebra,

$$\mathcal{L} = \langle X, Y, Z : [X, Y] = Y,\ [X, Z] = -Z,\ [Y, Z] = X\rangle.$$

In fact

$$\mathcal{L} = \mathbf{sl}(2, \mathbb{R})$$

under the correspondence $X \mapsto \frac{1}{2}H,\ E \mapsto Y, F \mapsto Z$.

(b) Alternatively, $\operatorname{ad} X$ has eigenvalues $0, \pm\rho i$ with some $\rho \geq 0$ for *every* $X \in \mathcal{L}$. (For otherwise we fall into the first case.) Choose one such X. As before, on replacing X by $\frac{1}{\rho}$ we may suppose that $\rho = 1$, ie $\operatorname{ad} X$ has eigenvalues $0, \pm i$. Taking the i-eigenvector of $\operatorname{ad} X$ to be $Z + iY$,

$$[X, Z + iY] = i(Z + iY) = -Y + iZ.$$

Thus

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

Suppose

$$[Y, Z] = aX + bY + cZ.$$

Jacobi's identity yields

$$[X, [Y, Z]]+[Y, [Z, X]]+[Z, [X, Y]] = bZ - cY + [Y, Y] + [Z, Z] = 0.$$

Thus $b = c = 0$, ie

$$[Y, Z] = aX.$$

Dividing Y, Z by $\sqrt{|a|}$, we may suppose that

$$[Y, Z] = \pm X.$$

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

$$\operatorname{ad} Y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

This has eigenvalues $0, \pm 1$, and so falls into the first case. Hence

$$\mathcal{L} = \langle X, Y, Z : [X,Y] = Z,\ [Z,X] = Y,\ [Y,Z] = X \rangle.$$

In other words,

$$\mathcal{L} = \mathbf{so}(3).$$

It remains to show that $\mathbf{so}(3) = \mathbf{su}(2)$ and $\mathbf{sl}(2, \mathbb{R})$ are not isomorphic. We shall see later that this follows from the fact that $\mathbf{so}(3)$ and $\mathbf{su}(2)$ are compact, while $\mathbf{sl}(2, \mathbb{R})$ is not; for we shall show that the compactness of a group G is reflected in its Lie algebra $\mathcal{L}G$. But we can reach the same result by a cruder argument. From our argument above, it is sufficient to show that $\operatorname{ad} X$ has eigenvalues $0, \pm\rho i$ for every $X \in \mathbf{so}(3)$. But it is readily seen that

$$\operatorname{ad} U = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},\ \operatorname{ad} V = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},\ \operatorname{ad} W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus, with respect to this basis, $\operatorname{ad} X$ is always represented by a skew-symmetric matrix. The result follows since the eigenvalues of such a matrix are either purely imaginary or 0.

Summary: To each linear group G there corresponds a Lie algebra $\mathcal{L}G$. Most of the information about G is ‘encoded’ in $\mathcal{L}G$; and since $\mathcal{L}G$ is far easier to analyse—using standard techniques of linear algebra—it provides a powerful tool in the study of linear groups.

Exercises

In Exercises 01–10 determine the Lie algebra of the given group

1.	$\mathbf{GL}(2,\mathbb{R})$	2.	$\mathbf{SL}(3,\mathbb{R})$	3.	$\mathbf{O}(2)$	4.	$\mathbf{O}(3)$	5.	$\mathbf{SO}(2)$
6.	$\mathbf{SO}(4)$	7.	$\mathbf{GL}(2,\mathbb{C})$	8.	$\mathbf{SL}(2,\mathbb{C})$	9.	$\mathbf{U}(2)$	10.	$\mathbf{SU}(3)$
11.	$\mathbf{Sp}(1)$	12.	$\mathbf{Sp}(2)$	13.	$\mathbf{O}(1,1)$	14.	$\mathbf{O}(1,3)$	15.	$\mathbf{O}(2,2)$

Chapter 5

Abelian Linear Groups

As a happy by-product, Lie theory gives us the structure of connected abelian linear groups.

Definition 5.1 *The Lie algebra $\mathcal{L}$ is said to be abelian if the Lie product is trivial, ie*

$$[X, Y] = 0 \quad \textit{for all } X, Y \in \mathcal{L}.$$

Proposition 5.1 *If G is a linear group then*

$$G \textit{ abelian} \implies \mathcal{L}G \textit{ abelian}.$$

If in addition G is connected then

$$\mathcal{L}G \textit{ abelian} \iff G \textit{ abelian}.$$

Proof ► Suppose G is abelian; and suppose

$$X, Y \in \mathcal{L}G.$$

Then e^{tX}, e^{tY} commute for all t. If t is sufficiently small, tX and tY will lie in the logarithmic zone U, so that

$$tX = \log e^{tX}, tY = \log e^{tY}$$

by Proposition 2.6. In particular tX, tY are expressible as power-series in e^{tX}, e^{tY} respectively. Hence tX, tY commute; and so X, Y commute, ie $\mathcal{L}G$ is abelian.

Conversely, suppose G is connected and $\mathcal{L}G$ is abelian, ie

$$[X, Y] = XY - YX = 0 \quad \forall X, Y \in \mathcal{L}G.$$

Then e^X, e^Y commute, by Proposition 2.2. and so therefore do any 2 products

$$e^{X_1} \dots e^{X_r}, e^{Y_1} \dots e^{Y_s} \qquad (X_1, \dots, X_r, Y_1, \dots, Y_s \in \mathcal{L}G).$$

Hence $G = G_0$ is abelian, by Proposition 3.3. ◄

Proposition 5.2 *If G is a connected abelian linear group then G is isomorphic to a cylinder group, ie*

$$\boxed{G = \mathbb{T}^r \times \mathbb{R}^s,}$$

where $\mathbb{T}$ is the torus

$$\mathbb{T} = \mathbb{R}/\mathbb{Z}.$$

In particular, the only compact connected abelian linear groups are the tori $\mathbb{T}^n$.

Lemma 5.1 *If G is a connected abelian linear group then the map*

$$X \mapsto e^X : \mathcal{L}G \to G$$

is a surjective homomorphism (of abelian groups) with discrete kernel.

Proof of Lemma ▷ The exponential map is a homomorphism in this case, by Proposition 2.2. It is surjective, by Proposition 3.3, since

$$e^{X_1} \dots e^{X_r} = e^{X_1 + \dots + X_r} \qquad (X_1, \dots, X_r \in \mathcal{L}G).$$

Its kernel is discrete, by Proposition 3.2; for the exponential map is one-one in the logarithmic zone. ◁

Proof ▶ By the Lemma,

$$G = \frac{\mathbb{R}^n}{K},$$

where K is a discrete subgroup of $\mathbb{R}^n$.

We shall show by induction on n that such a subgroup has a $\mathbb{Z}$-basis consisting of $m \leq n$ linearly independent vectors

$$v_1, \dots, v_m \in \mathbb{R}^n,$$

ie K consists of all linear combinations

$$a_1 v_1 + \dots + a_m v_m \quad (a_1, \dots, a_m \in \mathbb{Z}).$$

Let v_1 be the closest point of K to 0 (apart from 0 itself). We may suppose, on choosing a new basis for $\mathbb{R}^n$, that

$$v_1 = (1, 0, \dots, 0).$$

Now let

$$p : \mathbb{R}^n \to \mathbb{R}^{n-1}$$

be the projection onto the last $n-1$ coordinates, ie

$$p(x_1, x_2, \ldots, x_n) = (x_2, \ldots, x_n).$$

Then the subgroup

$$pK \subset \mathbb{R}^{n-1}$$

is discrete. For each point

$$(x_2, \ldots, x_n) \in pK$$

arises from a point

$$v = (x_1, x_2, \ldots, x_n) \in K;$$

and we may suppose, on adding a suitable integral multiple of v_1 to v, that

$$-\frac{1}{2} < x_1 \leq \frac{1}{2}.$$

But then v would clearly be closer to 0 than v_1 (contrary to hypothesis) if $x_2, \ldots, x_n$ were all very small.

Applying the inductive hypothesis we can find a $\mathbb{Z}$-basis for pK consisting of linearly independent vectors

$$u_2, \ldots, u_m \in \mathbb{R}^{n-1}.$$

Choose

$$v_2, \ldots, v_m \in \mathbb{R}^n$$

such that

$$pv_i = u_i.$$

Then it is easy to see that $v_1, v_2, \ldots, v_m$ are linearly independent, and form a $\mathbb{Z}$-basis for K.

Again, on choosing a new basis for $\mathbb{R}^n$ we may suppose that

$$v_1 = (1, 0, \ldots, 0), v_2 = (0, 1, \ldots, 0), \ldots, v_m = (0, \ldots, 1, \ldots, 0),$$

ie

$$K = \{(a_1, \ldots, a_m, 0, \ldots, 0) : a_i \in \mathbb{Z}\}.$$

Then

$$\begin{aligned} G &= \mathbb{R}^n/K \\ &= \mathbb{R}/\mathbb{Z} + \ldots + \mathbb{R}/\mathbb{Z} + \mathbb{R} + \ldots + \mathbb{R} \\ &= \mathbb{T}^m \times \mathbb{R}^{n-m}. \end{aligned}$$

◄

Remark: We shall find this result extremely useful later, in studying the structure of a general compact linear group G. For if we take any element $g \in G$ then the smallest *closed subgroup* of G containing g is abelian; and so the connected component of I in this group must be a torus.

Summary: If G is abelian then so is $\mathcal{L}G$; and the exponential map in this case is a homomorphism, mapping onto the connected component G_0 of G. We deduce that every connected abelian linear group is of the form $\mathbb{T}^m \times \mathbb{R}^n$.

Chapter 6

The Lie Functor

To each linear group G we have associated a Lie algebra $\mathcal{L}G$. But that is only half the story—to complete it we must show that each homomorphism $G \to H$ of linear groups gives rise to a homomorphism $\mathcal{L}G \to \mathcal{L}H$ of the associated Lie algebras.

Definition 6.1 *A homomorphism*

$$f : \mathcal{L} \to \mathcal{M}$$

of Lie algebras over k is a linear map which preserves the Lie product, ie

1. $f(aX) = a(fX)$
2. $f(X + Y) = fX + fY$
3. $f[X, Y] = [fX, fY]$

Proposition 6.1 *Suppose*

$$F : G \to H$$

is a continuous homomorphism of linear groups. Then there exists a unique homomorphism

$$f = \mathcal{L}F : \mathcal{L}G \to \mathcal{L}H$$

of the corresponding Lie algebras such that

$$\boxed{e^{fX} = F(e^X)}$$

for all $X \in \mathcal{L}G$.

Proof ► Suppose $X \in \mathcal{L}G$. Then the composition

$$t \mapsto e^{tX} \mapsto F(e^{tX}) : \mathbb{R} \to H$$

is a continuous homomorphism. Hence, by Proposition 2.5, there exists a unique element $fX \in \mathcal{L}H$ such that

$$e^{t(fX)} = F(e^{tX}) \quad \forall t \in \mathbb{R}.$$

We must show that the 3 conditions of Definition 1 are satisfied.

1. This follows at once from the definition of f.

2. We saw in the proof of Proposition 3.1 that

$$(e^{tX/m}e^{tY/m})^m \to e^{t(X+Y)}.$$

Applying the homomorphism F to each side,

$$(F(e^{tX/m})F(e^{tY/m}))^m \to F(e^{t(X+Y)}),$$

ie

$$(e^{tfX/m}e^{tfY/m})^m \to e^{tf(X+Y)}.$$

But by the same Lemma,

$$(e^{tfX/m}e^{tfY/m})^m \to e^{t(fX+fY)}.$$

Hence

$$e^{tf(X+Y)} = e^{t(fX+fY)} \quad \forall t;$$

and so, by Proposition 2.5,

$$f(X+Y) = fX + fY.$$

3. We saw in the proof of Proposition 4.1 that

$$(e^{tX}e^{tY}e^{-tX}e^{-tY})^{m^2} \to e^{t^2[X,Y]}.$$

The result follows from this as in (2).

