

# Lie groups and Lie algebras

Jonny Evans

May 4, 2019

# Lecture 1: Introduction

To the students, past, present and future, who have/are/will taken/taking/take this course and to those interested parties who just read the notes and gave me feedback: thank you for providing the background level of enthusiasm, pedantry and general confusion needed to force me to improve these notes. All comments and corrections are very welcome.

## 1.1 What is a representation?

**Definition 1.1.** A representation  $\rho$  of a group  $G$  on an  $n$ -dimensional vector space  $V$  over a field  $k$  is an assignment of a  $k$ -linear map  $\rho(g): V \rightarrow V$  to each group element  $g$  such that

- $\rho(gh) = \rho(g)\rho(h)$
- $\rho(1_G) = 1_V$ .

Here are some equivalent definitions. A representation is...

1. ...an assignment of an  $n$ -by- $n$  matrix  $\rho(g)$  to each group element  $g$  such that  $\rho(gh) = \rho(g)\rho(h)$  (i.e. the matrices multiply together like the group elements they represent). This is clearly the same as the previous definition once you have picked a basis of  $V$  to identify linear maps with matrices;
2. ...a homomorphism  $\rho: G \rightarrow GL(V)$ . The multiplicativity condition is now hiding in the definition of homomorphism, which requires  $\rho(gh) = \rho(g)\rho(h)$ ;
3. ...an action of  $G$  on  $V$  by linear maps. In other words an action  $\tilde{\rho}: G \times V \rightarrow V$  such that, for each  $g \in G$ , the map  $\rho(g): V \rightarrow V$  defined by  $\rho(g)(v) = \tilde{\rho}(g, v)$  is linear. The multiplicativity condition is now hiding in the definition of an action, which requires  $\tilde{\rho}(gh, v) = \tilde{\rho}(g, \tilde{\rho}(h, v))$ .

You should internalise these points of view as being all the same: I will switch between them. More confusingly still, I will sometimes say “let  $V$  be a representation of  $G$ ” or “consider the action  $\rho$  of  $G$  on  $V$ ”. This is the only warning you’ll get, so don’t be confused.

## 1.2 Why do we care?

Why do we study representations? How do representations arise in the real world (of mathematics)? Here is a motivating example.

**Example 1.2** (Invariant theory for binary quadratic forms). For simplicity, let’s work over  $\mathbb{C}$ . A *binary quadratic form* is an expression  $ax^2 + bxy + cy^2$  in two variables. One can add

and subtract quadratic forms and rescale them by a real number, so they form a vector space. Another way to write a quadratic form is:

$$ax^2 + bxy + cy^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =: \mathbf{x}^T M \mathbf{x}$$

so we can think of the vector space of binary quadratic forms as the (3-dimensional) vector space  $V$  of symmetric 2-by-2 matrices.

One very old question is: when can we make a linear change of coordinates to transform one conic into another? Of course, the answer depends on what kinds of coordinate changes we allow.

Let's suppose that we can change coordinates by any matrix in  $SL(2, \mathbb{C})$ , the group of invertible 2-by-2 matrices. If  $S \in SL(2, \mathbb{C})$  then, changing basis by  $S$  we get

$$M \mapsto M' = S^T M S$$

This gives a 3-dimensional representation of  $SL(2, \mathbb{C})$  on  $V$  (the vector space of symmetric 2-by-2 matrices). Note that  $\det(M') = \det(M)$  since  $\det(S^T M S) = \det(S)^2 \det(M)$  and, by assumption  $\det(S) = 1$ . We say that  $\Delta := \det(M) = ac - b^2/4 = -\frac{1}{4}(b^2 - 4ac)$  is an *invariant* of the quadratic form. It is often called the discriminant.

We can diagonalise  $M$ . Thus we see that any quadratic form is equivalent to one with  $M = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ . We can also swap the eigendirections using  $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$  which has the effect of swapping  $\lambda_1$  and  $\lambda_2$ .

We can then use the matrix  $S = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$  ( $\lambda \neq 0$ ) to get

$$M = \begin{pmatrix} \lambda^2 \lambda_1 & 0 \\ 0 & \lambda^{-2} \lambda_2 \end{pmatrix}.$$

Since  $\Delta = \lambda_1 \lambda_2$  is an invariant, this leaves us with the following possibilities:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \Delta & 0 \\ 0 & 1 \end{pmatrix}.$$

The important thing to take away from this example is the existence of an invariant  $\Delta = -\frac{1}{4}(b^2 - 4ac)$  and the fact that the invariant almost completely characterises quadratic forms up to equivalence. Observe that this invariant is itself a quadratic polynomial in the coefficients of  $M$ .

The relation to representation theory is the following: the quadratic polynomials in the entries of  $M$  themselves form a 6-dimensional<sup>1</sup> vector space  $Q$  whose general element is:

$$Aa^2 + Bb^2 + Cc^2 + Dab + Eac + Fbc$$

---

<sup>1</sup>6 coordinates  $A, B, C, D, E, F$ .

When you act using  $S$  on  $M$  you also act on the coefficients of such a polynomial, so we get a 6-dimensional representation  $R: SL(2, \mathbb{C}) \rightarrow GL(Q)$  of  $SL(2, \mathbb{C})$  on the vector space  $Q$ . The vector  $\Delta := -\frac{1}{4}(b^2 - ac)$  (i.e.  $A = C = D = F = 0$ ,  $B = -\frac{1}{4}$ ,  $E = \frac{1}{4}$ ) is *invariant* under this representation, that is it is fixed by  $R(g)$  for all  $g \in SL(2, \mathbb{C})$ .

This motivates the key question:

Given a representation, when can we find invariant vectors/subspaces?

We will be interested in invariant subspaces of any dimension, not only one-dimensional ones. Moreover, we will try to *decompose* an arbitrary representation into *irreducible* subrepresentations (which cannot be decomposed further).

Incredibly, this will turn out to have relevance for the structure of the hydrogen atom and the classification of hadrons (protons, neutrons, kaons, etc.)

### 1.3 Smoothness

We saw above that representations are a special case of homomorphisms (they are homomorphisms  $G \rightarrow GL(V)$ ). The kind of thing we would like to be able to do is to classify homomorphisms  $G \rightarrow GL(V)$ . Without further restriction, that's nigh-on impossible.

**Example 1.3.** Let  $\mathbf{R}$  be the group of real numbers under addition. What are the homomorphisms  $\mathbf{R} \rightarrow \mathbf{R}$ ? Certainly we have the linear maps  $F_\lambda: \mathbf{C} \rightarrow \mathbf{C}$ ,  $F_\lambda(x) = \lambda x$  for  $\lambda \neq 0$ . Suppose  $F: \mathbf{R} \rightarrow \mathbf{R}$  is an additive homomorphism and set  $\lambda = F(1)$ . Clearly  $F(n) = n\lambda$  and  $qF(p/q) = F(p) = p\lambda$  so  $F(q/p) = p\lambda/q$ . Therefore  $F$  is just a rescaling by  $\lambda$  on rational numbers  $\mathbf{Q}$ . However, if we consider  $\mathbf{R}$  as a vector space over  $\mathbf{Q}$  and pick a basis (invoking the Axiom of Choice) and rescale each basis element independently. In other words, let  $A$  be a basis for  $\mathbf{R}$  over  $\mathbf{Q}$  so each  $r \in \mathbf{R}$  can be written as  $\sum_{a \in A} c_a a$ ,  $c_a \in \mathbf{Q}$  (with only finitely many  $c_a$  nonvanishing) and let  $\lambda: A \rightarrow \mathbf{R}$  be an arbitrary function. Then  $\sum_{a \in A} c_a a \mapsto \sum_{a \in A} \lambda(a) c_a a$  defines a (pretty pathological) homomorphism  $\mathbf{R} \rightarrow \mathbf{R}$ .

But  $\mathbf{R}$  is more than just a group: it's a *Lie group*. This means it has a coordinate on it and one can ask for the homomorphism to be smooth with respect to this coordinate. Differentiating the homomorphism condition:

$$F(s+t) = F(s) + F(t)$$

with respect to  $t$  at  $t = 0$  gives:

$$\dot{F}(s) = 0 + \dot{F}(0)$$

so that  $\dot{F}$  is constant. Therefore smooth homomorphisms  $\mathbf{R} \rightarrow \mathbf{R}$  are just given by linear maps  $x \mapsto \lambda x$  for some  $\lambda$ .

In this course we are going to focus on Lie groups (precise definition to be given later) because then we have the whole gamut of tools from calculus at our disposal.

## Lecture 2: Examples and exponentials

### 2.1 The matrix exponential

The simplest Lie group is the group  $U(1)$  of unit complex numbers

$$U(1) = \{z \in \mathbf{C} : z\bar{z} = 1\}.$$

Since Euler, we have known how to parametrise the elements of this group:

$$z \in U(1) \Leftrightarrow z = e^{i\theta} = \cos(\theta) + i \sin(\theta).$$

In other words, every element of the group can be written as the **exponential** of a **purely imaginary** number. We would like to generalise this useful parametrisation to other groups of matrices.

**Definition 2.1.** The exponential of a matrix  $A$  is defined to be

$$\exp(A) = 1 + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k.$$

**Example 2.2.** Let  $A = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}$ . Then  $A^2 = -\theta^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ , so

$$\begin{aligned} \exp \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} &= \left(1 - \frac{\theta^2}{2} + \frac{\theta^4}{4!} - \cdots\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \\ &\quad + \left(\theta - \frac{1}{3!}\theta^3 + \frac{1}{5!}\theta^5 - \cdots\right) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \end{aligned}$$

### 2.2 Convergence

This power series definition makes sense in any ring provided you can make sense of the infinite sum. We can make sense of convergence for sequences of  $n$ -by- $n$  matrices using the *operator norm*

$$\|A\|^2 = \inf\{C \in \mathbf{R} : |Av| \leq C|v| \ \forall v \in \mathbf{R}^n\}$$

(so a sequence of matrices  $A_k$  converges to  $A$  if  $\|A_k - A\| \rightarrow 0$  as  $k \rightarrow \infty$ ).