◄

Proposition 6.2 *The assignment $G \to \mathcal{L}G$ is functorial, ie*

1. *If*
$$E : G \to H, \quad F : H \to K$$
are 2 continuous homomorphisms of linear groups then
$$\mathcal{L}(FE) = (\mathcal{L}F)(\mathcal{L}E).$$

2. *The identity on G induces the identity on LG,*
$$\mathcal{L}1_G = 1_{\mathcal{L}G}.$$

Corollary 6.1 *If the linear groups G and H are isomorphic then so are their Lie algebras:*
$$\boxed{G \cong H \Longrightarrow \mathcal{L}G \cong \mathcal{L}H.}$$

Remark: This is a most important result, since it allows us to extend Lie theory from linear to *linearisable* groups. Thus we can speak of the Lie algebra of a topological group G, provided G is *isomorphic to some linear group*—we need not specify this linear group or the isomorphism, so long as we have established that they exist. We shall return to this point later.

Proposition 6.3 *Suppose*
$$F : G \to H$$
is a continuous homomorphism of linear groups; and suppose G is connected. Then F is completely determined by $f = \mathcal{L}F$. More precisely, if
$$F_1, F_2 : G \to H$$
are 2 such homomorphisms then
$$\boxed{\mathcal{L}F_1 = \mathcal{L}F_2 \Longrightarrow F_1 = F_2.}$$

Proof ► Suppose $T \in G$. By Proposition 3.4,
$$T = e^{X_1} \dots e^{X_r}.$$
Hence
$$\begin{aligned} F(T) &= F(e^{X_1} \dots e^{X_r}) \\ &= F(e^{X_1}) \dots F(e^{X_r}) \\ &= e^{fX_1} \dots e^{fX_r}. \end{aligned}$$

Thus $F(T)$ is completely determined once f is known. ◄

Remark: This result shows that if G is connected (and if it is not we can always replace it by its connected component G_0) then there is at most one group homomorphism

$$F : G \to H$$

corresponding to a given Lie algebra homomorphism

$$f : \mathcal{L}G \to \mathcal{L}H.$$

One might say that *nothing is lost in passing from group homomorphism to Lie algebra homomorphism*.

Whether in fact f can be 'lifted' to a group homomorphism F in this way is a much more difficult question, which we shall consider in Chapter 7.

Proposition 6.4 *Suppose*

$$F : G \to H$$

is a continuous homomorphism of linear groups. Then

$$K = \ker F$$

is a linear group; and

$$\boxed{\mathcal{L}K = \ker(\mathcal{L}F).}$$

Proof ► Suppose G is a linear group, ie a closed subgroup of $\mathbf{GL}(n, \mathbb{R})$. Since K is closed in G, it is also closed in $\mathbf{GL}(n, \mathbb{R})$. Thus $K \subset \mathbf{GL}(n, \mathbb{R})$ is a linear group.

Moreover

$$\begin{aligned} X \in \mathcal{L}K &\implies e^{tX} \in K \quad \forall t \in \mathbb{R} \\ &\implies F(e^{tX}) = e^{tfX} = I \quad \forall t \in \mathbb{R} \\ &\implies fX = 0 \\ &\implies X \in \ker f. \end{aligned}$$

◄

Corollary 6.2 *Suppose*

$$F : G \to H$$

is a continuous homomorphism of linear groups. Then

$$\mathcal{L}F \textit{ injective} \iff \ker F \textit{ discrete}.$$

In particular

$$\boxed{F \textit{ injective} \implies \mathcal{L}F \textit{ injective}.}$$

Proposition 6.5 *Suppose*

$$F : G \to H$$

is a continuous homomorphism of linear groups; and suppose H is connected. Then

$$\boxed{\mathcal{L}F \text{ surjective} \iff F \text{surjective}.}$$

Proof ► Suppose first that $f = \mathcal{L}F$ is surjective; and suppose $T \in H$. Then by Proposition 3.4,

$$T = e^{Y_1} \dots e^{Y_r},$$

where $Y_1, \dots, Y_r \in \mathcal{L}H$. Since f is surjective we can find $X_1, \dots, X_r \in \mathcal{L}G$ such that

$$Y_i = f(X_i) \quad (i = 1, \dots, r)$$

Then

$$\begin{aligned} T &= e^{fX_1} \dots e^{fX_r} \\ &= F(e^{X_1}) \dots F(e^{X_r}) \\ &= F(e^{X_1} \dots e^{X_r}). \end{aligned}$$

Thus $T \in \operatorname{im} F$; and so F is surjective.

Now suppose conversely that F is surjective; and suppose $Y \in \mathcal{L}H$. We must show that there exists an $X \in \mathcal{L}G$ such that

$$fX = Y.$$

This is much more difficult.

Since f is linear, we can suppose Y so small that the line-segment

$$[0, 1]Y = \{tY : 0 \le t \le 1\}$$

lies inside the logarithmic zone in $\mathcal{L}H$. For the same reason it suffices to find X such that

$$fX = tY$$

for some non-zero t.

Our proof falls into 2 parts.

1: An enumerability argument. Let

$$C = F^{-1}(e^{[0,1]Y}) = \{T \in G : FT = e^{tY},\ 0 \le t \le 1\}.$$

We shall show that I is not isolated in C. If it were then every point of C would be isolated. For suppose $S, T \in C$ were close to one another; and suppose

$$S = e^{sY}, \quad T = e^{tY},$$

with $s < t$. Then

$$R = S^{-1}T$$

would be close to I; and it is in C since $F(R) = e^{(t-s)Y}$.

Thus if I were isolated in C we could find disjoint open subsets of $\mathbf{M}(n, \mathbb{R})$ surrounding each element of C. But C is non-enumerable, since it contains at least one point corresponding to each $t \in I$. Thus—always supposing I isolated in C—we would have a non-enumerable family of disjoint open sets in $\mathbf{M}(n, \mathbb{R}) = \mathbb{R}^N$ (where $N = n^2$). But that leads to a contradiction. For each subset will contain a rational point (ie a point with rational coordinates); and the number of rational points is enumerable. We conclude that I cannot be isolated in C.

2: A logarithmic argument. Since I is not isolated, we can find $T \in C$ arbitrarily close to I. In particular we can choose T so that

1. T lies in the logarithmic zone of G, with say

$$X = \log T;$$

2. fX lies in the logarithmic zone of $\mathcal{L}H$.

But then

$$FT = F(e^X) = e^X,$$

while on the other hand

$$FT = e^{tY}$$

for some $t \in [0, 1]$. Thus

$$e^{fX} = e^{tY}.$$

Since fX and tY both lie in the logarithmic zone, it follows that

$$fX = tY,$$

as required.

◄

The Corollary to Proposition 4 and Proposition 5 together give

Proposition 6.6 *Suppose*

$$F : G \to H$$

is a continuous homomorphism of linear groups; and suppose H is connected. Then

$$\boxed{\mathcal{L}F \textit{bijective} \iff \ker(F) \textit{discrete and } F \textit{surjective}.}$$

Suppose G, H, K are 3 linear groups; and suppose we are given homomorphisms

$$\alpha : G \to H, \ \beta : H \to K.$$

Recall that the sequence

$$G \to H \to K$$

is said to be *exact* if

$$\operatorname{im} \alpha = \ker \beta.$$

Proposition 6.7 *An exact sequence*

$$G \to H \to K$$

of linear groups yields a corresponding exact sequence

$$\mathcal{L}G \to \mathcal{L}H \to \mathcal{L}K$$

of Lie algebras.

Proof ► This follows at once from Propositions 6.4 and 6.5 above. ◄ *Remark:* To summarise our last 3 results: if H is connected, then

1. $\mathcal{L}F$ injective $\iff \ker F$ discrete;
2. $\mathcal{L}F$ surjective $\iff \operatorname{im} F$ open;
3. $\mathcal{L}F$ bijective $\iff \ker F$ discrete and Fsurjective.

Examples:

1. As we noted above, the functorial property of the Lie operator allows the theory to be extended from linear to *linearisable* groups.

   Consider for example the real projective group

$$\mathbf{PGL}(n, \mathbb{R}) = \mathbf{GL}(n+1, \mathbb{R})/\mathbb{R}^{\times},$$

ie the group of projective transformations

$$P(T) : \mathbf{PR}^n \to \mathbf{PR}^n$$

of n-dimensional real projective space $\mathbf{PR}^n$—that is, the space of 1-dimensional subspaces (or rays) in the $(n+1)$-dimensional vector space $\mathbb{R}^{n+1}$. Each non-singular linear map

$$T : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$$

defines a projective transformation

$$P(T) : \mathbf{PR}^n \to \mathbf{PR}^n :$$

two linear maps defining the same projective transformation if and only if each is a scalar multiple of the other, ie

$$P(T) = P(T') \iff T' = aT.$$

As it stands, $\mathbf{PGL}(n, \mathbb{R})$ is not a linear group. However, we can 'linearise' it in the following way. Suppose $T \in \mathbf{GL}(n, \mathbb{R})$. Consider the linear map

$$X \mapsto TXT^{-1} : \mathbf{M}(n+1, \mathbb{R}) \to \mathbf{M}(n+1, \mathbb{R}).$$

It is evident that any scalar multiple aT will define the same linear map. Conversely, suppose $T, U \in \mathbf{GL}(n+1, \mathbb{R})$ define the same map, ie

$$TXT^{-1} = UXU^{-1} \quad \text{for all } X \in \mathbf{M}(n+1, \mathbb{R}).$$

Let

$$V = T^{-1}U.$$

Then

$$VXV^{-1} = X,$$

ie

$$VX = XV \quad \forall X.$$

It follows that

$$V = aI,$$

ie

$$U = aT.$$

Thus we have defined an injective homomorphism

$$\Theta : \mathbf{PGL}(n, \mathbb{R}) \to \mathbf{GL}(N, \mathbb{R}),$$

where $N = (n+1)^2$, identifying the projective group with a linear group $G \subset \mathbf{GL}(N, \mathbb{R})$.

This identifies $\mathbf{PGL}(n, \mathbb{R})$ with a subgroup of $\mathbf{GL}(N, \mathbb{R})$. But that is not quite sufficient for our purposes; we must show that the subgroup is *closed*.

(The image of an injective homomorphism certainly need not be closed. Consider for example the homomorphism

$$\mathbb{R} \to \mathbb{T}^2 : t \mapsto (a \bmod 1, b \bmod 1)$$

from the reals to the 2-dimensional torus. The image is a path going round the torus like a billiard ball round a billiard table. There are 2 possibilities.

If the ratio $a : b$ is *rational*, the path runs repeatedly over the same closed subgroup of the torus. In particular, the homomorphism is not injective,

On the other hand, if the ratio $a : b$ is *irrational*, then the path will never return to the origin. The homomorphism is injective, and its image passes arbitrarily close to every point of the torus. Thus the closure of the image is the whole of the torus. However, it does not pass through every point of the torus, since eg it will only cut the 'circle' $(0, y \bmod 1)$ enumerably often. So the image group in this case is not closed.)

There are several ways of showing that $\operatorname{im}\Theta$ *is* closed. For example, we shall see later that since $\mathbf{SL}(n+1, \mathbb{R})$ is *semisimple* its image is necessarily closed.

But perhaps the simplest approach is to *identify* the subgroup $\operatorname{im}\Theta$. To this end, observe that for each $X \in \mathbf{GL}(n+1, \mathbb{R})$, the map

$$\alpha_T : X \mapsto T^{-1}XT : \mathbf{M}(n+1, \mathbb{R}) \to \mathbf{M}(n+a, \mathbb{R})$$

is an *automorphism* of the algebra $\mathbf{M}(n+1, \mathbb{R})$, ie it preserves mutliplication as well as addition and scalar multiplication. It also preserves the trace.

We shall show that every such automorphism is of this form, ie

$$\operatorname{im}\Theta = \operatorname{Aut}\mathbf{M}(n+1, \mathbb{R}).$$

Since the property of being a trace-preserving automorphism of $\mathbf{M}(n+1, \mathbb{R})$ can be defined by algebraic equations (albeit an infinity of them) the automorphism group is closed in $\mathbf{GL}(N, \mathbb{R})$, so if we can prove that it is in fact $\operatorname{im}\Theta$ we shall have achieved our objective.