**Lemma 2.3.** This power series converges absolutely, i.e. there exists  $K \in \mathbf{R}$  such that

$$\forall \epsilon > 0 \ \exists N : \left| \sum_{n=0}^N \frac{1}{n!} \|A^n\| - K \right| < \epsilon.$$

**Proof.** Note that  $\|AB\| \leq \|A\|\|B\|$  because  $|ABv| \leq \|A\|\|Bv\| \leq \|A\|\|B\|\|v\|$  for all  $v$ . Therefore

$$\sum_{n=0}^N \frac{1}{n!} \|A^n\| \leq \sum_{n=0}^N \frac{1}{n!} \|A\|^n \leq \exp \|A\|$$

so this sequence of partial sums is a monotonic sequence bounded above and hence converges (to some number, say  $K$ ).  $\square$

**Corollary 2.4.** \*This power series converges (indeed any absolutely convergent series converges).

**Proof.** We need to show that  $\sum_{n=0}^N \frac{1}{n!} A^n$  is a Cauchy sequence in the space of matrices with the operator norm, i.e.

$$\left\| \sum_{n=0}^N \frac{1}{n!} A^n - \sum_{n=0}^M \frac{1}{n!} A^n \right\| < \epsilon$$

for sufficiently large  $M, N$ . But we can rewrite this difference as

$$\left\| \sum_{n=M+1}^N \frac{1}{n!} A^n \right\|$$

which is bounded from above by

$$\sum_{n=M+1}^N \frac{1}{n!} \|A^n\|$$

(using Cauchy-Schwarz) and hence by  $\sum_{n=M+1}^N \frac{1}{n!} \|A\|^n$ . But  $\sum_{n=0}^N \frac{1}{n!} \|A\|^n$  is a Cauchy sequence (converging to  $\exp(\|A\|)$ ) and hence, for sufficiently large  $M, N$

$$\sum_{n=M+1}^N \frac{1}{n!} \|A\|^n < \epsilon.$$

Now it is a theorem from Analysis 4 that a Cauchy sequence in a finite-dimensional normed space converges (a finite-dimensional normed vector space is a Banach space).  $\square$

*Remark 2.5.* In fact, the same proof (along with the Weierstrass M-test) implies that, for any  $R$ , the sequence of functions  $F_N(A) = \sum_{n=0}^N \frac{1}{n!} A^n$  defined on  $\{A \in \mathfrak{gl}(n, \mathbf{R}) : \|A\| \leq R\}$  converges uniformly to a function  $\exp(A)$  and the same is true of the partial derivatives of  $F_N$  and hence  $\exp(A)$  is differentiable and can be differentiated term-by-term.

**Corollary 2.6.** The function  $t \mapsto \exp(tA)$  satisfies the differential equation

$$\frac{d}{dt} \exp(tA) = A \exp(tA).$$

**Proof.** Differentiating  $\exp(tA) = \sum_{n=0}^{\infty} \frac{1}{n!} t^n A^n$  term-by-term we get

$$\sum_{n=0}^{\infty} \frac{1}{n!} n t^{n-1} A^n = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} t^{n-1} A^n = A \sum_{m=0}^{\infty} \frac{1}{m!} t^m A^m.$$

$\square$

**Corollary 2.7** (Cauchy product formula). We have

$$\exp(A) \exp(B) = \sum_{k=0}^{\infty} \sum_{i+j=k} \frac{1}{i!j!} A^i B^j.$$

**Proof.** This always holds for the product of absolutely convergent series.  $\square$

**Corollary 2.8.** *We have*

(a)  $\exp(-A)\exp(A) = 1$  (so  $(\exp(A))^{-1} = \exp(-A)$ ).

(b) If  $AB = BA$  then  $\exp(A)\exp(B) = \exp(B)\exp(A)$ .

**Proof.** For (a), using the Cauchy product formula, we have

$$\begin{aligned}\exp(-A)\exp(A) &= \sum_{k=0}^{\infty} \sum_{n+m=k} \frac{1}{n!m!} A^n (-A)^m \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{n+m=k} \frac{k!}{n!m!} A^n (-A)^m \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} (A - A)^k \\ &= 1.\end{aligned}$$

For (b), using the Cauchy product formula, we have

$$\begin{aligned}\exp(A)\exp(B) &= \sum_{k=0}^{\infty} \sum_{i+j=k} \frac{1}{i!j!} A^i B^j \\ &= \sum_{k=0}^{\infty} \sum_{i+j=k} \frac{1}{i!j!} B^j A^i \\ &= \exp(B)\exp(A)\end{aligned}$$

where we use commutativity in the second line and to get to the third line we relabel  $i \leftrightarrow j$  and apply the Cauchy product formula again.  $\square$

## 2.3 $U(n)$

We will take a simple example: unitary matrices  $U(n)$ , the  $n$ -dimensional generalisation of unit complex numbers. Let  $A^\dagger$  denote the Hermitian transpose of  $A$ . We say that  $A \in U(n)$  if

$$AA^\dagger = 1.$$

**Lemma 2.9.** *A matrix  $B$  is skew-Hermitian if and only if  $\exp(tB) \in U(n)$  for all  $t \in \mathbf{R}$ .*

**Proof.** Since  $(\exp(tB))^\dagger = \exp(tB^\dagger)$  (taking  $\dagger$  term by term in the power series), if  $B^\dagger = -B$  then  $(\exp(tB))^\dagger = \exp(-tB) = (\exp(tB))^{-1}$  so  $\exp(tB)$  is unitary (for any  $t \in \mathbf{R}$ ).

Conversely, if  $\exp(tB^\dagger) = \exp(-tB)$  for all  $t$  we can differentiate this relation with respect to  $t$  (at  $t = 0$ ) and we get  $B^\dagger = -B$ .  $\square$

*Remark 2.10.* We used differentiation in the proof. Note that we can't just take logarithms: just like for complex numbers there's no globally-defined single-valued logarithm function.

**Definition 2.11.** If  $G \subset GL(n, \mathbf{R})$  is a subgroup, we define its Lie algebra  $\mathfrak{g}$  to be

$$\mathfrak{g} := \{B \in \mathfrak{gl}(n, \mathbf{R}) : \exp(tB) \in G \forall t \in \mathbf{R}\}.$$

So we have just seen that  $\mathfrak{u}(n)$ , the Lie algebra of  $U(n)$  is the space of skew-Hermitian matrices  $B^\dagger = -B$ .

## 2.4 $SU(2)$

My favourite example is the group of unitary 2-by-2 matrices with determinant 1:  $SU(2)$ . This has Lie algebra  $\mathfrak{su}(2)$ , the 3-dimensional space of skew-Hermitian matrices with trace zero, whose general element is:

$$M_v := \begin{pmatrix} ix & y + iz \\ -y + iz & -ix \end{pmatrix}, \quad v = (x, y, z) \in \mathbf{R}^3.$$

**Lemma 2.12.** *A general element of  $SU(2)$  has the form*

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

for  $a, b \in \mathbf{C}$  with  $|a|^2 + |b|^2 = 1$ .

**Proof.** If  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(2)$  then  $A^\dagger = A^{-1}$  implies

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

(since  $\det(A) = 1$ ). Thus  $d = \bar{a}$  and  $c = -\bar{b}$ . The  $\det(A) = 1$  condition now implies that  $1 = a\bar{a} - (-\bar{b}b) = |a|^2 + |b|^2$ .  $\square$

This means that  $SU(2)$  can be identified with the 3-dimensional sphere

$$S^3 = \{(x_1, x_2, x_3, x_4) \in \mathbf{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1\}$$

where  $a = x_1 + ix_2$  and  $b = x_3 + ix_4$ .

In Question Sheet 1 we will show that if  $u = (x, y, z) \in \mathbf{R}^3$  is a unit vector then

$$\exp(\theta M_u) = \cos \theta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin \theta M_u = \begin{pmatrix} \cos \theta + ix \sin \theta & y \sin \theta + iz \sin \theta \\ -y \sin \theta + iz \sin \theta & \cos \theta - ix \sin \theta \end{pmatrix}.$$

It is easy to see that this matrix is in  $SU(2)$ . In fact:

**Lemma 2.13.** *Any matrix in  $SU(2)$  can be written as  $\begin{pmatrix} \cos \theta + ix \sin \theta & y \sin \theta + iz \sin \theta \\ -y \sin \theta + iz \sin \theta & \cos \theta - ix \sin \theta \end{pmatrix}$  for some  $\theta, x, y, z$ .*

**Proof.** If  $a = x_1 + ix_2$  and  $b = x_3 + ix_4$  with  $|a|^2 + |b|^2 = 1$  then take  $\cos \theta$  to be  $x_1$  so that  $\sin^2 \theta$  equals  $x_2^2 + x_3^2 + x_4^2$ . Set  $\sin \theta$  to be one of the square roots  $\sqrt{x_2^2 + x_3^2 + x_4^2}$  and define  $(x, y, z) = (x_2/\sin \theta, x_3/\sin \theta, x_4/\sin \theta)$ . Now  $\theta$  is determined (modulo  $2\pi\mathbf{Z}$ ) by  $\cos \theta$  and  $\sin \theta$ .  $\square$

This tells us that the exponential map takes  $\mathfrak{su}(2)$  to  $SU(2)$  surjectively, but it is clearly not injective. For example,  $\exp(\pi M_u) = -1$  for all unit vectors  $u$ .



## Lecture 3: Local logarithm

The aim is to prove that in a neighbourhood of  $0 \in U' \subset \mathfrak{gl}(n, \mathbf{R})$  and  $1 \in V' \subset GL(n, \mathbf{R})$  the exponential map  $\exp: \mathfrak{gl}(n, \mathbf{R}) \rightarrow GL(n, \mathbf{R})$  admits an inverse  $\log: V' \rightarrow U'$ .

### 3.1 Calculus of several variables: review

**Definition 3.1.** Suppose that  $U \subset \mathbf{R}^m$  and  $V \subset \mathbf{R}^n$  are open sets, that  $(x_1, \dots, x_n)$  are coordinates on  $U$  and that  $F: U \rightarrow V$  is a map. We say that  $F$  is smooth if, for each component  $F_i$  of  $F$ , all possible partial derivatives  $\frac{\partial^k F}{\partial x_{i_1} \dots \partial x_{i_k}}$  exist and are continuous.

**Definition 3.2.** We write  $d_p F$  for the matrix of partial derivatives at  $p \in U$ :

$$d_p F = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(p) & \cdots & \frac{\partial F_n}{\partial x_1}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_1}{\partial x_m}(p) & \cdots & \frac{\partial F_n}{\partial x_m}(p) \end{pmatrix}.$$

We can think of this as a linear map  $\mathbf{R}^m \rightarrow \mathbf{R}^n$  which is the best linear approximation to  $F$  at  $p$  in the sense that

$$F(p + v) = F(p) + d_p F(v) + o(|v|).$$

Another way to think of  $d_p F(v)$  is as the directional derivative

$$d_p F(v) = \left. \frac{d}{dt} \right|_{t=0} F(p + tv)$$

of  $F$  at  $p$  in the  $v$ -direction.

**Example 3.3.** Consider the map  $F(x) = x^2$ . We have

$$F(x + t) = (x + t)^2 = x^2 + 2xt + \mathcal{O}(t^2),$$

so  $d_x F(t) = 2xt$  (which is linear in  $t$ ).