Again, there are several ways of proving this. One approach, which might appeal to the geometrically-minded but which we shall not pursue, starts by noting that we can represent subspaces of the projective space $\mathbf{PR}^n$ by projections (idempotents) in $\mathbf{M}(n{+}1, \mathbb{R})$. An automorphism of $\mathbf{M}(n{+}1, \mathbb{R})$ then defines a transformation

$$\mathbf{PR}^n \to \mathbf{PR}^n$$

which sends subspaces to subspaces of the same dimension, in such a way that all incidence relations are preserved. Now it is a well-known proposition of projective geometry that such a transformation is necessarily projective, ie it can be defined by a linear transformation $T \in \mathbf{GL}(n+1, \mathbb{R})$. But one could say that this approach is somewhat masochistic, since we have thrown away a good deal of information in passing from an automorphism of $\mathbf{M}(n{+}1, \mathbb{R})$ to the corresponding tranformation of projective space $\mathbf{PR}^n$.

The following proof may not be the simplest, but it has the advantage of being based on ideas from the representation theory of finite groups, studied in Part 1.

Let F be any finite or compact group having an absolutely simple $n+1$-dimensional representation; that is, one that remains simple under complexification.

For example, we saw in Part 1 that the symmetric group S_{n+2} has such a representation—its natural representation θ in $\mathbb{R}^{n+2}$ splitting into 2 absolutely simple parts

$$\theta = 1 + \sigma$$

where σ is the representation in the $(n+1)$-dimensional subspace

$$x_1 + \cdots + x_{n+2} = 0.$$

We can turn this on its head and say that we have a finite subgroup F of $\mathbf{GL}(n+1, \mathbb{R})$ whose natural representation ν in $\mathbb{R}^n$ is simple.

Now suppose

$$\alpha : \mathbf{M}(n+1, \mathbb{R}) \to \mathbf{M}(n+1, \mathbb{R})$$

is a trace-preserving automorphism of the algebra $\mathbf{M}(n+1, \mathbb{R})$. (Actually every automorphism of $\mathbf{M}(n+1, \mathbb{R})$ is trace-preserving. But it is easier for our purposes to add this condition than to prove it.) We want to show that α is an *inner* automorphism, ie of the form

$$\alpha = \alpha_T$$

for some $T \in \mathbf{GL}(n+1, \mathbb{R})$.

To this end, note that the composition

$$\alpha\nu : F \to \mathbf{GL}(n+1, \mathbb{R})$$

yields a second representation of F in $\mathbb{R}^{n+1}$.

The 2 representations $\nu, \alpha\nu$ haves the same character (since α preserves trace). Hence they are equivalent, ie there exists a $T \in \mathbf{GL}(n+1, \mathbb{R})$ such that

$$\alpha\nu(g) = \alpha_T \nu(g)$$

for all $g \in F$.

It follows that

$$\alpha(X) = \alpha_T(X)$$

for all $X \in \mathbf{M}(n+1, \mathbb{R})$ in the matrix algebra generated by the matrices of F, ie all X expressible as linear combinations of the elements of F:

$$X = a_1 F_1 + \cdots + a_m F_m.$$

But we saw in Part 1 that every matrix $X \in \mathbf{M}(n+1, \mathbb{R})$ is of this form, ie the matrices in an absolutely simple representation span the matrix space. (For a simple proof of this, consider the representation $\nu^* \times \nu$ of the product-group $F \times G$ in the space of matrices $\mathbf{M}(n+1, \mathbb{R})$. We know that this representation is simple. Hence the matrices

$$\nu(f_1)^{-1} I \nu(f_2) = f_1^{-1} f_2$$

span the representation space $\mathbf{M}(n+1, \mathbb{R})$.) Thus

$$\alpha(X) = \alpha_T(X)$$

for *all* $X \in \mathbf{M}(n+1, \mathbb{R})$, ie $\alpha = \alpha_T$.

So we have identified $\mathbf{PGL}(n, \mathbb{R})$ with the group of trace-preserving automorphisms of $\mathbf{M}(n+1, \mathbb{R})$. Since the property of being an automorphism, and of preserving the trace, can be defined by polynomial equations (albeit an infinity of them), this automorphism group is a *closed* subgroup of $\mathbf{GL}(N, \mathbb{R})$, as required.

Hence $\mathbf{PGL}(n, \mathbb{R})$ is *linearisable*, and we can speak of its Lie algebra $\mathbf{pgl}(n, \mathbb{R})$.

By definition,

$$\mathbf{pgl}(n, \mathbb{R}) = \mathcal{L}G,$$

where G is the linear group above. However, we need not compute $\mathcal{L}G$ explicitly; there is a much simpler way of determining $\mathbf{pgl}(n, \mathbb{R})$.

Consider the homomorphism

$$T \mapsto P(T) : SL(n+1, \mathbb{R}) \to \mathbf{PGL}(n, \mathbb{R}).$$

If n is even then P is in fact an isomorphism. For on the one hand

$$\begin{aligned} T \in \ker P &\implies T = aI \\ &\implies a^n = 1 \\ &\implies a = 1 \end{aligned}$$

since $a \in \mathbb{R}$ and $n+1$ is odd; while on the other hand if $P = P(T)$ is a projective transformation we can suppose that $T \in SL(n+1, \mathbb{R})$, since we can always find a scalar a such that

$$\det aT = a^n \det T = 1.$$

Thus if n is even,

$$\mathbf{PGL}(n, \mathbb{R}) = \mathbf{SL}(n+1, \mathbb{R}).$$

So it is evident in this case that $\mathbf{PGL}(n, \mathbb{R})$ is linearisable—we don't need the long-winded argument above—and that

$$\mathbf{pgl}(n, \mathbb{R}) = (n+1, \mathbb{R}).$$

In fact this last result still holds if n is odd. For in that case the homomorphism

$$\Theta : \mathbf{SL}(n+1, \mathbb{R}) \to \mathbf{PGL}(n, \mathbb{R})$$

is not bijective. However, it has kernel $\pm I$; and its image is the subgroup of $\mathbf{PGL}(n, \mathbb{R})$ of index 2, consisting of those projective transformations $P(T)$ defined by $T \in \mathbf{GL}(n+1, \mathbb{R})$ with $\det T > 0$. We can summarise this in the exact sequence

$$1 \to C_2 \to \mathbf{SL}(n+1, \mathbb{R}) \to \mathbf{PGL}(n, \mathbb{R}) \to C_2 \to 1.$$

But now Proposition 6.7 above tells us that the corresponding Lie algebra homomorphism

$$\mathcal{L}\Theta : \mathbf{sl}(n+1, \mathbb{R}) \to \mathbf{pgl}(n, \mathbb{R})$$

is in fact an isomorphism.

So we find that in all cases

$$\mathbf{pgl}(n, \mathbb{R}) = \mathbf{sl}(n+1, \mathbb{R}).$$

2. Now consider the Euclidean group $E(2)$, ie the isometry group of the Euclidean plane E^2. As it stands, $E(2)$ is not a linear group. However, on choosing coordinates, we can identify E^2 with $\mathbb{R}^2$; and $E(2)$ can then be identified with the group of transformations

$$E : (x, y) \mapsto (ax + cy + e, bx + dy + f)$$

where

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathbf{O}(2).$$

We can extend this transformation E to a projective transformation

$$P(T) : \mathbf{PR}^2 \to \mathbf{PR}^2,$$

namely that defined by the matrix

$$T = \begin{pmatrix} a & c & e \\ b & d & f \\ 0 & 0 & 1 \end{pmatrix}$$

This defines an injective homomorphism

$$F : E(2) \to \mathbf{PGL}(2, \mathbb{R}),$$

allowing us to identify $E(2)$ with a closed subgroup of the 2-dimensional projective group. Since we have already established that the projective group can be linearised, it follows that $E(2)$ can be also.

Explicitly, we have identified $E(2)$ with the group G of 3×3 matrices described above. By definition,

$$e(2) = \mathcal{L}G.$$

To determine $\mathcal{L}G$, we adopt our usual technique. Suppose

$$X = \begin{pmatrix} p & u & x \\ q & v & y \\ r & w & z \end{pmatrix} \in \mathcal{L}G.$$

Then

$$I + tX = \begin{pmatrix} 1 + tp & tu & tx \\ tq & 1 + tv & ty \\ tr & tw & 1 + tz \end{pmatrix}$$

satisfies the conditions on G to the first order. Hence

$$p = r = v = w = z = 0, q + u = 0,$$

ie

$$X = \begin{pmatrix} 0 & q & x \\ -q & 0 & y \\ 0 & 0 & 0 \end{pmatrix}$$

Conversely, it is readily verified that if X is of this form then

$$e^X \in G.$$

The space $\mathcal{L}G$ has basis

$$L = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

By computation,

$$[L, M] = N, \quad [L, N] = -M, \quad [M, N] = 0.$$

Thus

$$e(2) = \langle L, M, N : [L, M] = N, [L, N] = -M, [M, N] = 0 \rangle.$$

Chapter 7

Representations

> Since a representation is—from one point of view—just a particular kind of homomorphism, Lie theory certainly applies. But there is one small problem: the Lie algebra of a linear group is *real*, while we are interested almost exclusively in *complex* representations. Overcoming this problem brings an unexpected reward, by disclosing a surprising relation between apparently unrelated groups. This allows us to extend the representation theory of compact groups to a much wider class of linear groups.

Definition 7.1 *A representation of a Lie algebra $\mathcal{L}$ over k (where $k = \mathbb{R}$ or $\mathbb{C}$) in the vector space V over k is defined by a bilinear map*

$$\mathcal{L} \times V \to V,$$

which we denote by $(X, v) \to Xv$, satisfying

$$\boxed{[X, Y]v = X(Yv) - Y(Xv).}$$

Remark: Notice that we only consider *real* representations of *real* Lie algebras, or *complex* representations of *complex* algebras—we do not mix our scalars. This might seem puzzling, since:

1. we are primarily interested (as always) in *complex* representations of linear groups; but
2. the Lie algebra of a linear group is always *real*.

The explanation is found in the following Definition and Proposition.

Definition 7.2 *Suppose $\mathcal{L}$ is a real Lie algebra. We denote by $\mathbb{C}\mathcal{L}$ the complex Lie algebra derived from $\mathcal{L}$ by extension of the scalars (from $\mathbb{R}$ to $\mathbb{C}$).*

Remark: Suppose $\mathcal{L}$ has structure constants c_{ij}^k with respect to the basis $e_1, ..., e_n$, ie

$$[e_i, e_j] = \sum_k c_{ij}^k e_k.$$

Then we can take the same basis and structure constants for $\mathbb{C}\mathcal{L}$.

Proposition 7.1 1. *Each real representation α of the linear group G in U gives rise to a representation $\mathcal{L}\alpha$ of the corresponding Lie algebra $\mathcal{L}G$ in U, uniquely characterised by the fact that*

$$\boxed{\alpha(e^X) = e^{\mathcal{L}\alpha X} \quad \forall X \in \mathcal{L}G.}$$

2. *Each complex representation α of G in V gives rise to a representation $\mathcal{L}\alpha$ of the complexified Lie algebra $\mathbb{C}\mathcal{L}G$ in V, uniquely characterised by the fact that*

$$\boxed{\alpha(e^X) = e^{\mathcal{L}\alpha X} \quad \forall X \in \mathcal{L}G.}$$

In either case, if G is connected *then α is uniquely determined by $\mathcal{L}\alpha$, ie*

$$\boxed{\alpha = \beta \Longrightarrow \mathcal{L}\alpha = \mathcal{L}\beta.}$$

Proof ► The real and complex representations of G are defined by homomorphisms

$$G \to \mathbf{GL}(n, \mathbb{R}),\ G \to \mathbf{GL}(n, \mathbb{C}).$$

These in turn give rise to Lie algebra homomorphisms

$$\mathcal{L}G \to \mathbf{gl}(n, \mathbb{R}),\ \mathcal{L}G \to \mathbf{gl}(n, \mathbb{C}).$$

The first of these defines the required representation of $\mathcal{L}G$ by

$$(X, u) \to Xu.$$

The second needs a little more care.