**Example 3.4.** Let  $H$  be the space of Hermitian matrices  $\{A \in \mathfrak{gl}(n, \mathbf{C}) : A^\dagger = A\}$  and consider the map  $F: \mathfrak{gl}(n, \mathbf{C}) \rightarrow H$  defined by  $F(A) = A^\dagger A$ . We have

$$F(A + B) = (A + B)^\dagger (A + B) = A^\dagger A + (B^\dagger A + A^\dagger B) + \mathcal{O}(|B|^2)$$

so  $d_A F(B) = B^\dagger A + A^\dagger B$ .

This is going to be the easiest way to compute derivatives in this course! If you think too much about it you will realise you're working with matrices whose entries are matrices (whose entries might well be matrices!) and your mind will blow.

The biggest advantage to thinking of derivatives as linear maps is the following beautiful form of the chain rule:

**Lemma 3.5.** If  $U_i \subset \mathbf{R}^{n_i}$ ,  $i = 1, 2, 3$ , are open sets and  $U_1 \xrightarrow{F_1} U_2 \xrightarrow{F_2} U_3$  is a sequence of maps with composite  $F_3 = F_2 \circ F_1$  then

$$d_x F_3(v) = d_{F(x)} F_2(d_x F_1(v)),$$

that is  $d_x F_3 = d_{F(x)} F_2 \circ d_x F_1$  where  $\circ$  denotes matrix product.

Calculus is all about approximating functions by linear maps and the local properties of a function by properties of its linearisation. One of the best (and most useful) examples of this is the inverse function theorem, which states that a smooth inverse function exists if the linearisation is an invertible matrix.

**Theorem 3.6.** Let  $U$  and  $V$  be open subsets of  $\mathbf{R}^n$  and  $F: U \rightarrow V$  be a smooth map. If  $d_p F$  is an invertible matrix then there are neighbourhoods  $p \in U' \subset U$  and  $f(p) \in V' \subset V$  such that  $F|_{U'}: U' \rightarrow V'$  is a bijection with smooth inverse.

“Bijection with smooth inverse” is a bit of a mouthful which will come up again and again, so we introduce a name for it.

**Definition 3.7 (Diffeomorphism).** Let  $U, V \subset \mathbf{R}^m$  be open sets. A smooth map  $F: U \rightarrow V$  is a diffeomorphism if it is bijective with smooth inverse.

“Diffeomorphism” is still a bit of a mouthful, but it’s the word we use.

## 3.2 A local logarithm

In the last lecture we defined the matrix exponential

$$\exp: \mathfrak{gl}(n, \mathbf{R}) \rightarrow GL(n, \mathbf{R})$$

by

$$\exp(A) = 1 + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k.$$

The first thing we will prove here is that, near  $0 \in \mathfrak{gl}(n, \mathbf{R})$ , the matrix exponential is a diffeomorphism: in other words there is a unique (smooth) logarithm function defined on a neighbourhood of the identity with the property that  $\log(\exp A) = A$ .

**Theorem 3.8.** There exists a neighbourhood  $U'$  of  $0 \in \mathfrak{gl}(n, \mathbf{R})$  and a neighbourhood  $V'$  of  $1 \in GL(n, \mathbf{R})$  such that  $\exp|_{U'}: U' \rightarrow V'$  is a diffeomorphism.

**Proof.** We have  $\exp(A) = 1 + A + \mathcal{O}(\|A\|^2)$  so  $d_0 \exp(A) = A$ . Therefore  $d_0 \exp = \text{Id}: \mathfrak{gl}(n, \mathbf{R}) \rightarrow \mathfrak{gl}(n, \mathbf{R})$ . The identity is certainly invertible, so the result follows from the inverse function theorem.  $\square$

*Remark 3.9.* More explicitly, the matrix entries  $A_{ij}$  are coordinates, and in terms of these we have

$$\exp(A) = \begin{pmatrix} 1 + A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & 1 + A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & 1 + A_{nn} \end{pmatrix} + \mathcal{O}(A^2)$$

so

$$\frac{\partial \exp(A)}{\partial A_{11}}(0) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

and more generally  $\frac{\partial \exp(A)}{\partial A_{ij}}(0) = E_{ij}$  (where  $E_{ij}$  is the matrix with a 1 in the  $i$ th row and  $j$ th column). We should think of  $\partial \exp(A)/\partial A_{ij}(0)$  as the directional derivative  $d_0 \exp(E_{ij})$  of  $\exp$  at  $0 \in \mathfrak{gl}(n, \mathbf{R})$  in the  $E_{ij}$  direction. By taking linear combinations  $v = \sum v_{ij} E_{ij}$  we get  $d_0 \exp(v) = v$  so  $d_0 \exp$  is the identity.

Of course we give the name  $\log$  to the local inverse to  $\exp$  - we say “local” because it is not globally defined (for example,  $\exp \begin{pmatrix} 0 & -2\pi \\ 2\pi & 0 \end{pmatrix} = \begin{pmatrix} \cos 2\pi & -\sin 2\pi \\ \sin 2\pi & \cos 2\pi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \exp \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ , so  $\exp$  is not globally a bijection).

### 3.3 The Baker-Campbell-Hausdorff formula

**Lemma 3.10.**  $\log$  has a power series expansion around  $1 \in GL(n, \mathbf{R})$ :

$$\log(1 + X) = X - \frac{1}{2}X^2 + \frac{1}{3}X^3 - \cdots$$

This power series has radius of convergence 1 (since  $\log(1 - 1)$  is ill-defined).

**Proof of Lemma.** We know that a local  $\log(1 + X)$  exists for small  $X$  (with respect to the operator norm) and that its first derivative is the identity so we can try to compute its power series. If

$$\log(1 + X) = X + b_2X^2 + b_3X^3 + \cdots$$

and  $\exp(X) = 1 + X + a_2X^2 + a_3X^3 + \cdots$  then

$$\begin{aligned} X &= \log(1 + X + a_2X^2 + \cdots) \\ &= X + a_2X^2 + a_3X^3 + \cdots \\ &\quad + b_2(X + a_2X^2 + a_3X^3 + \cdots)^2 + \cdots \\ &= X + (a_2 + b_2)X^2 + (a_3 + 2b_2a_2)X^3 + \cdots \end{aligned}$$

which gives a sequence of recurrence relations for  $b_i$  in terms of the known coefficients  $a_i$ . These relations are completely independent of whether  $X$  is a matrix or a number, hence the usual logarithm Taylor coefficients solve this recurrence relation.  $\square$

Since

$$\begin{aligned} \exp(A) \exp(B) &= \left(1 + A + \frac{1}{2}A^2 + \cdots\right) \left(1 + B + \frac{1}{2}B^2 + \cdots\right) \\ &= 1 + A + B + AB + \frac{1}{2}(A^2 + B^2) + \cdots \end{aligned}$$

we get

$$\begin{aligned}
\log(\exp(A)\exp(B)) &= A + B + AB + \frac{1}{2}(A^2 + B^2) + \dots \\
&\quad - \frac{1}{2} \left( A + B + AB + \frac{1}{2}(A^2 + B^2) + \dots \right)^2 + \dots \\
&= A + B + AB + \frac{1}{2}(A^2 + B^2) \\
&\quad - \frac{1}{2}(A^2 + AB + BA + B^2 + \dots) \\
&= A + B + \frac{1}{2}(AB - BA) + \dots
\end{aligned}$$

so:

- to first order in  $A$  and  $B$  we get the usual law of logarithms,
- to second order we get a correction term which involves the commutator  $[A, B] = AB - BA$ .

**Lemma 3.11.** *The next term is*

$$\frac{1}{12} ([A, [A, B]] - [B, [A, B]]) .$$

**Proof.** See Question Sheet 2. □

In general we have:

**Theorem 3.12** (Baker-Campbell-Hausdorff). *The higher order terms can also be expressed in terms of iterated commutators. Explicitly,  $\log(\exp(A)\exp(B))$  is given by:*

$$\sum_{n>0} \frac{(-1)^{n-1}}{n} \sum_{r_i+s_i>0, 1\leq i\leq n} \frac{(\sum_{i=1}^n (r_i + s_i))^{-1}}{r_1!s_1! \cdots r_n!s_n!} \text{ad}_A^{r_1} \text{ad}_B^{s_1} \cdots \text{ad}_A^{r_{n-1}} \text{ad}_B^{s_{n-1}} K_{r_n, s_n}$$

where  $\text{ad}_X Y = [X, Y] = XY - YX$  and

$$K_{r_n, s_n} = \begin{cases} \text{ad}_A^{r_n} B & \text{if } s_n = 1 \\ A & \text{if } r_n = 1, s_n = 0 \\ 0 & \text{otherwise} \end{cases} .$$

*Remark 3.13.* The explicit formula is not hugely important: the moral of the theorem is that the anticommutator bracket on  $\mathfrak{gl}(n, \mathbf{R})$  determines the group law on  $GL(n, \mathbf{R})$ . This bracket is called the *Lie bracket* and the pair  $(\mathfrak{gl}(n, \mathbf{R}), [\cdot, \cdot])$  is an example of a *Lie algebra*.

### 3.4 Lie algebras

We now abstract the important properties of the Lie bracket and define *Lie algebras* in general.

**Definition 3.14** (Lie algebra). Let  $V$  be a  $\mathbf{K}$ -vector space and let  $[\cdot, \cdot]: V \times V \rightarrow V$  be a bilinear bracket satisfying

$$[X, X] = 0, \quad [[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0.$$

The structure  $(V, [\cdot, \cdot])$  is called a *Lie algebra* over  $\mathbf{K}$ . The three-term relation here is called the *Jacobi identity*.

*Remark 3.15.* Note that if  $\text{char } \mathbf{K} \neq 2$  then  $[X, X] = 0$  is equivalent to antisymmetry:  $[X, Y] = -[Y, X]$ . If we define  $\text{ad}_X Y = [X, Y]$  then the Jacobi identity is equivalent to the more natural-looking:

$$\text{ad}_{[X, Y]} = \text{ad}_X \text{ad}_Y - \text{ad}_Y \text{ad}_X.$$

More on the Jacobi identity later.

# Lecture 4: Matrix groups

## 4.1 Definition

The operator norm on matrices turns  $GL(n, \mathbf{R})$  into a metric space and we can talk about convergent sequences of matrices.

**Definition 4.1.** A subset  $G \subset GL(n, \mathbf{R})$  is *closed* if it has the following property: if  $A_n \in G$  is a sequence and  $A_n$  converges to  $A \in GL(n, \mathbf{R})$  then  $A \in G$ .

**Definition 4.2.** A *matrix group*  $G$  is a subgroup  $G \subset GL(n, \mathbf{R})$  which is closed.

Don't get confused between "closed" in the group theory sense (i.e.  $g, h \in G$  implies  $gh \in G$ ) and "closed" in the topological sense above. The next lemma is (hopefully) the only place where both will appear.