Note first that $\mathbf{gl}(n, \mathbb{C}) = \mathbf{M}(n, \mathbb{C})$ has a natural structure as a complex Lie algebra. If we use $\mathbf{M}(n, \mathbb{C})$ to denote this algebra then

$$\mathbf{gl}(n, \mathbb{C}) = \mathbb{R}\mathbf{M}(n, \mathbb{C}),$$

where the "realification" $\mathbb{R}\mathcal{L}$ of a complex Lie algebra $\mathcal{L}$ is defined in the usual forgetful way.

Recall the following simple (if confusing) result from linear algebra. If U is a real vector space and V a complex vector space then each real linear map

$$F : U \to \mathbb{R}V$$

extends—uniquely—to a complex linear map

$$F : \mathbb{C}U \to V.$$

Applying this with $U = \mathcal{L}G$ and $V = \mathbf{M}(n, \mathbb{C})$, the *real* Lie algebra homomorphism

$$\mathcal{L}G \to \mathbf{gl}(n, \mathbb{C})$$

yields a *complex* Lie algebra homomorphism

$$\mathbb{C}\mathcal{L}G \to \mathbf{M}(n, \mathbb{C}).$$

This defines the required representation of $\mathbb{C}\mathcal{L}G$ by

$$(X, v) \to Xv.$$

Finally, if G is connected then each $g \in G$ is expressible as a product

$$g = e^{X_1} \dots e^{X_r}.$$

But from the equation defining the relation between α and $\mathcal{L}\alpha$ the action of e^X on U or V is defined by the action of X. Hence the action of $\mathcal{L}G$ or $\mathbb{C}\mathcal{L}G$ completely determines the action of each $g \in G$. ◀

Remark: By "abuse of notation" we shall call a representation of $\mathbb{C}\mathcal{L}$ (where $\mathcal{L}$ is a real Lie algebra) a complex representation of $\mathcal{L}$. With this understanding Proposition 1 can be summarised as follows:

Each representation α of $\mathcal{L}$ defines a corresponding representation $\mathcal{L}\alpha$ of $\mathcal{L}G$ in the same space. Moreover if G is connected then $\mathcal{L}\alpha$ uniquely determines α, ie

$$\alpha = \beta \Longrightarrow \mathcal{L}\alpha = \mathcal{L}\beta.$$

Corollary 7.1 *Suppose $\mathcal{L}$ and $\mathcal{L}'$ are 2 real Lie algebras. Then an isomorphism between their complexifications*

$$\mathbb{C}\mathcal{L} \longleftrightarrow \mathbb{C}\mathcal{L}'$$

sets up a $1-1$ correspondence between the complex representations of $\mathcal{L}$ and $\mathcal{L}'$.

Example: We shall show that, for each n, the Lie algebras

$$\mathbf{sl}(n, \mathbb{R}) = \langle X \in \mathbf{M}(n, \mathbb{R}) : \operatorname{tr} X = 0 \rangle$$

and

$$\mathbf{su}(n) = \langle X \in \mathbf{M}(n, C) : X^* = X, \operatorname{tr} X = 0 \rangle$$

have the same complexification:

$$\boxed{\mathbb{C}\mathbf{sl}(n, \mathbb{R}) = \mathbb{C}\mathbf{su}(n).}$$

Certainly these 2 algebras have the same dimension:

$$\dim \mathbf{sl}(2, \mathbb{R}) = \dim \mathbf{su}(n) = n^2 - 1.$$

Moreover, we can regard both $\mathbf{sl}(n, \mathbb{R})$ and $\mathbf{su}(n)$ as real subalgebras of the *complex* Lie algebra $\mathbf{M}(n, \mathbb{C})$; and the injections

$$\mathbf{sl}(n, \mathbb{R}) \to \mathbf{M}(n, \mathbb{C}), \mathbf{su}(n) \to \mathbf{M}(n, \mathbb{C})$$

define complex Lie algebra homomorphisms

$$\Phi : \mathbb{C}\mathbf{sl}(n, \mathbb{R}) \to \mathbf{M}(n, \mathbb{C}), \Psi : \mathbb{C}\mathbf{su}(n) \to \mathbf{M}(n, \mathbb{C}).$$

It is not obvious a priori that Φ and Ψ are injections. However, that will follow if we can establish that

- $\operatorname{im} \Phi = \operatorname{im} \Psi$;
- $\dim_{\mathbb{C}} \operatorname{im} \Phi = n^2 - 1.$

Indeed, this will also prove the desired result

$$\mathbb{C}\mathbf{sl}(n, \mathbb{R}) = \operatorname{im} \Phi = \mathbb{C}\mathbf{su}(n).$$

But $\operatorname{im} \Phi$ is just the complex linear hull of $\mathbf{sl}(n, \mathbb{R})$ in $\mathbf{M}(n, \mathbb{C})$, ie the subspace formed by the linear combinations, with complex coefficients, of the elements of $\mathbf{sl}(n, \mathbb{R})$; and similarly $\operatorname{im} B$ is the complex hull of $\mathbf{su}(n)$.

But it is easy to see that these hulls are both equal to the complex subspace

$$V = \{X \in \mathbf{M}(n, \mathbb{C}) : \operatorname{tr} X = 0\}.$$

For $\mathbf{sl}(n, \mathbb{R})$ this is a consequence of the elementary proposition that the complex solutions of a real linear equation are just the linear combinations, with complex coefficients, of real solutions.

For $\mathbf{su}(n)$, the result follows on noting that any element $X \in V$ can be written as

$$X = Y + iZ,$$

with

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

both in $\mathbf{su}(n)$.

In conclusion, since

$$\dim_{\mathbb{C}} V = n^2 - 1$$

we have established that

$$\mathbb{C}\mathbf{sl}(n, \mathbb{R}) = \mathbb{C}\mathbf{su}(n).$$

As a concrete illustration, consider the case $n = 2$. We have seen that

$$\mathbf{sl}(n, \mathbb{R}) = \{X \in \mathbf{M}(2, \mathbb{R}) : \operatorname{tr} X = 0\} = \langle H, E, F \rangle,$$

with

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad E = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad F = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

while

$$\mathbf{su}(n) = \{X \in \mathbf{M}(2, \mathbb{C}) : X^* = X, \operatorname{tr} X = 0\} = \langle A, B, C \rangle,$$

with

$$A = \begin{pmatrix} i & 0 \\ 0 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

The isomorphism between the complexifications is defined by

$$A \longleftrightarrow iH,\ B \longleftrightarrow (E - F),\ C \longleftrightarrow i(E + F).$$

All the basic notions of group representation theory have Lie algebra analogues. These are summarised in

Definition 7.3 1. *The representation α of $\mathcal{L}$ in V is said to be simple if there are no non-trivial stable subspaces, ie if U is a subspace of V such that*

$$X \in \mathcal{L}, u \in U \Longrightarrow Xu \in U$$

then $U = 0$ or V.

2. *The sum $\alpha + \beta$ of 2 representations α and β of $\mathcal{L}$ in U and V is the representation in the direct sum $U \oplus V$ defined by*

$$X(u + v) = Xu + Xv.$$

3. *The product* $\alpha\beta$ *of* α *and* β *is the representation in the tensor product* $U \otimes V$ *defined by*
$$X(u \otimes v) = Xu \otimes v + u \otimes Xv.$$

4. *The dual* α^* *of* α *is the representation in* V^* *defined by*
$$(X, p) \to -X'p.$$

5. *A representation is said to be semisimple if it is expressible as a sum of simple representations.*

Proposition 7.2 *Suppose* α, β *are representations of the connected linear group* G. *Then*

1. α *simple* $\iff \mathcal{L}\alpha$ *simple*

2. $L(\alpha + \beta) = \mathcal{L}\alpha + \mathcal{L}\beta$

3. $\mathcal{L}(\alpha\beta) = (\mathcal{L}\alpha)(\mathcal{L}\beta)$

4. $\mathcal{L}(\alpha^*) = (\mathcal{L}\alpha)^*$

5. α *semisimple* $\iff \mathcal{L}\alpha$ *semisimple*

Chapter 8

Simply Connected Linear Groups

To each homomorphism $F : G \to H$ of linear groups the Lie functor associates a homomorphism$f = \mathcal{L}F : \mathcal{L}G \to \mathcal{L}H$ of the corresponding Lie algebras. But which Lie algebra homomorphisms arise in this way? Which can be *lifted* to group homomorphisms? This question is of a rather different nature to those we have been considering. For while Lie theory is a *local* theory, this is a *global* question. Indeed, every Lie algebra homomorphism can be lifted *locally*. The question is: do these local bits fit together? That depends, as we shall see, on the *fundamental group* (or first homotopy group) $\pi_1(G)$ of the linear group G. If this homotopy group is trivial—that is, G is *simply-connected* then every Lie algebra homomorphism $f : \mathcal{L}G \to \mathcal{L}H$ *can* be lifted.

Proposition 8.1 *Suppose G and H are linear groups; and suppose G is simply connected. Then every Lie algebra homomorphism*

$$f : \mathcal{L}G \to \mathcal{L}H$$

can be lifted to a unique group homomorphism

$$F : G \to H$$

such that

$$\boxed{\mathcal{L}F = f.}$$

Remark: Recall that a topological space X is said to be *simply connected* if it is arcwise connected and if in addition every loop in X can be shrunk to a point, ie every continuous map

$$u : S^1 \to X$$

from the circumference S^1 of the 2-ball B^2 (the circle with its interior) can be extended to a continuous map

$$u : B^2 \to X$$

from the whole ball.

This is equivalent to saying that the first homotopy group (or fundamental group) of X is trivial:

$$\pi_1(X) = \{e\}.$$

Proof ▶ Since G is connected, each $T \in G$ is expressible in the form

$$T = e^{X_r} \dots e^{X_1},$$

by Proposition 3.4. If F exists, $F(T)$ must be given by

$$\begin{aligned} F(T) &= F(e^{X_r}) \dots F(e^{X_1}) \\ &= e^{fX_r} \dots e^{fX_1} \end{aligned}$$

In particular, the existence of F requires that

$$\boxed{e^{X_r} \dots e^{X_1} = I \Longrightarrow e^{fX_r} \dots e^{fX_1} = I.}$$

Conversely, if this condition is satisfied, then $F(T)$ is defined unambiguously by (*); and the map $F : G \to H$ defined in this way is clearly a homomorphism with $\mathcal{L}F = f$.

It is sufficient therefore to show that condition (**) is always satisfied. This we do in 2 stages.

1. First we show that condition (**) is always satisfied *locally*, ie for sufficiently small $X_1, \dots, X_r$. This does not require that G be simply-connected, or even connected. We may say that f always lifts to a *local* homomorphism, defined on a neighbourhood of $I \in G$.

2. Then we show that if G is simply-connected, every local homomorphism can be extended to the whole of G.

These 2 stages are covered in the 2 lemmas below. But first we see how relations on linear groups, like those in (**) above, can be represented by closed paths, or *loops*.

Let us call a path on G of the form

$$I = [0, 1] \to G : t \mapsto e^{tX} g$$

an *exponential path* joining g to $e^X g$. Then the relation

$$e^{X_r} \dots e^{X_1} = I$$

defines an *exponential loop* starting at any point $g \in G$, with vertices

$$g_0 = g, g_1 = e^{X_1} g, g_2 = e^{X_2} e^{X_1} g, \dots, g_r = e^{X_r} \dots e^{X_1} g = g,$$

each successive pair g_{i-1}, g_i of vertices being joined by the exponential path

$$e^{tX_i} g_{i-1} \quad (0 \le t \le 1).$$

(If we chose a different starting-point, we would get a "congruent" loop, ie a transform Pg of the first loop P by a group element g. In effect, we are only concerned with paths or loops "up to congruence".)