**Lemma 4.3.** Suppose that  $G \subset GL(n, \mathbf{R})$  is a subgroup. Then the topological closure  $\overline{G}$  is also a subgroup. (Recall that  $\overline{G} = \{g \in GL(n, \mathbf{R}) : \exists g_i \in G \text{ with } g_i \rightarrow g\}$ ).

**Proof.** Suppose that  $g = \lim_{i \rightarrow \infty} g_i$  and  $h = \lim_{i \rightarrow \infty} h_i$  are two points in  $\overline{G}$ . Since multiplication depends continuously (in fact bilinearly) on matrix entries,  $g_i h_i \rightarrow gh$ . Hence  $gh \in \overline{G}$  (so the topological closure is closed as a group!). Also, the matrix entries of  $g_i^{-1}$  are rational functions of the matrix entries of  $g_i$ , in particular continuous functions. So  $x \mapsto x^{-1}$  is continuous, hence  $g_i^{-1}$  is a convergent sequence with limit  $c \in G$ . Since  $g_i^{-1} g_i = 1$  and multiplication is continuous, we have  $cg = 1$  and hence  $c = g^{-1}$ . Thus  $g \in \overline{G}$  implies  $g^{-1} \in \overline{G}$ .  $\square$

*Remark 4.4.* If  $G$  is abelian then so is  $\overline{G}$ , because  $g_i h_i \rightarrow gh$ ,  $h_i g_i \rightarrow hg$  and  $g_i h_i = h_i g_i$  for sequences  $g_i, h_i \in G$ .

## 4.2 More examples

### 4.2.1 Stabilisers of quadratic forms

Here is a recipe for constructing many examples of matrix groups. Let  $Q$  be an  $n$ -by- $n$  matrix and define

$$G = \{A \in GL(n, \mathbf{R}) : A^T Q A = Q.\}$$

**Lemma 4.5.** The group  $G$  is closed.

**Proof.** Let  $A_i$  be a sequence in  $G$  which converges to  $A \in GL(n, \mathbf{R})$ . The map  $A \mapsto A^T Q A$  is continuous with respect to the operator norm:

$$\begin{aligned} \|(A+B)^T Q (A+B) - A^T Q A\| &= \|B^T Q A + A^T Q B + B^T Q B\| \\ &\leq \|B\|(2\|A\| + \|B\|)\|Q\| \\ &< \epsilon \text{ for } \|B\| < \min(\|A\|, \epsilon/3\|Q\|\|A\|). \end{aligned}$$

Since  $A_i^T Q A_i = Q$ , continuity implies  $A^T Q A = Q$ , so  $A \in G$ .  $\square$

**Example 4.6.** •  $Q = 1$ : We get the orthogonal group  $O(n)$  ( $A^T A = 1$ ).

- $Q = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ : We get the *Lorentz group*  $O(1, n-1)$ . The case  $O(1, 3)$  is relevant for special relativity.

- If  $Q$  is the  $2n$ -by- $2n$  matrix  $J = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots \\ -1 & 0 & 0 & 0 & \cdots & \cdots \\ 0 & 0 & \ddots & & & \\ 0 & 0 & & \ddots & & \\ \vdots & \vdots & & & 0 & 1 \\ \vdots & \vdots & & & -1 & 0 \end{pmatrix}$ : We get the *symplectic group*  $Sp(2n, \mathbf{R})$ .

We will see that the Lie algebra in these cases is the space of matrices  $B$  such that  $B^T Q + QB = 0$ .

**Example 4.7.** Consider the group  $O(1, 1)$ . We claimed above that its Lie algebra consisted of matrices  $B$  such that

$$B^T \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = - \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} B.$$

If  $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  then this implies  $a = d = 0$  and  $b = c$ , so the Lie algebra is

$$\mathfrak{o}(1, 1) = \left\{ \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix} : b \in \mathbf{R} \right\}.$$

The exponential of this matrix is

$$\begin{pmatrix} \cosh(b) & \sinh(b) \\ \sinh(b) & \cosh(b) \end{pmatrix}.$$

#### 4.2.2 $GL(n, \mathbf{C}) \subset GL(2n, \mathbf{R})$

Note that  $GL(n, \mathbf{C})$  is isomorphic to  $G := \{A \in GL(2n, \mathbf{R}) : AJ = JA\}$  for  $J$  as above. The isomorphism simply replaces each complex matrix entry  $a + ib$  with the 2-by-2 real matrix  $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ . If we write  $F: GL(n, \mathbf{C}) \rightarrow G$  for this isomorphism then it has the property that

$$F(A^\dagger) = (F(A))^T.$$

### 4.3 The Lie algebra

Given a matrix group  $G$ , our goal is to understand the Lie algebra  $\mathfrak{g}$ , that is the set of all matrices  $A$  such that  $\exp(tA) \in G$  for all  $t \in \mathbf{R}$ . This set will turn out to be the vector subspace of  $\mathfrak{gl}(n, \mathbf{R})$  comprising vectors tangent to  $G$ .

We first prove a short technical lemma which will be useful.

**Lemma 4.8.** *Let  $H$  be a matrix group. If  $h_n$  is a sequence in  $\mathfrak{gl}(n, \mathbf{R})$  such that  $\exp(h_n) \in H$ ,  $h_n \rightarrow 0$  and  $\frac{h_n}{|h_n|} \rightarrow v$  then  $\exp(tv) \in H$  for all  $t \in \mathbf{R}$ .*

**Proof.** Fix  $t \in \mathbf{R}$ . Let  $m_n$  be the largest integer less than  $t/|h_n|$  (since  $|h_n| \rightarrow 0$ ,  $m_n \rightarrow \infty$ ). Then

$$\frac{t}{|h_n|} - 1 \leq m_n \leq \frac{t}{|h_n|} + 1$$

so

$$t - |h_n| \leq m_n |h_n| \leq t + |h_n|$$

and hence  $m_n |h_n| \rightarrow t$ . Therefore  $\exp(m_n h_n) \rightarrow \exp(tv)$ . Since

$$\exp(m_n h_n) = (\exp(h_n))^{m_n} \in H,$$

and  $H$  is a closed subset of  $GL(n, \mathbf{R})$ , we deduce that  $\exp(tv) = \lim_{n \rightarrow \infty} \exp(m_n h_n) \in H$ .  $\square$

**Theorem 4.9.** *Let  $G \subset GL(n, \mathbf{R})$  be a matrix group. Define*

$$\mathfrak{g} = \{v \in \mathfrak{gl}(n, \mathbf{R}) : \exp(tv) \in G \text{ for all } t \in \mathbf{R}\}.$$

*Then  $\mathfrak{g}$  is a vector subspace of  $\mathfrak{gl}(n, \mathbf{R})$ .*

**Proof.** We need to check that we can:

1. rescale elements of  $\mathfrak{g}$  (i.e.  $w \in \mathfrak{g}, t \in \mathbf{R}$  implies  $tw \in \mathfrak{g}$ ) - this is obvious from the definition.
2. add elements of  $\mathfrak{g}$  (i.e.  $w_1, w_2 \in \mathfrak{g}$  implies  $w_1 + w_2 \in \mathfrak{g}$ ).

So we need to prove that if  $w_1$  and  $w_2$  satisfy  $\exp(tw_1) \in G$  and  $\exp(tw_2) \in G$  for all  $t$  then  $\exp(t(w_1 + w_2)) \in G$  for all  $t$ .

First note that  $\gamma(t) = \exp(tw_1) \exp(tw_2) \in G$  and for sufficiently small  $t$ ,  $\gamma(t)$  is contained in the image of  $\exp$ , so  $\gamma(t) = \exp(f(t))$  for some  $f(t) \rightarrow 0$  as  $t \rightarrow 0$ . Since  $\exp(tw_i) = 1 + tw_i + \mathcal{O}(t^2)$  for  $i = 1, 2$  we have

$$\exp(f(t)) = 1 + f(t) + \cdots = 1 + t\dot{f}(0) + \mathcal{O}(t^2) = 1 + t(w_1 + w_2) + \mathcal{O}(t^2)$$

and hence  $\lim_{t \rightarrow 0} \frac{f(t)}{t} = w_1 + w_2$ . In particular,  $|f(t)| = t|w_1 + w_2| + \mathcal{O}(t^2)$ , so  $\lim_{t \rightarrow 0} \frac{f(t)}{|f(t)|} = \lim_{t \rightarrow 0} \frac{f(t)}{t} \frac{t}{|w_1 + w_2|} = \frac{w_1 + w_2}{|w_1 + w_2|} =: v$ . So if we set  $h_n = f(1/n)$  then Lemma 4.8 tells us that  $\exp(tv) \in H$  for all  $t$  and hence  $\exp(t(w_1 + w_2)) \in H$  for all  $t$ .  $\square$

**Definition 4.10.** Given a matrix group  $G \subset GL(n, \mathbf{R})$ , the vector space

$$\mathfrak{g} = \{v \in \mathfrak{gl}(n, \mathbf{R}) : \exp(tv) \in G \text{ for all } t \in \mathbf{R}\} \subset \mathfrak{gl}(n, \mathbf{R})$$

is called the *Lie algebra* of  $G$ .



**Example 4.11.** •  $G = U(n)$ , we have seen that  $\mathfrak{g} = \mathfrak{u}(n)$ , the space of skew-Hermitian matrices.

- $G = SL(n, \mathbf{R})$ , we will see in Question Sheet 2 that  $\mathfrak{g} = \mathfrak{sl}(n, \mathbf{R})$ , the space of trace-free matrices.
- $G = SU(n)$ , we get  $\mathfrak{g} = \mathfrak{su}(n)$ , the space of tracefree skew-Hermitian matrices.
- $G = O(n)$ , we will see in Question Sheet 2 that  $\mathfrak{g} = \mathfrak{so}(n)$ , the space of antisymmetric matrices.
- $G = SO(n)$ , we get again that  $\mathfrak{g} = \mathfrak{so}(n)$ . This is because  $O(n) = SO(n) \cup \tau SO(n)$  where  $\tau \in O(n)$  is a reflection matrix with determinant  $-1$ ; the elements in  $\tau SO(n)$  can never be in the image of the exponential map.

## 4.4 Exponential charts on matrix groups

So much of what we will say in this course involves working in local coordinates. The exponential map allows us to define some very convenient coordinates on a matrix group near the identity.

**Definition 4.12.** If  $G$  is a matrix group with Lie algebra  $\mathfrak{g}$  and  $0 \in B \subset \mathfrak{g}$ ,  $1 \in C \subset G$  are neighbourhoods such that  $\exp: B \rightarrow C$  is a bijection then we call  $\exp: B \rightarrow C$  an *exponential chart* for  $G$  near 1.