Each exponential path

$$e^{tX} g$$

in G defines an exponential path in H, once we have settled on a starting point h:

$$e^{tX} g \mapsto e^{tfX} h.$$

More generally, each path in G made up of exponential segments—let us call it a "piecewise-exponential" path—maps onto a piecewise-exponential path in H, starting from any given point.

In this context, condition (**) becomes: Every loop in G maps into a loop in H. Or: if a path in G closes on itself, then so does its image in H.

Similarly, the local version reads: Every sufficiently small loop in G maps into a loop in H. It is sufficient in this case to consider exponential *triangles*, made of 3 exponential paths. For the loop

$$P_r P_{r-1} \dots P_0$$

can be split into the triangles

$$P_0 P_r P_{r-1}, P_0 P_{r-1} P_{r-2}, \dots, P_0 P_2 P_1.$$

In algebraic terms, given a relation

$$e^{X_r} \dots e^{X_1} = I,$$

we set

$$e^{Y_i} = e^{X_i} e^{X_{i-1}} \dots e^{X_1} \quad 1 \le i \le r,$$

with

$$e^{Y_r} = e^{Y_0} = I;$$

the given relation then splits into the triangular relations

$$e^{Y_i} = e^{X_i} e^{Y_{i-1}},$$

ie

$$e^{-Y_i} e^{X_i} e^{Y_{i-1}} = I.$$

If each of these relations is preserved by F, then the product of their images in H gives the required relation:

$$e^{fX_r} \ldots e^{fX_0} = (e^{-fY_r} e^{X_r} e^{Y[r-1}) \ldots (e^{-fY_1} e^{X_1} e^{Y_0}) = I.$$

Lemma 8.1 *Suppose G and H are linear groups; and suppose*

$$f : \mathcal{L}G \to \mathcal{L}H$$

is a Lie algebra homomorphism. Then the map

$$e^X \mapsto e^{fX}$$

is a local homomorphism, ie there is a constant $C > 0$ such that if $|X|, |Y|, |Z| < C$ then

$$\boxed{e^X e^Y = e^Z \Longrightarrow e^{fX} e^{fY} = e^{fZ}.}$$

Proof of Lemma ▷ This falls into 2 parts:

1. 1. If the triangle is small of size (or side) d then the 'discrepancy' in its image, ie the extent to which it fails to close, is of order d^3.
2. 2. A reasonably small triangle, eg one lying within the logarithmic zone, can be divided into n^2 triangles, each of $(1/n)$th the size. Each of these maps into a 'near-triangle' with discrepancy of order $1/n^3$. These sum to give a total discrepancy of order $n^2/n^3 = 1/n$. Since n can be taken arbitrarily large the discrepancy must in fact be 0.

1. If

   $$|X|, |Y|, |Z| < d$$

   then

   $$e^X e^Y = e^Z \Longrightarrow Z = \log(e^X e^Y).$$

   But

$$\begin{aligned} e^X e^Y &= (I + X + X^2/2)(I + Y + Y^2/2) + O(d^3) \\ &= I + X + Y + XY + X^2/2 + Y^2/2 + O(d^3). \end{aligned}$$

Hence

$$\begin{aligned} Z &= \log(e^X e^Y) \\ &= X + Y + XY + X^2/2 + Y^2/2 - (X+Y)^2/2 + O(d^3) \\ &= X + Y + [X, Y]/2 + O(d^3). \end{aligned}$$

Since by hypothesis f preserves the Lie product it follows that

$$fZ = fX + fY + [fX, fY]/2 + O(d^3).$$

Hence (working backwards)

$$e^{fX} e^{fY} = e^{fZ} + O(d^3).$$

2. We have not yet said exactly what we mean by the discrepancy $D(R)$ of a relation R on G. It is convenient to define $D(R)$ to be the sup-norm of the gap in H:

$$D(R) = |F(R)|_0,$$

where

$$|T|_0 = \sup_{x \neq 0} \frac{|Tx|}{|x|}.$$

This has the advantage, for our present purpose, that

$$|P^{-1}TP|_0 = |T|_0.$$

This means that if we take a different vertex on a loop as starting-point, or equivalently, take a conjugate form

$$e^{X_i} e^{X_{i-1}} \ldots e^{X_1} e^{X_r} \ldots e^{X_{r+1}} = I$$

of the given relation, the discrepancy remains the same.

It is clear that any triangle within the logarithmic zone can be shrunk to a point within that zone. In fact, such a deformation can be brought about by a sequence of 'elementary deformations', each consisting either of replacing the path round 2 sides of a small exponential triangle by the 3rd side, or conversely replacing 1 side by the other 2.

More precisely, if we start with a triangle of side $< d$ then this can be divided into n^2 exponential triangles of side $< d/n$; and so the original triangle can be shrunk to a point by at most n^2 elementary deformations, each involving an exponential triangle of side $< d/n$.

Actually, we really need the converse constuction. Starting from the trivial relation $I = I$, we can re-construct the original triangle by a sequence of elementary deformations, each involving a small triangle of side $< d/n$.

The discrepancy caused by each of these will be of order $1/n^3$ by (1). This perhaps calls for clarification. We can assume that each of these deformations involves the last edge of the loop, since the deformation is unaltered if we change our base-vertex. Suppose then the relation

$$R = e^{X_r} \dots e^{X_1} = I$$

is "deformed" into

$$e^Y e^Z e^{X_{r-1}} \dots e^{X_1},$$

where

$$e^Y e^Z = e^{X_r}.$$

The new relation can be written

$$TR = (e^Y e^Z e^{-X_r})(e^{X_r} \dots e^{X_1}) = I.$$

The corresponding gap in H is

$$F(TR) - I = F(R)(F(T) - I) + (F(R) - I);$$

and so

$$D(TR) \leq |F(R)|_0 D(T) + D(R).$$

This shows (by induction) both that $F(R)$ is bounded, and that the descrepancy changes at each deformation by an amount of order $1/n^3$.

In sum therefore the n^2 deformations will cause a change in the discrepancy of order $1/n$, ie an arbitrarily small change. Since the discrepancy was initially 0 (with the triangle shrunk to a point), it must be 0 finally, ie

$$\boxed{e^X e^Y = e^Z \Longrightarrow e^{fX} e^{fY} = e^{fZ}.}$$

◁

Corollary 8.1 *There is an open neighbourhood U of I in G such that every exponential loop in U maps into a loop in H.*

Lemma 8.2 *Suppose G and H are linear groups; and suppose G is simply connected. Then every local homomorphism*

$$U \to H,$$

where U is a neighbourhood of I in G, has a unique extension to the whole of G.

Proof of Lemma ▷ Any exponential path

$$e^{tX}g$$

can be split into arbitrarily small sub-paths, with vertices

$$g, e^{X/m}g, e^{2X/m}g, \ldots, e^{X}g.$$

It follows that we can always suppose the edges of the various loops we consider small enough to come within the ambit of Lemma 1.

Any loop can be shrunk to a point, by hypothesis. We can suppose this shrinkage occurs as a sequence of small shrinks,

$$P_r \ldots P_0 \mapsto P'_r \ldots P'_0$$

say. These steps can in turn be split into a number of sub-steps, each involving a small rectangular deformation, of the form

$$P_r \ldots P_{i+1}P_iP'_i \ldots P'_0P_0 \to P_r \ldots P_{i+1}P'_{i+1}P'_i \ldots P'_0P_0$$

Since by the Corollary to Lemma 1 each such deformation leaves the discrepancy unchanged, and since the discrepancy finally vanishes, it must vanish initially, ie

$$\boxed{e^{X_1} \ldots e^{X_r} = I \Longrightarrow e^{fX_1} \ldots e^{fX_r} = I.}$$

◁ ◀

Proposition 8.2 *Suppose the linear group* G is simply connected. *Then every representation* α *of* $\mathcal{L}G$ *can be lifted uniquely to a representation* α' *of* G *such that*

$$\boxed{\alpha = \mathcal{L}\alpha'.}$$

Proof ▶ If α is real then by Proposition 1 the Lie algebra homomorphism

$$\alpha : \mathcal{L}G \to \mathbf{gl}(n, \mathbb{R})$$

can be lifted (uniquely) to a group homomorphism

$$\alpha' : G \to \mathbf{GL}(n, \mathbb{R}).$$

On the other hand, suppose α is complex, ie a complex Lie algebra homomorphism

$$\alpha : \mathbb{C}\mathcal{L}G \to \mathbf{M}(n, \mathbb{C}).$$

Since $\mathcal{L}G < \mathbb{C}\mathcal{L}G$, this restricts to a real Lie algebra homomorphism

$$\alpha : \mathcal{L}G \to \mathbf{gl}(n, \mathbb{C});$$

and the result again follows by Proposition 1. ◀

Corollary 8.2 1. *Suppose $\mathcal{L}$ is a complex Lie algebra. If there exists a simply connected compact Lie group G such that*

$$\mathcal{L} = \mathbb{C}\mathcal{L}G$$

then every representation of $\mathcal{L}$ is semisimple.

2. *Suppose G is a linear group. If there exists a simply connected compact linear group H with the same complexification,*

$$\mathbb{C}\mathcal{L}G = \mathbb{C}\mathcal{L}H,$$

then every representation of G is semisimple.

Proof ►

1. Every representation of $\mathcal{L}$ arises from a representation of G, by Proposition 2, which we know from Part B is semisimple.

2. Every representation of G arises from a representation of $\mathcal{L} = \mathbb{C}\mathcal{L}G$, which by (1) is semisimple.

◄

Example: Every representation of the group $\mathbf{SL}(2,\mathbb{R})$ *is semisimple*, since its Lie algebra $\mathbf{sl}(2,\mathbb{R})$ has the same complexification as the Lie algebra $\mathbf{su}(2)$ of the simply connected compact linear group $\mathbf{SU}(2)$.

Chapter 9

The Representations of $\mathbf{sl}(2,\mathbb{R})$

As we know, $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$ share the same Lie algebra; and this algebra has the same complexification, and so the same representation theory, as that of $\mathbf{SL}(2,\mathbb{R})$. So in studying the Lie theory of any one of these groups we are in effect studying all three; and we can choose whichever is most convenient for our purpose. This turns out to be the algebra $\mathbf{sl}(2,\mathbb{R})$.

Proposition 9.1 *The Lie algebra*

$$\mathbf{sl}(2,\mathbb{R}) = \langle H, E, F : [H,E] = 2E, [H,F] = -2F, [E,F] = H\rangle$$

has just 1 simple representation (over $\mathbb{C}$) of each dimension $1, 2, 3, \ldots$. If we denote the representation of dimension $2j+1$ by D_j (for $j = 0, 1/2, 1, 3/2, \ldots$) then

$$D_j = \langle e_j, e_{j-1}, \ldots, e_{-j}\rangle,$$

with

$$\boxed{\begin{array}{rcl} He_k &=& 2ke_k, \\ Ee_k &=& (j-k)e_{k+1}, \\ Fe_k &=& (j+k)e_{k-1}. \end{array}}$$

for $k = j, j-1, \ldots, -j$, setting $e_{j+1} = e_{-j-1} = 0$.

Remark: Note that we have already proved the existence and uniqueness of the D_j in Part II. For

1. The representations of $\mathbf{sl}(2,\mathbb{R})$ and $\mathbf{su}(2)$ are in 1–1 correspondence, since these algebras have the same complexification;

2. The representations of $\mathbf{su}(2)$ and $\mathbf{SU}(2)$ are in 1–1 correspondence, since the group $\mathbf{SU}(2)$ is simply-connected;

3. We already saw in Part II that $\mathbf{SU}(2)$ possessed a unique representation D_j of dimension $2j+1$ for each half integer $j = 0, 1/2, 1, \ldots$.

However, we shall re-establish this result by purely algebraic means.

Proof ▶ Suppose α is a simple representation of $\mathbf{sl}(2, \mathbb{R})$ in the (finite-dimensional) complex vector space V.