Theorem 3.8 tells us precisely that there is an exponential chart  $\exp: U' \rightarrow V'$  for  $GL(n, \mathbf{R})$ . If  $G \subset GL(n, \mathbf{R})$ , we would like to use this to define a coordinate chart  $U' \cap \mathfrak{g} \xrightarrow{\exp} V' \cap G$  on  $G$ , which identifies a neighbourhood of  $1 \in G$  with a neighbourhood of  $0 \in \mathfrak{g}$ . This turns out to be slightly tricky.

**Theorem 4.13.** *There exist neighbourhoods  $0 \in U' \subset \mathfrak{gl}(n, \mathbf{R})$  and  $1 \in V' \subset GL(n, \mathbf{R})$  such that*

$$\exp|_{U' \cap \mathfrak{g}}: U' \cap \mathfrak{g} \rightarrow G \cap V'$$

*is an exponential chart.*

**Proof.** If  $U'$  and  $V'$  are given by Theorem 3.8 then certainly this map is injective (as  $\exp|_{U'}$  is injective). However, since we have restricted the domain it is not clear that it is surjective. Suppose that the result were not true (we will find a contradiction); then for any  $U', V'$  we could find a point  $g \in V' \cap G$  not contained in  $\exp(\mathfrak{g})$ . By letting the radius of  $V'$  shrink we can ensure that there is a sequence  $g_i \in G \setminus \exp(\mathfrak{g})$  such that  $g_i \rightarrow 1$ .

To rule out the existence of such a sequence we make use of the following lemma:

**Lemma 4.14.** *Suppose that  $\mathfrak{gl}(n, \mathbf{R}) = W_1 \oplus W_2$  for a pair of complementary subspaces  $W_1, W_2 \subset \mathfrak{gl}(n, \mathbf{R})$ . Then there are open neighbourhoods  $U'$  of  $0 \in W_1 \oplus W_2$  and  $V'$  of  $1 \in GL(n, \mathbf{R})$  such that the map*

$$F: W_1 \oplus W_2 \rightarrow GL(n, \mathbf{R}), \quad F(w_1, w_2) = \exp(w_1) \exp(w_2)$$

*is a diffeomorphism  $F|_{U'}: U' \rightarrow V'$ .*

**Proof.** The proof is similar to Theorem 3.8. See Question Sheet 2. □

In our case, we take  $W_1 = \mathfrak{g}$  and  $W_2$  is a complement. Our sequence  $g_i$  eventually lies in the image of  $F$ , and so  $g_i = \exp(w_{1,i}) \exp(w_{2,i})$  for some sequences  $w_{1,i} \in \mathfrak{g}$ ,  $0 \neq w_{2,i} \in W_2$ . Note that  $g_i$  and  $\exp(-w_{1,i})$  are both in  $G$ , hence  $\exp(w_{2,i}) \in G$  even though  $w_{2,i}$  is nonzero in the complement of  $\mathfrak{g}$ .

The sequence  $\tilde{w}_i = w_{2,i}/\|w_{2,i}\|$  satisfies  $\|\tilde{w}_i\| = 1$  and therefore has a convergent subsequence with limit  $w \in W_2$  satisfying  $\|w\| = 1$ . Lemma 4.8 now implies that  $\exp(tw) \in G$  for all  $t \in \mathbf{R}$ . Therefore  $w \in \mathfrak{g}$  by definition. This contradicts the fact that  $w \in W_2$ .  $\square$

*Remark 4.15.* One obtains an exponential chart near any  $g \in G$  using  $v \mapsto g \exp(v)$ . The proof is the same as for the chart near  $1 \in G$ , except that in Lemma 4.14 you use the map

$$F_g: W_1 \oplus W_2 \rightarrow GL(n, \mathbf{R}), \quad F_g(w_1, w_2) = g \exp(w_1) \exp(w_2).$$

## 4.5 Tangent spaces

There is a more geometric characterisation of  $\mathfrak{g}$ .

**Definition 4.16.** Let  $\gamma = (\gamma_1, \dots, \gamma_n): (-\epsilon, \epsilon) \rightarrow \mathbf{R}^n$  be a (continuously differentiable) path. Recall that the tangent vector  $\dot{\gamma}(t)$  to  $\gamma$  at  $\gamma(t)$  is the vector

$$\dot{\gamma}(t) = \left( \frac{d\gamma_1}{dt}(t), \dots, \frac{d\gamma_n}{dt}(t) \right).$$

Note that by definition,  $\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon}(\gamma(\epsilon) - \gamma(0)) \rightarrow \dot{\gamma}(0)$ .

**Definition 4.17.** Let  $X \subset \mathbf{R}^n$  be a subset and  $x \in X$ . If  $\gamma: (-\epsilon, \epsilon) \rightarrow \mathbf{R}^n$  is a path with  $\gamma(-\epsilon, \epsilon) \subset X$  and  $\gamma(0) = x$  then we say  $\dot{\gamma}(0)$  is a tangent vector to  $X$  at  $x$ . The tangent cone of  $X$  at the point  $x \in X$  is the set of all tangent vectors to  $X$  at  $x$ . In the case when this cone is actually a subspace we call it the tangent space.

**Proposition 4.18.** *The vector space  $\mathfrak{g}$  is the tangent space of  $G$  at 1. More generally, the vector space  $g\mathfrak{g}$  is the tangent space of  $G$  at  $g \in G$ .*

**Proof.** Suppose that  $v \in \mathfrak{g}$ . Then  $g \exp(tv) \in G$  for all  $t$  so  $\gamma(t) = g \exp(tv)$  is a path in  $G$  with  $\gamma(0) = g$  and  $\dot{\gamma}(t) = gv \exp(tv)$  so  $\dot{\gamma}(0) = gv$ . Thus  $v$  is a tangent vector at  $g$ .

Conversely, suppose that  $gv$  is a tangent vector to  $G$  at  $g$ . Then there exists  $\gamma: (-\epsilon, \epsilon) \rightarrow G$  with  $\gamma(0) = g$  and  $\dot{\gamma}(0) = gv$ . The path  $g^{-1}\gamma$  satisfies  $g^{-1}\gamma(0) = 1$  and  $g^{-1}\dot{\gamma}(0) = v$ . By Theorem 3.8, for sufficiently small  $t$ ,  $g^{-1}\gamma(t) = \exp(f(t))$  for some  $f(t) \in \mathfrak{g}$  with  $f(0) = 0$ .

Define  $h_n = f(1/n)$ . Then  $h_n = \log(g^{-1}\gamma(1/n)) = g^{-1}\gamma(1/n) + \mathcal{O}(1/n^2) = g^{-1}\dot{\gamma}(0)/n + \mathcal{O}(1/n^2)$  so  $|h_n| = |g^{-1}\dot{\gamma}(0)|/n + \mathcal{O}(1/n^2)$ . Therefore

$$\lim_{n \rightarrow \infty} \frac{1}{|h_n|} (\exp(h_n) - 1) = \lim_{n \rightarrow \infty} \frac{1}{|g^{-1}\dot{\gamma}(0)|} \frac{g^{-1}\gamma(1/n) - g^{-1}\gamma(0)}{1/n} = \frac{g^{-1}\dot{\gamma}(0)}{|g^{-1}\dot{\gamma}(0)|}.$$

Let us define  $v := \frac{\dot{\gamma}(0)}{|\dot{\gamma}(0)|}$ . We also have, by Taylor expanding  $\exp$ , that  $\lim_{n \rightarrow \infty} \frac{1}{|h_n|} (\exp(h_n) - 1) = \lim_{n \rightarrow \infty} \frac{h_n}{|h_n|}$ . Thus

$$\lim_{n \rightarrow \infty} \frac{h_n}{|h_n|} = v$$

and the sequence  $h_n$  satisfies the assumptions of Lemma 4.8. We deduce that  $\exp(tv) \in G$  for all  $t$ . Hence  $v \in \mathfrak{g}$ , hence  $\dot{\gamma}(0) \in g\mathfrak{g}$ .  $\square$

Here is a quick way to find tangent spaces; for a proof, see multivariable calculus.

**Lemma 4.19.** *Let  $F: \mathbf{R}^n \rightarrow \mathbf{R}^m$  be a smooth map with  $F(p) = q$  and  $F(p+v) = q + d_p F(v) + \dots$ . Here the linear map  $d_p F$  is the one whose matrix is*

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1}(p) & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}.$$

*If  $d_p F$  is surjective then the tangent space to  $F^{-1}(q)$  at  $p$  is  $\ker(d_p F)$ .*

**Corollary 4.20.** *The tangent space of  $U(n)$  at 1 is the space of skew-Hermitian matrices*

$$\mathfrak{u}(n) = \{v \in \mathfrak{gl}(n, \mathbf{C}) : v^\dagger = -v\}.$$

**Proof.** Let  $H = \{v \in \mathfrak{gl}(n, \mathbf{C}) : v^\dagger = v\}$  be the space of Hermitian matrices. Consider the map

$$F: \mathfrak{gl}(n, \mathbf{C}) \rightarrow H, \quad F(A) = A^\dagger A.$$

By definition,  $U(n) = F^{-1}(1)$ . We have

$$F(1 + tB) = (1 + tB)(1 + tB^\dagger) = 1 + tB + tB^\dagger + \mathcal{O}(t^2)$$

so  $d_1 F(B) = B + B^\dagger$ . This is surjective: if  $C \in H$  then  $C = (\frac{1}{2}C) + (\frac{1}{2}C)^\dagger = d_1 F(C/2)$ . Therefore the tangent space of  $U(n)$  at 1 is

$$\ker d_1 F = \{B \in \mathfrak{gl}(n, \mathbf{C}) : B + B^\dagger = 0\} = \mathfrak{u}(n).$$

□

## 4.6 The Lie bracket

**Lemma 4.21.** *Given a matrix group  $G$ , the subspace  $\mathfrak{g}$  is preserved by the Lie bracket.*

**Proof.** If  $X, Y \in \mathfrak{g}$  then  $C_{s,t} = [\exp(sX), \exp(tY)] = \exp(sX) \exp(tY) \exp(-sX) \exp(-tY) \in G$ . By the part of the BCH formula we actually checked, we have

$$\exp(sX) \exp(tY) = \exp\left(sX + tY + st\frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]]\right),$$

so

$$C_{s,t} = \exp(st[X, Y] + st(\mathcal{O}(s) + \mathcal{O}(t))) \in G.$$

If we set  $\gamma(u) := C_{\sqrt{u}, \sqrt{u}}$  then  $\gamma(u) = \exp(u[X, Y] + \mathcal{O}(u^{3/2}))$  defines a path in  $G$  whose tangent vector at  $\gamma(0) = 1$  is  $[X, Y]$ . Therefore  $[X, Y] \in \mathfrak{g}$ . □

Let us summarise everything we have learned in one theorem:

**Theorem 4.22.** *If  $G$  is a matrix group, the set*

$$\mathfrak{g} := \{v \in \mathfrak{gl}(n, \mathbf{R}) : \exp(tv) \in G \forall t \in \mathbf{R}\}$$

*is a vector space, indeed it is the tangent space of  $G$  at the identity. Moreover it is closed under the operation of Lie bracket, so inherits the structure of a Lie algebra. The exponential map carries  $\mathfrak{g}$  to  $G$  and is a local diffeomorphism from a neighbourhood of  $0 \in \mathfrak{g}$  to a neighbourhood of  $1 \in G$ .*

# Lecture 5: Smooth homomorphisms

## 5.1 Smoothness in exponential charts

Let  $G$  and  $H$  be matrix groups and suppose that  $\phi: G \rightarrow H$  is a homomorphism. We wanted a smoothness condition on this homomorphism. A homomorphism is said to be smooth if it is smooth when written in exponential charts:

**Definition 5.1.** Let  $G_1$  and  $G_2$  be Lie groups with Lie algebras  $\mathfrak{g}_1$  and  $\mathfrak{g}_2$  respectively. Let  $F: G_1 \rightarrow G_2$  be a homomorphism. Let  $\exp: B_i \rightarrow C_i$ ,  $B_i \subset \mathfrak{g}_i$ ,  $C_i \subset G_i$  be exponential charts with  $F(C_1) = C_2$  (this can always be achieved by shrinking  $C_1$  and  $C_2$ ). A homomorphism  $F: G_1 \rightarrow G_2$  is *smooth* if

$$f := \exp^{-1} \circ F \circ \exp: B_1 \rightarrow C_1 \rightarrow C_2 \rightarrow B_2$$

is smooth.