To fit in with our subsequent nomenclature, we shall term the eigen-values of H (or rather, of αH) the *weights* of the representation α; and we shall call the corresponding eigen-vectors *weight-vectors.* We denote the *weight-space* formed by the weight-vectors of weight ω by

$$W(\omega) = \{v \in V : Hv = \omega v\}.$$

Suppose ω is a weight of α; say

$$He = \omega e.$$

(Note that α has at least 1 weight, since a linear transformation over $\mathbb{C}$ always possesses at least 1 eigen-vector.) Since

$$[H, E] = 2E,$$

we have

$$\begin{aligned} & (HE - EH)e = 2Ee \\ \Longrightarrow\ & HEe - \omega Ee = 2Ee \\ \Longrightarrow\ & H(Ee) = (\omega + 2)Ee. \end{aligned}$$

Thus either $Ee = 0$; or else Ee is also a weight-vector, but of weight $\omega + 2$. In any case,

$$\boxed{e \in W(\omega) \Longrightarrow Ee \in W(\omega + 2).}$$

Similarly,

$$\begin{aligned} (HF - FH)e &= [H, F]e = -2Fe \\ \Longrightarrow H(Fe) &= (\omega - 2)Fe. \end{aligned}$$

Thus either $Fe = 0$; or else Fe is also a weight-vector, but of weight $\omega - 2$. In any case,

$$\boxed{e \in W(\omega) \Longrightarrow Fe \in W(\omega - 2).}$$

This is sometimes expressed by saying that E is a "raising operator", which raises the weight by 2; while F is a "lowering operator", which lowers the weight by 2.

A finite-dimensional representation can only possess a finite number of weights, since weight-vectors corresponding to distinct weights are necessarily linearly independent. Let μ be the maximal weight, and let e be a corresponding weight-vector. Then

$$Ee = 0,$$

since $\mu + 2$ is not a weight.

Repeated action on e by F will take us 2 rungs at a time down the "weight-ladder", giving us weight vectors

$$e, Fe, F^2e, \ldots$$

of weights

$$\mu, \mu - 2, \mu - 4, \ldots$$

until finally the vectors must vanish (since there are only a finite number of weights); say

$$F^{2j}e \neq 0, F^{2j+1}e = 0,$$

for some half-integer j. Set

$$v_\mu = e, v_{\mu-2} = Fe, \ldots, v_{\mu-4j} = F^{2j}e;$$

and let

$$v_\omega = 0 \text{ if } \omega > \mu \text{ or } \omega < \mu - 4j.$$

Then

$$\begin{aligned} Hv_\omega &= \omega v_\omega, \\ Fv_\omega &= v_{\omega-2} \end{aligned}$$

for $\omega = \mu, \mu - 2, \ldots$

Since F takes us 2 rungs down the weight-ladder, while E takes us 2 rungs up, EF and FE both leave us at the same level:

$$v \in W(\omega) \Longrightarrow EFv, FEv \in W(\omega).$$

We shall show that these 2 new weight-vectors are in fact the same, up to scalar multiples, as the one we started from, ie

$$EFv_\omega = a(\omega)v_\omega,\ FEv_\omega = b(\omega)v_\omega.$$

Notice that each of these results implies the other, since

$$\begin{aligned} (EF - FE)v_\omega &= [E, F]v_\omega \\ &= Hv_\omega \\ &= \omega v_\omega. \end{aligned}$$

So eg the second result follows from the first, with

$$b(\omega) = a(\omega) - \omega.$$

At the top of the ladder,

$$FEv_\mu = 0.$$

Thus the result holds there, with

$$b(\mu) = 0, \text{ ie } a(\mu) = \mu.$$

On the other hand, at the bottom of the ladder,

$$EFv_{\mu-4j} = 0.$$

So the result holds there too, with

$$a(\mu - 4j) = 0.$$

We establish the general result by induction, working down the ladder. Suppose it proved for

$$v_\mu, v_{\mu-2}, \dots, v_{\omega+2}.$$

Then

$$\begin{aligned} FEv_\omega &= FEFv_{\omega+2} \\ &= Fa(\omega+2)v_{\omega+2} \\ &= a(\omega+2)v_\omega \end{aligned}$$

Thus the result also holds for v_ω, with

$$\begin{aligned} b(\omega) &= a(\omega+2), \\ \text{ie } a(\omega) &= a(\omega+2) + \omega. \end{aligned}$$

This establishes the result, and also gives $a(\omega)$ by recursion:

$$\begin{aligned} a(\mu) &= \mu, \\ a(\mu-2) &= a(\mu) + \mu - 2 \\ a(\mu-4) &= a(\mu-2) + \mu - 4, \\ &\dots \\ a(\omega) &= a(\omega+2) + \omega. \end{aligned}$$

Hence

$$\begin{aligned} a(\omega) &= m + (m-2) + \dots + \omega \\ &= (m - \omega + 2)(m + \omega)/4, \end{aligned}$$

while

$$\begin{aligned} b(\omega) &= a(\omega) - \omega \\ &= (m-\omega)(m+\omega+2)/4. \end{aligned}$$

At the bottom of the ladder,

$$a(\mu - 4j) = 0 \text{ ie } (2j+1)(\mu - 2j) = 0.$$

Hence

$$\mu = 2j.$$

Thus the weights run from $+2j$ down to $-2j$; and

$$a(\omega) = (j - \omega/2 + 1)(j + \omega/2), \quad b(\omega) = (j - \omega/2)(j + \omega/2 + 1).$$

In particular,

$$\begin{aligned} Ev_\omega &= EFe_{\omega+2} \\ &= a(\omega+2)e_{\omega+2} \\ &= (j-\omega/2)(j+\omega/2+1)e_{\omega+2}. \end{aligned}$$

The space

$$U = \langle v_{2j}, v_{2j-2}, \ldots, v_{-2j} \rangle$$

spanned by the weight-vectors is stable under H, E and F. and is therefore the whole of V, since the representation was supposed simple. On the other hand U is simple; for any subspace stable under $\mathbf{sl}(2,\mathbb{R})$ must contain a weight-vector. This must be a scalar multiple of one of the v_ω; and all the others can then be recovered by the action of E and F.

This establishes the result; it only remains to "prettify" the description of D_j, by

- Indexing the basis weight-vectors by $\kappa = \omega/2$ in place of ω;
- Renormalising these vectors (now christened e_κ) so that

  $$Fe_\kappa = (j+\kappa)e_{\kappa-1}.$$

  It then follows that

  $$\begin{aligned} (j+\kappa+1)Ee_\kappa &= EFe_{\kappa+1} \\ &= a(\kappa+1)e_{\kappa+1} \\ &= (j-\kappa)(j+\kappa+1)e_{\kappa+1} \\ \text{ie } Ee_\kappa &= (j-\kappa)e_{\kappa+1}, \end{aligned}$$

  as stated.

◀

Remark: Note in particular that all the weights of all the representations of $\mathbf{sl}(2,\mathbb{R})$ are *integral.* The Lie algebra

$$\mathbf{su}(2) = \langle H, U, V : [H,U] = V, [H,V] = -U, [U,V] = H\rangle$$

has just 1 simple representation (over $\mathbb{C}$) of each dimension $1, 2, 3, \ldots$ If we denote the representation of dimension $2j+1$ by D_j (for $j = 0, 1, 1, 3, \ldots$) then

$$D_j = \langle e_j, e_{j-1}, \ldots, e_{-j}\rangle,$$

with

$$\begin{aligned} He_\kappa &= 2\kappa e_\kappa, \\ Ue_\kappa &= (j-\kappa)e_{\kappa+1} + (j+\kappa)e_{\kappa-1}, \\ Ve_\kappa &= i(j-\kappa)e_{\kappa+1} - i(j+\kappa)e_{\kappa-1}. \end{aligned}$$

Proposition 9.2 *Every representation of* $\mathbf{sl}(2,\mathbb{R})$ *is semisimple.*

Proof ▶ The representations of $\mathbf{sl}(2,\mathbb{R})$ are in 1–1 correspondence with the representations of $\mathbf{su}(2)$, since these 2 Lie algebras have the same complexification.

Moreover, the representations of $\mathbf{su}(2)$ are in 1-1 correspondence with the representations of the group $\mathbf{SU}(2)$, since $\mathbf{SU}(2)$ is simply-connected.

But the representations of $\mathbf{SU}(2)$ are all semisimple, since this group is compact. So therefore are the representations of $\mathbf{su}(2)$, and hence of $\mathbf{sl}(2,\mathbb{R})$. ◀

Proposition 9.3 *For all half-integers* j, k,

$$D_j D_k = D_{j+k} + D_{j+k-1} + \ldots + D_{|j-k|}$$

Remark: We have already proved this result—or rather the corresponding result for $\mathbf{SU}(2)$—using character theory. But it is instructive to give an algebraic proof.

Proof ▶ We know that $D_j D_k$ is semisimple, and so a sum of D_j's.

Lemma 9.1 *Suppose* α, α' *are representations of* $\mathbf{sl}(2,\mathbb{R})$ *in* V, V'*; and suppose* e, e' *are weight-vectors in* V, V' *of weights* ω, ω' *respectively. Then the tensor product* $e \otimes e'$ *is a weight-vector of* $\alpha\alpha'$*, with weight* $\omega + \omega'$.

Proof of Lemma ▷ By the definition of the action of a Lie algebra on a tensor product,

$$\begin{aligned} H(e \otimes e') &= (He) \otimes e' + e \otimes (He') \\ &= (\omega + \omega')(e \otimes e') \end{aligned}$$

◁

Corollary 9.1 *If Ω is the set of weights of the representation α, and Ω' that of α', then the set of weights of the product representation $\alpha\alpha'$ is the sum-set*

$$\Omega + \Omega' = \{\omega + \omega' : \omega \in \Omega, \omega' \in \Omega'\}.$$

Remark: When we speak of "sets" of weights, it is understood that each weight appears with a certain multiplicity.
Proof of Proposition 3 (resumed). Let us denote the weight-ladder joining $+2j$ to $-2j$ by

$$L(j) = \{2j, 2j-2, \ldots, -2j\}.$$

By the Corollary above, the weights of $D_j D_k$ form the sum-set

$$L(j) + L(k).$$

We must express this set as a union of ladders. (Nb sum-sets must not be confused with unions.)

A given set of weights (with multiplicities) can be expressed in at most 1 way as a union of ladders, as may be seen by successively removing ladders of maximal length.

To express the sum-set above as such a union, note that

$$L(k) = \{2k, -2k\} \cap L(k-1).$$

We may assume that $k \leq j$. Then

$$L(j) + \{2k, -2k\} = L(j+k) \cap L(j-k).$$

Hence

$$\begin{aligned} L(j) + L(k) &= L(j) + (\{2k, -2k\} \cap L(k-1)) \\ &= (L(j) + 2k, -2k) \cap (L(j) + L(k-1)) \\ &= L(j+k) \cap L(j-k) \cap (L(j) + L(k-1)). \end{aligned}$$

It follows by induction on k that

$$L(j) + L(k) = L(j+k) \cap L(j+k-1) \cap \ldots \cap L(j-k).$$

But this is the weight-set of

$$D_{j+k} + D_{j+k-1} + \ldots + D_{j-k}.$$

The result follows. ◀

Chapter 10

The Representations of $\mathbf{su}(3)$

The representation theory of $\mathbf{sl}(2,\mathbb{R})$ (and $\mathbf{su}(2)$) outlined above provides both a model, and a starting-point, for the representation theory of the much larger class of groups considered below. (We cannot at the moment define this class precisely—but it includes the whole of the "classical repertoire" catalogued in Chapter 1.)

As an illustration of the techniques involved, we take a quick look at the representations of $\mathbf{sl}(3,\mathbb{R})$. (This informal account will be properly "proofed" later.) Recall that since $\mathbf{su}(3)$ has the same complexification as $\mathbf{sl}(3,\mathbb{R})$, we are at the same time studying the representations of this algebra.

The only innovation in passing from $\mathbf{sl}(2,\mathbb{R})$ to $\mathbf{sl}(3,\mathbb{R})$ is that we must now consider weights with respect, not just to a single element H, but to a whole commuting family.