*Remark 5.2.* Note that by definition we have

$$\exp(f(X)) = F(\exp X)$$

for any  $X$  in the domain of an exponential chart.

This may seem like a lot of effort to go to just to define a smooth map between matrix groups, but we only know how to define smooth maps between open subsets of vector spaces (like  $B_1, B_2$ ) where we have Euclidean coordinates and can take partial derivatives.

**Example 5.3.** Consider the homomorphism  $R: SU(2) \rightarrow SO(3)$  which sends  $\exp(\theta M_u)$  to the rotation through  $2\theta$  around  $u = (x, y, z)$ , studied in Question Sheet 1. We saw that if

$K_u = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \in \mathfrak{so}(3)$  then  $\exp(2\theta K_u) \in SO(3)$  also rotates around  $u$  by  $2\theta$ , so

$$R(\exp(\theta M_u)) = \exp(2\theta K_u).$$

This means that, in an exponential neighbourhood, defining  $r = \exp^{-1} \circ R \circ \exp$ , we have

$$r \begin{pmatrix} ix & y + iz \\ -y + iz & -ix \end{pmatrix} = 2 \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}.$$

In this example, the map  $R$  looks extremely simple in an exponential chart: it is linear! We will see shortly that this is no coincidence: a smooth homomorphism of Lie groups is linear in an exponential chart.

## 5.2 One-parameter subgroups

The group  $\mathbf{R}$  under addition can be realised as a closed subgroup of  $GL(2, \mathbf{R})$ :

$$\mathbf{R} \cong \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} : x \in \mathbf{R} \right\}, \quad \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x+y \\ 0 & 1 \end{pmatrix}.$$

**Definition 5.4.** A smooth homomorphism  $\phi: \mathbf{R} \rightarrow G$  is called a one-parameter subgroup.

**Example 5.5.** If  $X \in \mathfrak{g}$  then  $\phi(t) = \exp(tX)$  is a one-parameter subgroup, because

$$\exp((s+t)X) = \exp(sX) \exp(tX). \quad (1)$$

This is manifestly clear for matrix exponentials because  $sX$  and  $tX$  commute.

*Remark 5.6.* Another way to see it is this, using the fact that  $\dot{\phi}(t) = X\phi(t)$ :

- $\psi(s) := \phi(s+t)$  solves  $\frac{d\psi}{ds} = X\psi(s)$  and  $\psi(0) = \phi(t)$ ,
- $\theta(s) = \phi(s)\phi(t)$  solves  $\frac{d\theta}{ds} = X\theta(s)$  and  $\theta(0) = \phi(t)$ ,
- an ODE has a unique solution with given initial condition. Therefore  $\theta = \psi$  and Equation (1) holds, so  $\phi$  is a homomorphism.

This is a powerful strategy of proof, crucially using calculus and the uniqueness of solutions to ODEs with given initial conditions. Here we use it again, to prove that *all* one-parameter subgroups arise as  $\exp(tX)$  for some  $X \in \mathfrak{g}$ .

**Proposition 5.7.** Suppose that  $\phi: \mathbf{R} \rightarrow G$  is a one-parameter subgroup. Then  $\phi(t) = \exp(tX)$  for some  $X \in \mathfrak{g}$ .

**Proof.** Differentiate the homomorphism condition,  $\phi(s+t) = \phi(s)\phi(t)$ , with respect to  $s$  at  $s = 0$ . This gives

$$\dot{\phi}(t) = \dot{\phi}(0)\phi(t)$$

If we set  $X := \dot{\phi}(0)$  then we see that  $\phi(t)$  satisfies the ODE  $\dot{\phi}(t) = X\phi(t)$  and the initial condition  $\phi(0) = 1$ . Thus  $\phi(t) = \exp(tX)$ .  $\square$

## 5.3 Linearity in exponential charts

**Theorem 5.8.** Suppose that  $F: G_1 \rightarrow G_2$  is a smooth homomorphism of matrix groups whose Lie algebras are  $\mathfrak{g}_i$ , let  $\exp: B_i \rightarrow C_i$ ,  $i = 1, 2$ , be exponential charts and let  $f: B_1 \rightarrow B_2$  be the map  $F$  viewed in exponential coordinates. Then  $f$  is a linear map.

**Proof.** For any  $X \in \mathfrak{g}_1$ ,  $\exp(tX)$  is a one-parameter subgroup of  $G_1$  and, since  $F$  is a smooth homomorphism,  $F(\exp(tX))$  is a one-parameter subgroup of  $G_2$ . By Proposition 5.7, there exists  $Y \in \mathfrak{g}_2$  such that  $F(\exp(tX)) = \exp(tY)$ .

Since  $\exp(f(X)) = F(\exp X)$  for  $X \in B_1$  we have that  $\exp(f(tX)) = F(\exp(tX)) = \exp(tY)$  for all sufficiently small  $t$ . Taking logarithms (as we are in the image of an exponential chart) we see that  $f(tX) = tY$ . Differentiating this at  $t = 0$  we see that  $d_0 f(X) = Y$ . Therefore  $f(tX) = d_0 f(tX)$ . So  $f$  equals its own linearisation,  $d_0 f$ . In particular it is linear.  $\square$

**Definition 5.9.** If  $F: G_1 \rightarrow G_2$  is a homomorphism, denote by  $F_*: \mathfrak{g}_1 \rightarrow \mathfrak{g}_2$  the linear map  $d_0 f$ . We call this variously the *linearisation* of  $F$  at 1, the *induced map on Lie algebras* or the *differential* of  $F$  at the identity. We make this change in notation because  $f$  is only defined locally in an exponential chart.

Recall that by definition,  $\exp(f(X)) = F(\exp(X))$  for all  $X \in B_1$  and hence  $F(\exp(X)) = \exp(F_* X)$  for all  $X \in B_1$ . In fact:

**Lemma 5.10.**

$$F(\exp(X)) = \exp(F_*(X))$$

for all  $X \in \mathfrak{g}_1$ .

**Proof.** Since  $F(\exp(tX))$  is a one-parameter subgroup it has the form  $\exp(tY)$  and for sufficiently small  $t$ ,  $tX \in B_1$  so  $F(\exp(tX)) = \exp(tF_*(X))$ , therefore  $tF_*(X) = tY$  and  $Y = F_*(X)$ . Now setting  $t = 1$  gives the result.  $\square$

## 5.4 Lie algebra homomorphisms

So any smooth homomorphism  $F$  of matrix groups has the form  $F(\exp X) = \exp(F_*(X))$  for some linear map  $F_*$  on Lie algebras. Which linear maps  $F_*$  arise? Since  $F$  has to preserve the group structure, there should be an algebraic structure on the Lie algebra (hence the name) which is preserved by  $F_*$ . Given the BCH formula, it should be no surprise that this turns out to be the Lie bracket.

**Definition 5.11.** A homomorphism of Lie algebras  $\mathfrak{g}, \mathfrak{h}$  is a linear map  $f: \mathfrak{g} \rightarrow \mathfrak{h}$  such that  $[f(X), f(Y)] = f([X, Y])$  for all  $X, Y \in \mathfrak{g}$ .

**Theorem 5.12.** If  $F: G_1 \rightarrow G_2$  is a smooth homomorphism of matrix groups and  $F_*: \mathfrak{g}_1 \rightarrow \mathfrak{g}_2$  is its linearisation then

$$F_*[X, Y] = [F_*X, F_*Y].$$

In other words, the linearisation of a smooth homomorphism of matrix groups is a Lie algebra homomorphism.

**Proof.** Recall from the proof of Lemma 4.21 that

$$\exp(tX) \exp(tY) \exp(-tX) \exp(-tY) = \exp(t^2[X, Y] + o(t^2)) \quad (2)$$

Applying  $F$  and using the fact it is a homomorphism we get

$$F(\exp(tX))F(\exp(tY))F(\exp(-tX))F(\exp(-tY)) = F(\exp(t^2[X, Y] + o(t^2))).$$

Using the fact that  $F(\exp(X)) = \exp(F_*X)$  we get

$$\exp(tF_*X) \exp(tF_*Y) \exp(-tF_*X) \exp(-tF_*Y) = \exp(t^2F_*[X, Y] + o(t^2)).$$

Applying Equation (2) in  $G_2$  to the left-hand side we get

$$\exp(tF_*X) \exp(tF_*Y) \exp(-tF_*X) \exp(-tF_*Y) = \exp(t^2[F_*X, F_*Y] + o(t^2))$$

and comparing the two expressions for  $\exp(tF_*X) \exp(tF_*Y) \exp(-tF_*X) \exp(-tF_*Y)$  we see that

$$F_*[X, Y] = [F_*X, F_*Y].$$

$\square$

## Lecture 6: Lie's theorem

We have seen that linearisations of smooth homomorphisms are Lie algebra homomorphisms. Do all Lie algebra homomorphisms arise this way?

### 6.1 Example: $U(1)$

**Example 6.1.** Let  $F: U(1) \rightarrow U(1)$  be a homomorphism. The Lie algebra is  $i\mathbf{R}$  (skew-Hermitian 1-by-1 matrices!) and the linear map  $F_*: i\mathbf{R} \rightarrow i\mathbf{R}$  is just  $x \mapsto \lambda x$  for some  $\lambda \in \mathbf{R}$ . The Lie bracket vanishes identically on  $i\mathbf{R}$  since  $[ix, iy] = xy[i, i] = 0$ , so the Lie algebra homomorphism condition is vacuous.

Since  $F(\exp(ix)) = \exp(F_*(ix)) = \exp(i\lambda x)$ , we see that there is an extra condition that has to be satisfied in order for this equation to make sense. This comes from the fact that  $\exp(i2\pi) = 1$  and  $F(1) = 1$ , so  $\exp(i2\pi\lambda) = 1$ . This means  $\lambda \in \mathbf{Z}$ .

Any  $n \in \mathbf{Z}$  does indeed define a homomorphism  $F_n: e^{ix} \mapsto e^{inx}$ . So we have proved

**Proposition 6.2.** *Given a smooth homomorphism  $F: U(1) \rightarrow U(1)$  there exists an integer  $n$  such that  $F = F_n$ .*

### 6.2 Simply-connectedness

In this example it was not the case that every Lie algebra homomorphism exponentiated to give a group homomorphism: there were  $\mathbf{R}$  possible homomorphisms but only a discrete set  $\mathbf{Z} \subset \mathbf{R}$  exponentiated successfully. The problem is actually to do with the topology of the group  $U(1)$ : loops in  $U(1)$  have a well-defined winding number around the central hole. This problem disappears if we make an extra topological assumption about the group:

**Definition 6.3.** A subset  $X \subset \mathbf{R}^n$  is simply-connected if for any loop  $\gamma: [0, 1] \rightarrow X$  with  $\gamma(0) = \gamma(1) = x \in X$  there is a continuous map  $H: [0, 1] \times [0, 1] \rightarrow X$  such that

- $H(0, t) = H(1, t) = x$ ,
- $H(s, 0) = x, H(s, 1) = \gamma(s)$ .

We say that  $H$  is a nullhomotopy of  $\gamma$  and we should imagine that  $H$  contracts  $\gamma(s)$  to a point through a family of continuous loops  $\gamma_t(s) = H(s, t)$ . Note that we could work with smooth paths and smooth homotopies (just by approximating continuous maps by smooth maps using the Stone-Weierstrass theorem in local coordinates).

This fails for  $X = U(1) \subset \mathbf{C}$  because the loop  $\gamma(\theta) = e^{i2\pi\theta}$  is not contractible: the winding number does not change when you continuously deform  $\gamma$  but it equals 1 whereas the winding number for the constant path equals 0.

It turns out, amazingly, that this is the only obstruction to exponentiating homomorphisms

**Theorem 6.4** (Lie's theorem on homomorphisms). *Let  $G_1$  and  $G_2$  be path-connected matrix groups with Lie algebras  $\mathfrak{g}_1$  and  $\mathfrak{g}_2$ . Suppose that  $G_1$  is simply-connected. If  $f: \mathfrak{g}_1 \rightarrow \mathfrak{g}_2$  is a Lie algebra homomorphism then there exists a Lie group homomorphism  $F: G_1 \rightarrow G_2$  such that  $F_* = f$ .*

We will see a proof of this theorem on Question Sheet 4.

So does the theorem have any content? Are there any simply-connected groups?

**Example 6.5** (Topological aside). The 3-dimensional sphere is simply-connected: if  $\gamma$  is a smooth loop in  $S^3$  then  $\gamma([0, 1])$  cannot fill out the whole  $S^3$  as the image of  $\gamma$  is a measure zero set. Take  $p \in S^3 \setminus \gamma([0, 1])$  and stereographically project  $\gamma$  to  $\mathbb{R}^3$ ; any loop in  $\mathbb{R}^3$  is contractible to the origin (via radial rescaling) and projecting this nullhomotopy stereographically back to  $S^3$  gives a nullhomotopy of the original loop  $\gamma$ . Since  $SU(2)$  is topologically  $S^3$  this says that  $SU(2)$  is simply-connected. In fact if  $n > 1$  the quotient  $SU(n)/SU(n-1)$  is a sphere of dimension  $2n-1$  and hence simply-connected by the same argument; one can use this to show inductively that  $SU(n)$  is simply-connected (this argument uses the homotopy long exact sequence of the fibration  $SU(n-1) \rightarrow SU(n) \rightarrow S^{2n-1}$ ).

**Example 6.6.** Given any matrix group  $G$  one can take its simply-connected *universal cover*  $\tilde{G}$ . This is again a group, *locally isomorphic to  $G$*  in the sense that there are neighbourhoods  $1 \in C_1 \subset G$  and  $1 \in C_2 \subset \tilde{G}$  and a bijection  $F: C_1 \rightarrow C_2$  which satisfies  $F(gh) = F(g)F(h)$  whenever  $gh \in C_1$ . However, it is not necessarily a matrix group! For example, it turns out that the universal cover of  $SL(2, \mathbb{R})$  is not a matrix group (it doesn't embed homomorphically into any  $GL(n, \mathbb{R})$ ). Nonetheless, the local geometry/algebra looks very similar to that of a matrix group: one can define an exponential map and a Lie algebra. This motivates the introduction of *Lie groups*: smooth manifolds with a group structure where the multiplication and inversion maps are smooth. "Smooth manifold" just means that there is a system of local coordinate charts (like the exponential charts we built) which allow you to do calculus in local coordinates. In this general context, one can no longer multiply Lie algebra elements with group elements or with one another directly because they are no longer matrices. Instead one reinterprets all our constructions in terms of manifolds, pushforward maps and flows of vector fields. We will not do this in this course.

One benefit of working with Lie groups rather than matrix groups is:

**Theorem 6.7** (The Lie correspondence). *For any Lie algebra  $\mathfrak{g}$  there exists a path-connected, simply-connected Lie group  $G$  with Lie algebra  $\mathfrak{g}$ .*

One way to prove this is to first prove:

**Theorem 6.8** (Another theorem of Lie). *Suppose that  $\mathfrak{g}$  is the Lie algebra of  $G$ . For any Lie subalgebra  $\mathfrak{h} \subset \mathfrak{g}$  there exists a Lie subgroup  $H \subset G$  with Lie algebra  $\mathfrak{h}$ .*



**Theorem 6.9** (Ado-Hochschild). *Any finite-dimensional Lie algebra occurs as a subalgebra of  $\mathfrak{gl}(n, \mathbb{R})$  for some  $n$ .*

Now take a matrix group with Lie algebra  $\mathfrak{g}$  and take its universal cover to get Lie's corresponding group.

# Lecture 7: Representations of Lie groups

## 7.1 Definitions and examples

Let  $\mathbf{K}$  be a field. A  $\mathbf{K}$ -representation of a group  $G$  comprises a  $\mathbf{K}$ -vector space  $V$  and a homomorphism

$$R: G \rightarrow GL(V).$$

In other words, a representation is an assignment of a matrix to each group element such that the matrices multiply like the group elements.

**Example 7.1.** Any group admits a *zero representation*  $R = 0: G \rightarrow GL(0)$  on the 0-dimensional vector space. In fact any group admits a *trivial representation*  $R: G \rightarrow GL(V)$  on any vector space  $V$  such that  $R(g) = 1$  for all  $g \in G$ .

**Example 7.2.** The inclusion  $U(n) \rightarrow GL(n, \mathbf{C})$  is a  $\mathbf{C}$ -representation of  $U(n)$  on  $\mathbf{C}^n$ . The inclusion  $SO(n) \rightarrow GL(n, \mathbf{R})$  is a  $\mathbf{R}$ -representation of  $SO(n)$  on  $\mathbf{R}^n$ . These are called the standard representations of  $U(n)$  and  $SO(n)$ .

**Example 7.3.** Let  $G$  be a matrix group with Lie algebra  $\mathfrak{g}$ . We define the *adjoint representation* of  $G$  on  $\mathfrak{g}$  by

$$\text{Ad}: G \rightarrow GL(\mathfrak{g}), \quad g \mapsto \text{Ad}_g, \quad \text{Ad}_g v = g v g^{-1}.$$

We studied the adjoint representation of  $SU(2)$  extensively in Question Sheet 1. We saw that it is a 2-to-1 homomorphism  $SU(2) \rightarrow SO(3) \subset GL(3, \mathbf{R})$  (given by  $\exp(\theta M_u) \mapsto \exp(2\theta K_u)$ , where  $M_u = \begin{pmatrix} ix & y+iz \\ -y+iz & -ix \end{pmatrix}$  and  $K_u = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}$ ).

The adjoint representation will turn out to be the most important representation of all; understanding it will allow us to classify all representations of a given group and even classify semisimple groups.

We want to be able to compare different representations and to say when two are *isomorphic*.

**Definition 7.4.** Let  $R_1: G \rightarrow GL(V)$  and  $R_2: G \rightarrow GL(W)$  be representations. A *morphism of representations* comprises a linear map  $L: V \rightarrow W$  such that

$$L(R_1(g)v) = R_2(g)L(v)$$

for all  $g \in G$  and  $v \in V$ . A map of representations is called an isomorphism if  $L$  is an isomorphism.

**Example 7.5.** Let  $V$  be the standard representation  $\mathbf{R}^3$  of  $SO(3)$ , in which  $\exp(\theta K_u)$  acts like rotation by  $\theta$  around a unit vector  $u$ . Let  $W$  be the adjoint representation of  $SO(3)$  on  $\mathfrak{so}(3)$ , the Lie algebra of antisymmetric matrices  $K_v$ . Let  $L: V \rightarrow W$  be the map  $v \mapsto K_v$ . To check that this is an isomorphism of representations we need to prove

$$\text{Ad}_{\exp(\theta K_u)} K_v = K_{\exp(\theta K_u)v}, \quad \text{if } |u|^2 = 1.$$

The left-hand side is  $\exp(\theta K_u)K_v \exp(-\theta K_u)$  and we also have

$$\exp(\theta K_u) = 1 + K_u \sin \theta + (1 - \cos \theta)K_u^2, \quad [K_u, K_v] = K_{u \times v}, \quad K_u K_v K_u = -(u \cdot v)K_u.$$

Using these facts it's easy (but tedious) to check that

$$\exp(\theta K_u)K_v \exp(-\theta K_u) = K_v \cos \theta + K_{u \times v} \sin \theta + (1 - \cos \theta)(u \cdot v)K_u$$

which is then equal to  $K_{\exp(\theta K_u)v}$  by Rodrigues's formula. Therefore the standard representation of  $SO(3)$  is isomorphic to its adjoint representation.

This proof by calculation was so messy I that I didn't write it out. Later we will see a much easier way to prove that two representations of  $SU(2)$  are isomorphic - indeed we will classify representations of  $SU(2)$ .