In general, suppose

$$H = \{H_1, H_2, \ldots, H_r\}$$

is a family of commuting operators:

$$[H_i, H_j] = H_i H_j - H_j H_i = 0 \quad \forall i, j.$$

Then we say that a vector e is a *weight-vector* (always with respect to the given family H) if

$$H_i e = \omega_i e \text{ for } i = 1, \ldots, r.$$

The *weight* of e is the r-tuple

$$\omega = (\omega_1, \ldots, \omega_r)$$

We denote the space formed by the weight-vectors of weight ω (together with the vector 0) by

$$\boxed{W(\omega_1, \ldots, \omega_r) = \{v \in V : H_i v = \omega_i v \text{ for } i = 1, \ldots, r\}.}$$

Note that a commuting family always possesses at least 1 weight-vector. For certainly H_1 possesses an eigen-vector, with eigen-value ω_1, say. But then the eigen-space

$$U = \{v \in V : H_1 v = \omega_1 v\}$$

is non-trivial; and it is moreover stable under $H_2, \ldots, H_r$, since

$$\begin{aligned} v \in U &\implies H_1 H_i v = H_i H_1 v = \omega_1 H_i v \\ &\implies H_i v \in U. \end{aligned}$$

We may therefore suppose, by induction on r, that $H_2, \ldots, H_r$, acting on U, possess a common eigen-vector; and this will then be a common eigen-vector of $H_1, \ldots, H_r$, ie a weight-vector of the family H.

Example: Let us take an informal look at the representation theory of $\mathbf{sl}(3, \mathbb{R})$. Although this algebra is 8-dimensional, it is convenient to work with the following 9-member spanning set rather than a basis, so that we can preserve the symmetry between the 3 matrix coordinates:

$$H = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad F = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Notice that

$$H + J + K = 0;$$

this being the only linear relation between the 9 elements.

The Lie products are readily computed:

$$\begin{array}{lllllllll}
[H,A] & = & 2A & [H,B] & = & -B & [H,C] & = & -C \\
[J,A] & = & -A & [J,B] & = & 2B & [J,C] & = & -C \\
[K,A] & = & -A & [K,B] & = & -B & [K,C] & = & 2C \\
\\
[H,D] & = & -2D & [H,E] & = & E & [H,F] & = & F \\
[J,D] & = & D & [J,E] & = & -2E & [J,F] & = & F \\
[K,D] & = & D & [K,E] & = & E & [K,F] & = & -2F \\
\\
[A,D] & = & H & [B,D] & = & 0 & [C,D] & = & 0 \\
[A,E] & = & 0 & [B,E] & = & J & [C,E] & = & 0 \\
[A,F] & = & 0 & [B,F] & = & 0 & [C,F] & = & K \\
\\
[B,C] & = & D & [C,A] & = & E & [A,B] & = & F \\
[E,F] & = & -A & [F,D] & = & -B & [G,E] & = & -C.
\end{array}$$

The 3 elements (H, J, K) form a commuting family (since they are diagonal). *All weights and weight-vectors will be understood to refer to this family.*

Notice that if (x, y, z) is such a weight then

$$H + J + K = 0 \Longrightarrow x + y + z = 0.$$

Thus the weights all lie in the plane section

$$\{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$$

of 3-dimensional space.

We shall make considerable use of the natural isomorphism between the following 3 sub-algebras and $\mathbf{sl}(2, \mathbb{R})$:

$$L_1 = \langle H, A, D\rangle, \quad L_2 = \langle J, B, E\rangle, \quad L_3 = \langle K, C, F\rangle \;\longleftrightarrow\; \mathbf{sl}(2, \mathbb{R}).$$

In particular, suppose (x, y, z) is a weight of (H, J, K). Then x is a weight of $L_1 = \langle H, A, D\rangle$. But we saw above that all the weights of $\mathbf{sl}(2, \mathbb{R})$ are integral. It follows that x is an integer. And so, similarly, are y and z. Thus all weights (x, y, z) are *integer triples*.

Now suppose α is a simple representation of $\mathbf{sl}(3, \mathbb{R})$ in V; and suppose e is a weight-vector of weight (x, y, z). Then

$$(HA - AH)e = [H, A]e = 2Ae$$

ie

$$H(Ae) = (x + 2)Ae.$$

Similarly

$$\begin{aligned} J(Ae) &= (y-1)Ae, \\ K(Ae) &= (z-1)Ae. \end{aligned}$$

Thus

$$\begin{aligned} e \in W(x,y,z) &\implies Ae \in W(x+2, y-1, z-1), \\ e \in W(x,y,z) &\implies Be \in W(x-1, y+2, z-1) \\ &\implies Ce \in W(x-1, y-1, z+2) \\ &\implies De \in W(x-2, y+1, z+1) \\ &\implies Ee \in W(x+1, y-2, z+1) \\ &\implies Fe \in W(x+1, y+1, z-2). \end{aligned}$$

Our argument is merely a more complicated version of that for $\mathbf{sl}(2,\mathbb{R})$. Let us define a *maximal* weight to be one maximising x. Then if e is a corresponding weight-vector we must have

$$Ae = 0, Ee = 0, Fe = 0.$$

Our aim is to construct a stable subspace by acting on e with the operators A, B, C, D, E, F.

In fact the subspace spanned by the weight-vectors

$$W(j,k) = B^j C^k e$$

is stable under $\mathbf{sl}(3,\mathbb{R})$, ie for each operator $X \in A, B, C, D, E, F$

$$XW(j,k) = xW(j', k')$$

for some scalar x, and appropriate j', k'.

This may readily be shown by induction on $j+k$. As an illustration, take $X = A$, and suppose $j > 0$. Then

$$\begin{aligned} AW(j,k) &= ABW(j-1,k) \\ &= BAW(j-1,k) + [A,B]W(j-1,k), \end{aligned}$$

and the inductive hypothesis may be applied to each term.

We conclude that there is at most 1 simple representation with a given maximal weight.

A rather different point of view throws an interesting light on the *weight-diagram*, ie the set of weights, of a representation. Consider the restriction of the representation to L_1. From the representation theory of $\mathbf{sl}(2, \mathbb{R})$, the weights divide into L_1-ladders

$$(x, y, z), (x-2, y+1, z+1), \ldots, (-x, y+x, z+x).$$

Since $x + y + z = 0$, this ladder in fact joins

$$(x, y, z) \text{ to } (-x, -z, -y),$$

and is sent into itself by the reflection

$$R : (x, y, z) \mapsto (-x, -z, -y).$$

in the line

$$x = 0, y + z = 0.$$

It follows that the whole weight-diagram is sent into itself by R. Similarly (taking L_2 and L_3 in place of L_1), the diagram is symmetric under the reflections

$$S : (x, y, z) \mapsto (-z, -y, -x), \quad T : (x, y, z) \mapsto (-y, -x, -z);$$

and so under the group W formed by the identity 1, R, S, T and the compositions

$$ST : (x, y, z) \mapsto (z, x, y) \text{ and } TS : (x, y, z) \mapsto (y, z, x).$$

We note that, starting from any non-zero weight ω, just 1 of the 6 transforms $g\omega$ (with $g \in W$) of a given non-zero weight ω lies in the "chamber"

$$\{(x, y, z) : x + y + z = 0, x > 0, y < 0, z < 0\}.$$

Our argument shows that, starting from any integer triple ω in this chamber, we can construct a weight-diagram having ω as maximal weight, by taking the 6 transforms of ω, and filling in all the ladders that arise.

In this way it may be seen that there is indeed a simple representation of $\mathbf{sl}(3, \mathbb{R})$ having a given integer triple (x, y, z), with $x > 0, y < 0, z < 0$, as maximal weight.

In conclusion, we note that *every representation of* $\mathbf{sl}(3, \mathbb{R})$ *is semisimple*. For the restrictions to L_1, L_2 and L_3 are semisimple, from the representation theory of $\mathbf{sl}(2, \mathbb{R})$. Moreover, from that theory we see that H, J and K are each diagonalisable.

But a commuting family of matrices, each of which is diagonalisable, are simultaneously diagonalisable. (That follows by much the same argument—restricting to the eigenspaces of one of the operators—used earlier to show that such a family possesses at least 1 weight-vector.) Thus every representation of $\mathbf{sl}(3, \mathbb{R})$ is spanned by its weight-vectors. The semisimplicity of the representation follows easily from this; for we can successively add simple parts, choosing at each stage the simple representation corresponding to a maximal remaining weight—until finally the sum must embrace the whole representation.

Chapter 11

The Adjoint Representation

Definition 11.1 *The adjoint representation* ad *of the Lie algebra* $\mathcal{L}$ *is the representation in* $\mathcal{L}$ *itself defined by*

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

Remark: We should verify that this does indeed define a representation of $\mathcal{L}$. It is clearly bilinear; so it reduces to verifying that

$$\mathrm{ad}[X, Y] = \mathrm{ad}\, X\, \mathrm{ad}\, Y - \mathrm{ad}\, Y\, \mathrm{ad}\, X,$$

ie

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

for all $Z \in \mathcal{L}$. But this is just a re-arrangement of Jacobi's identity.

Chapter 12

Compactness and the Killing Form

As we know, every representation of a compact group carries an invariant positive-definite quadratic form. When we find that the adjoint representation of a Lie algebra also carries an invariant form, it is natural to ask—at least in the case of a compact linear group—whether these are in fact the same. If that is so, then we should be able to determine the compactness of a linear group from its Lie algebra.

Definition 12.1 *Suppose α is a representation of the Lie algebra $\mathcal{L}$ in the vector space V. Then the trace form of α is the quadratic form on $\mathcal{L}$ defined by*

$$T(X) = \operatorname{tr}\left((\alpha X)^2\right).$$

In particular, the Killing form of $\mathcal{L}$ is the trace form of the adjoint representation, ie the quadratic form on $\mathcal{L}$ defined by

$$K(X) = \operatorname{tr}\left((\operatorname{ad} X)^2\right).$$

Theorem 12.1 *Suppose G is a connected linear group.*

1. *If G is compact, then the trace form of every representation of $\mathcal{L}G$ is negative (ie negative-definite or negative-indefinite). In particular the Killing form of $\mathcal{L}G$ is negative:*

$$G \text{ compact } \Longrightarrow K \leq 0.$$

2. *If the Killing form on $\mathcal{L}G$ is negative-definite, then G is compact:*

$$K < 0 \Longrightarrow G \text{ compact}.$$

Proof ► 1. Suppose G is compact; and suppose α is a representation of G in V. Then we can find a positive-definite form $P(v)$ on V invariant under G—and therefore also under $\mathcal{L}G$. By change of coordinates we may suppose that

$$P(x_1, x_2, \dots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2.$$

In other words, $\alpha(G) \subset O(n)$. But then

$$\alpha'(\mathcal{L}G) \subset o(n) = \{X \in \mathbf{M}(n, \mathbb{R}) : X' = -X\},$$

where $\alpha' = \mathcal{L}\alpha$. Thus

$$X^2 = -X'X,$$

where "by abuse of notation" we write X for $\alpha' X$; and so

$$K(X) = \operatorname{tr}(X^2) = -\operatorname{tr}(X'X) \leq 0.$$

2. It is much more difficult to establish the converse. Suppose $K < 0$.

Lemma 12.1 *If the Killing form K of the Lie algebra $\mathcal{L}$ is non-singular, then the Lie algebra homomorphism*

$$\operatorname{ad} : \mathcal{L} \to \operatorname{der}(\mathcal{L})$$

is in fact an isomorphism.

Proof ► Firstly, ad is *injective*. For

$$\begin{aligned} X \in \ker(\operatorname{ad}) &\implies \operatorname{ad}(X) = 0 \\ &\implies K(X) = 0; \end{aligned}$$

and by hypothesis $K(X) = 0$ only if $X = 0$.

Secondly, ad is *surjective*, ie every derivation of $\mathcal{L}$ is of the form $\operatorname{ad}(D)$, for some $D \in \mathcal{L}$.

For suppose $d \in \operatorname{der}(\mathcal{L})$, ie $d : \mathcal{L} \to \mathcal{L}$ is a linear map satisfying

$$d([X, Y]) = [d(X), Y] + [X, d(Y)].$$

Consider the map

$$X \to \operatorname{tr}\left((\operatorname{ad} X)d\right).$$

Since K is non-singular, we can find $D \in \mathcal{L}$ such that

$$\operatorname{tr}\left((\operatorname{ad} X)d\right) = K(X, D) = \operatorname{tr}\left((\operatorname{ad} X)(\operatorname{ad} D)\right).$$

Setting $d' = d - \mathrm{ad}(D)$,

$$\mathrm{tr}\left((\mathrm{ad}\,X)d'\right) = 0$$

for all $X \in \mathcal{L}$.

Now

$$\mathrm{ad}(dY) = d(\mathrm{ad}\,Y) - (\mathrm{ad}\,Y)d = [d, \mathrm{ad}\,Y]$$

for any derivation d of $\mathcal{L}$, since

$$\begin{aligned}(\mathrm{ad}(dY))\,X &= [dY, X] \\ &= d\left([Y, X]\right) - [Y, dX] \\ &= d\left((\mathrm{ad}\,Y)X\right) - (\mathrm{ad}\,Y)(dX) \\ &= [d, \mathrm{ad}\,Y](X).\end{aligned}$$

Hence, substituting d' for d,

$$\begin{aligned}K(X, d'Y) &= \mathrm{tr}\left((\mathrm{ad}\,X)\,\mathrm{ad}(d'Y)\right) \\ &= \mathrm{tr}\left(\mathrm{ad}\,X)[d', \mathrm{ad}\,Y]\right) \\ &= \mathrm{tr}\left((\mathrm{ad}\,X)d'(\mathrm{ad}\,Y)\right) - \mathrm{tr}\left((\mathrm{ad}\,X)(\mathrm{ad}\,Y)d'\right) \\ &= \mathrm{tr}\left((\mathrm{ad}\,Y\,\mathrm{ad}\,X - \mathrm{ad}\,X\,\mathrm{ad}\,Y)d'\right) \\ &= -\,\mathrm{tr}\left(\mathrm{ad}[X, Y]d'\right) \\ &= 0.\end{aligned}$$

We conclude that $d'Y = 0$ for all Y; and so $d' = 0$, ie $d = \mathrm{ad}\,D$. ◄

Corollary 12.1 *If the linear group G is connected, and its Killing form $K < 0$, then the homomorphism*

$$\mathrm{Ad} : G \to \mathrm{Aut}(\mathcal{L}G)_0$$

is a covering.

Proof ► The Lie algebra homomorphism associated to this group homomorphism is just

$$\mathrm{ad} : \mathcal{L}G \to \mathcal{L}\left(\mathrm{Aut}(\mathcal{L}G)\right) = \mathrm{der}(\mathcal{L}G).$$

◄

We may assume, after suitable choice of basis, that

$$K(x_1, \ldots, x_n) = -x_1^2 - \cdots - x_n^2.$$

Thus

$$\mathrm{Aut}(\mathcal{L}G) \subset O(n)$$

is compact. The result will therefore follow if we can show that $Z = \ker(\mathrm{Ad})$ is finite.

First we establish that Z is finitely-generated.

Lemma 12.2 *Suppose*

$$\theta : G \to G_1$$

is a covering of a compact group G_1 by a connected linear group G. Then we can find a compact neighbourhood C of I in G such that

$$\theta(C) = G_1.$$

Proof ► Set

$$E = \{g = e^X : X \in \mathcal{L}G, \|X\| \le 1\}.$$

Since G is connected, it is generated by the exponentials e^X; thus

$$G = E \cup E^2 \cup E^3 \cup \dots$$

Hence

$$G_1 = \theta(E) \cup \theta(E^2) \cup \dots$$

Since G_1 is compact, $\theta(E^r) = G_1$ for some r. We can therefore set $C = E^r$. ◄

Corollary 12.2 *With the same assumptions,*

$$G = CZ,$$

where $Z = \ker\theta$.

Lemma 12.3 *Suppose*

$$\theta : G \to G_1$$

is a covering of a compact group G_1 by a connected linear group G. Then $\ker\theta$ is finitely-generated.

Proof ► From above, $G = CZ$, where C is compact and $Z = \ker\theta$. Then C^2 is compact, and so we can find a finite subset $\{z_1, \dots, z_r\} \subset Z$ such that

$$C^2 = Cz_1 \cup \dots \cup Cz_r.$$

It follows that

$$C^r \subset C\langle z_1, \dots, z_r\rangle$$

for all r. Since G is connected, and the union $\bigcup C^r$ is an open subgroup of G,

$$G = \bigcup C^r = C\langle z_1, \dots, z_r\rangle.$$

Consequently

$$Z = (C \cap Z)\langle z_1, \dots, z_r\rangle.$$

Since $C \cap Z$ is finite (being compact and discrete), we conclude that Z is finitely-generated. ◄

Lemma 12.4 *Suppose*

$$\theta : G \to G_1$$

is a covering of a compact group G_1 by a connected linear group G, with kernel Z. Then any homomorphism

$$f : Z \to \mathbb{R}$$

can be extended to a homomorphism

$$f : G \to \mathbb{R}.$$

Proof ► With the notation above, let $G = CZ$, with C compact. Let

$$u : G \to \mathbb{R}$$

be a continuous non-negative function with compact support containing C:

$$u(c) > 0 \quad \text{for all } c \in C.$$

If we define the function $v : G \to \mathbb{R}$ by

$$v(g) = \frac{u(g)}{\sum_{z \in Z} u(gz)}$$

then

$$\sum_{z \in Z} v(gz) = 1$$

for each $g \in G$.

Now set

$$t(g) = \sum_z v(gz) f(z).$$

Note that

$$\begin{aligned} t(gz) &= \sum_{z'} v(gzz') f(z') \\ &= \sum_{z'} v(gz') f(z'z^{-1}) \\ &= \sum_{z'} v(gz') f(z') - \sum_{z'} v(gz') f(z) \\ &= t(g) - f(z). \end{aligned}$$

Let us define the function $T : G \times G \to \mathbb{R}$ by

$$T(g, h) = t(gh) - t(g).$$

Then

$$\begin{aligned} T(gz,h) &= t(ghz)-t(gz) \\ &= t(gh)-f(z)-t(g)+f(z) \\ &= t(gh)-t(g) \\ &= T(g,h). \end{aligned}$$

Thus $T(g,h)$ depends only on $\theta(g)$ and h; so we have a function $S : G_1 \times G \to \mathbb{R}$ such that

$$T(g,h) = S(\theta g, h).$$

We can now define the sought-for function by integrating S over G_1:

$$F(g) = \int_{G_1} S(g_1, g) dg_1.$$

To verify that F has the required properties, we note in the first place that

$$\begin{aligned} T(g,z) &= t(gz)-t(g) \\ &= f(z). \end{aligned}$$

Thus

$$S(g_1, z) = f(z)$$

for all $g_1 \in G_1$, and so

$$F(z) = f(z).$$

Secondly,

$$\begin{aligned} T(g,hh') &= t(ghh')-t(g) \\ &= t(ghh')-t(gh)+t(gh)-t(g) \\ &= T(gh,h')+T(g,h). \end{aligned}$$

Hence

$$S(g_1, hh') = S\left(g_1\theta(h), h'\right) + S\left(g_1, h\right).$$

for all $g_1 \in G_1$, and so on integration

$$F(hh') = F(h) + F(h').$$

◄

We have almost reached the end of our marathon! We want to show that $Z = \ker(\mathrm{Ad})$ is finite. Suppose not. We know that Z is finitely-generated. Thus by the structure theory of finitely-generated abelian groups,

$$Z = \mathbb{Z}^r \times F,$$

where F is finite. So if Z is not finite, it has a factor $\mathbb{Z}$; and the natural injection $\mathbb{Z} \to \mathbb{R}$ extends to a homomorphism

$$f : Z \to \mathbb{R}.$$

By our last lemma, this in turn extends to a homomorphism

$$F : G \to \mathbb{R}.$$

Since G is connected, the image of this homomorphism is a connected subgroup of $\mathbb{R}$ containing $\mathbb{Z}$, which must be $\mathbb{R}$ itself, ie F is surjective.

The corresponding Lie algebra homomorphism

$$\mathcal{L}F : \mathcal{L}G \to \mathcal{L}R$$

is therefore also surjective; so its kernel is an $n-1$-dimensional ideal of $\mathcal{L}G$. We can use the non-singular Killing form to construct a complementary 1-dimensional ideal $\langle J \rangle$.

$$\mathcal{L}G = \ker(\mathcal{L}F) \bigoplus \langle J \rangle.$$

But if $X \in \ker(\mathcal{L}F)$,

$$[J, X] \in \ker(\mathcal{L}F) \cap \langle J \rangle = \{0\},$$

since both are ideals. On the other hand, $[J, J] = 0$; so

$$[J, X] = 0$$

for all $X \in \mathcal{L}G$; and so $\mathrm{ad}(J) = 0$, and in particular

$$K(J, X) = 0$$

for all X, contradicting the non-singularity of K. ◄

Remark: In view of the length and complexity of the proof above, a brief resumé may be in place.

- We start with the homomorhism

$$\mathrm{Ad} : G \to \mathrm{Aut}(\mathcal{L}G)_0.$$

- We want to show that this is a covering. In Lie algebra terms, we have to establish that the homomorphism

$$\mathrm{ad} : \mathcal{L}G \to \mathcal{L}(\mathrm{Aut}\, \mathcal{L}G) = \mathrm{der}(\mathcal{L}G)$$

  is an isomorphism. This is in fact true for any Lie algebra $\mathcal{L}$ with non-singular Killing form.

- Injectivity follows at once from the fact that

  $$\ker(\mathrm{ad}) = Z\mathcal{L},$$

  since the radical of the Killing form of $\mathcal{L}$ contains $Z\mathcal{L}$:

  $$X \in Z\mathcal{L} \implies K(X, Y) = 0 \text{ for all } Y.$$

- Surjectivity is more difficult. We have to show that every derivation of $\mathcal{L}$ is of the form $\mathrm{ad}(X)$ for some $X \in \mathcal{L}$.

  When $\mathcal{L} = \mathcal{L}G$, this is equivalent to showing that every automorphism of $\mathcal{L}G$ in the connected component of the identity is of the form $X \mapsto gXg^{-1}$ for some $g \in G$. This in turn implies that every automorphism of G in the connected component of the identity is inner.

  However, that is of little help in proving the result. Our proof was somewhat formal and unmotivated. The result is perhaps best understood in the context of the cohomology of Lie algebras and their modules (or representations), In this context, the derivations of $\mathcal{L}$ consititute the 1-cocycles of the $\mathcal{L}$-module $\mathcal{L}$, while the derivations of the form $\mathrm{ad}(X)$ form the 1-coboundaries:

  $$Z^1(\mathcal{L}) = \mathrm{der}(\mathcal{L}), \quad B^1(\mathcal{L}) = \mathrm{ad}(\mathcal{L}).$$

  Thus the result reflects the fact that $H^1(\mathcal{L}) = 0$ for a semisimple Lie algebra $\mathcal{L}$.

- Having established that

  $$Ad : G \to G_1 = \mathrm{Aut}(G)_0$$

  is a covering, it remains to be shown that—if $K < 0$—this covering is finite.

- The fact that $K < 0$ implies that G_1 is compact. That is not sufficient in itself—a compact group can have an infinite covering, as the covering $\mathbb{R} \to \mathbb{T}$ of the torus shows.

- Again, our proof that $\ker(Ad)$ is finite was somewhat formal and unmotivated. And again, the result is probably best understood in the context of cohomology—in this case the cohomology of groups.

  For G is a central extension of G_1; and such extensions correspond to the second cohomology group $H^2(G_1)$. Now if K is non-singular, $H^2(G_1, \mathbb{R}) = 0$; from which it follows that every essential extension of G_1 is finite.

Proposition 12.1 *Suppose G is a compact linear group. Then*

$$\mathcal{L}G = [\mathcal{L}G, \mathcal{L}G] \oplus Z\mathcal{L}G.$$

Moreover, the Killing form vanishes on $Z\mathcal{L}G$, and is positive-definite on $[\mathcal{L}G, \mathcal{L}G]$.