## 7.2 Subrepresentations; irreducibility

A *subrepresentation* is a subspace  $W \subset V$  such that  $R(g)(w) \in W$  for all  $g \in G, w \in W$ . This defines a homomorphism  $\text{Res}_W R: G \rightarrow GL(W)$  where

$$\text{Res}_W R(g) = R(g)|_W.$$

**Example 7.6.** The group  $U(n)$  acts on  $\mathfrak{gl}(n, \mathbb{C})$  by conjugation, which gives a representation

$$U(n) \rightarrow GL(\mathfrak{gl}(n, \mathbb{C})).$$

Since the conjugate of an anti-Hermitian matrix by a unitary matrix is still anti-Hermitian, the matrices in the image of this representation preserve the subspace  $\mathfrak{u}(n) \subset \mathfrak{gl}(n, \mathbb{C})$  so this is a subrepresentation (indeed it is the adjoint representation of  $U(n)$ ). Moreover, since the conjugate of a tracefree matrix is tracefree these matrices also preserve the subspace  $\mathfrak{su}(n) \subset \mathfrak{u}(n)$  so this is another subrepresentation (both of  $\mathfrak{gl}(n, \mathbb{C})$  and of  $\mathfrak{u}(n)$ ).

**Example 7.7.** If  $R: G \rightarrow GL(V)$  is a representation of  $G$  then (a)  $R$  is a subrepresentation of  $R$  and (b) the zero representation  $G \rightarrow GL(0)$  is a subrepresentation (because the origin in  $V$  is fixed by every element  $R(g)$ ). A subrepresentation of  $R$  which is not equal to  $R$  or  $0$  is called a *proper subrepresentation*.

**Definition 7.8.** A representation is called *irreducible* if it admits no proper subrepresentations.

## 7.3 New representations from old

### 7.3.1 Direct sums

Given two representations  $R_1: G \rightarrow GL(V)$  and  $R_2: G \rightarrow GL(W)$  we can construct a third representation

$$R_1 \oplus R_2: G \rightarrow GL(V \oplus W)$$

by defining

$$(R_1 \oplus R_2)(g) = \begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix},$$

that is  $(R_1 \oplus R_2)(g)(v, w) = (R_1(g)v, R_2(g)w)$ . It is easy to see that if  $R_1 \neq 0$  and  $R_2 \neq 0$  then  $R_1 \oplus R_2$  is not irreducible: it contains  $R_1$  and  $R_2$  as proper subrepresentations.

### 7.3.2 Duals

If  $V$  is a finite-dimensional  $\mathbf{K}$ -vector space then its  $\mathbf{K}$ -dual is the space  $V^*$  of  $\mathbf{K}$ -linear functionals

$$f: V \rightarrow \mathbf{K}.$$

If  $R: G \rightarrow GL(V)$  is a representation then we can define a representation

$$R^*: G \rightarrow GL(V^*)$$

by defining  $R^*(g)f$  to be the linear functional whose value on  $v$  is

$$(R^*(g)f)(v) = f(R(g^{-1})v).$$

It is an exercise to check that this defines a representation. Observe that  $(R^*(g)f)(R(g)v) = f(v)$ , so  $R^*(g)f$  is the function  $f$  viewed in the new frame of reference where vectors  $v$  look like  $R(g)v$ .

If we pick a basis of  $V$  and identify it with column vectors in  $\mathbf{K}^n$  then its dual  $V^*$  is identified with the space of row vectors: a row vector  $(x_1 \ \cdots \ x_n)$  gives a linear functional on  $V$  via

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \mapsto (x_1 \ \cdots \ x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i y_i.$$

If  $R(g)$  is an  $n$ -by- $n$  matrix then we can let it act from the left on column vectors (giving a representation because  $R(gh)v = (R(g)R(h))v = R(g)(R(h)v)$ ) or from the right on row vectors (giving an “antirepresentation” because  $v^T R(gh) = v^T (R(g)R(h)) = (v^T R(g))R(h)$  instead of  $(v^T R(h))R(g)$ ). To get a representation we use  $g \mapsto R(g)^{-1}$  because  $R(gh)^{-1} = R(h)^{-1}R(g)^{-1}$ .

### 7.3.3 Tensor products

Given two vector spaces  $V$  and  $W$  over  $\mathbf{K}$  we can form the tensor product  $V \otimes W$  as follows. Let  $e_i$  and  $f_j$  be bases of  $V$  and  $W$  and take the vector space  $V \otimes W$  with a basis given by the symbols  $\{e_i \otimes f_j\}_{i \in I, j \in J}$ .

**Lemma 7.9.** *The bilinear map  $\psi: V \times W \rightarrow V \otimes W$ :*

$$\psi \left( \sum_i v_i e_i, \sum_j w_j f_j \right) = \sum_{i,j} (v_i w_j) (e_i \otimes f_j)$$

has the following universal property: any bilinear map  $h: V \times W \rightarrow X$  factors uniquely as  $h' \circ \psi$  for a linear map  $h': V \otimes W \rightarrow X$ :

$$\begin{array}{ccc} V \times W & \xrightarrow{\psi} & V \otimes W \\ & \searrow h & \downarrow h \\ & & X \end{array}$$

**Proof.** Let  $\{g_k\}_{k \in K}$  be a basis for  $X$ . Elementary linear algebra tells us that a linear map

$$h: \left( \sum_i v_i e_i, \sum_j w_j f_j \right) \mapsto \sum_k A_k(v, w) g_k$$

is bilinear if  $A_k(v, w) = \sum_{i,j} A_{ijk} v_i w_j$  (just expand the coefficient of  $w_j$  as linear functions of the  $v_i$ ). Thus

$$h \left( \sum_i v_i e_i, \sum_j w_j f_j \right) = \sum_{i,j,k} A_{ijk} v_i w_j g_k$$

factors as  $h' \circ \psi$  where

$$h'(e_i \otimes f_j) = \sum_k A_{ijk} g_k.$$

Uniqueness follows because if we used  $h''(e_i \otimes f_j) = \sum_k B_{ijk} g_k$  we would get different coefficients in our bilinear map ( $B_{ijk}$  instead of  $A_{ijk}$ ) and these coefficients can be recovered as the  $g_k$  component of  $h(e_i, f_j)$ .  $\square$

**Lemma 7.10.** Suppose there were another vector space  $M$  and a map  $m: V \times W \rightarrow M$  with the universal property that any bilinear map  $g: V \times W \rightarrow X$  factored uniquely as  $g' \circ m$  for a linear map  $g': M \rightarrow X$ . Then  $M$  would be canonically isomorphic to  $V \otimes W$ .

**Proof.** Compare the diagrams

$$\begin{array}{ccc} V \times W & \xrightarrow{\psi} & V \otimes W \\ & \searrow m & \downarrow m' \\ & & M \\ & \searrow \psi & \downarrow \psi \\ & & V \otimes W \end{array}$$

Clearly  $\psi' \circ m' \circ \psi$  is a factorisation of  $\psi$  but there is a unique such factorisation, namely  $\text{id} \circ \psi$ . Therefore  $\psi' \circ m' = \text{id}$  and, by the same argument,  $m' \circ \psi' = \text{id}$ . Therefore  $m'$  is an isomorphism  $V \otimes W \rightarrow M$  with inverse  $\psi'$ .  $\square$

**Remark 7.11.** It is vitally important that you don't go away thinking every element of  $V \otimes W$  is of the form  $v \otimes w$ . Such elements are called *pure tensors*. But for example,  $e_1 \otimes f_1 + e_2 \otimes f_2$  is a perfectly good element which is not a pure tensor. The pure tensors don't even span a vector subspace, instead they form a *subvariety* cut out by equations called the *Plücker relations* (see Sheet 5). Nonetheless, we will usually define linear maps on  $V \otimes W$  by defining them on a basis of pure tensors and extending linearly. For example...

**Definition 7.12.** If  $R_1: G \rightarrow GL(V)$  and  $R_2: G \rightarrow GL(W)$  are also representations of  $G$  then we get a representation  $R_1 \otimes R_2: G \rightarrow GL(V \otimes W)$  by defining

$$(R_1 \otimes R_2)(g)v \otimes w = R_1(g)v \otimes R_2(g)w$$

and extending  $\mathbf{K}$ -linearly to  $V \otimes W$ .

### 7.3.4 Hom spaces

Let  $\text{Hom}(V, W)$  denote the space of linear maps from  $V$  to  $W$ . If  $R: G \rightarrow GL(V)$  and  $S: G \rightarrow GL(W)$  are representations then one can define a representation

$$T: G \rightarrow GL(\text{Hom}(V, W))$$

by

$$(T(g)F)(v) = S(g)F(R(g^{-1})v), \quad F \in \text{Hom}(V, W), \quad v \in V.$$

We will see in Sheet 5 that this is isomorphic to the representation  $V^* \otimes W$ .

### 7.3.5 Symmetric powers

In this section we will work over a field  $\mathbf{K}$  of characteristic zero.

**Definition 7.13.** Consider the action of  $S_n$  on  $V^{\otimes n}$  which is defined on pure tensors by

$$\sigma(v_1 \otimes \cdots \otimes v_n) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}, \quad \sigma \in S_n.$$

The  $n$ th symmetric power of  $V$  is the subspace  $\text{Sym}^n(V) \subset V^{\otimes n}$  consisting of tensors which are fixed by every  $\sigma \in S_n$  under this action.

**Example 7.14.** Suppose that  $V = \mathbf{K}\langle e_1, e_2 \rangle$ . The tensor  $e_1 \otimes e_2 + e_2 \otimes e_1$  is an element of  $\text{Sym}^2(V)$  but the tensor  $e_1 \otimes e_2$  is not.

**Definition 7.15.** We define the *averaging map*

$$\text{Av}: V^{\otimes n} \rightarrow V^{\otimes n}, \quad \text{Av}(v) = \frac{1}{n!} \sum_{\sigma \in S_n} \sigma(v).$$

Note that  $\text{Av}$  is a projection onto its image  $\text{Sym}^n(V) \subset V^{\otimes n}$ .

**Example 7.16.** If  $V = \mathbf{K}\langle e_1, e_2 \rangle$ ,  $\text{Av}(e_1 \otimes e_2) = \frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1)$ .

**Proposition 7.17.** If  $R: G \rightarrow GL(V)$  is a representation then  $\text{Sym}^n(V)$  is a subrepresentation of  $V^{\otimes n}$ .

*Remark 7.18.* This implies that tensor products are never irreducible: they always admit subrepresentations.

**Proof.** We will prove that  $\text{Av}$  is a morphism of representations  $V^{\otimes n} \rightarrow V^{\otimes n}$ . On Sheet 5 we will see that the image of a morphism of representations  $V_1 \rightarrow V_2$  is a subrepresentation of  $V_2$ , concluding the proof.

To check that it is a morphism, we need to prove that

$$\text{Av}(R^{\otimes n}(g)v) = R^{\otimes n}(g) \text{Av}(v).$$

It suffices to check this on pure tensors  $v = v_1 \otimes \cdots \otimes v_n$ . Expanding out the right-hand side gives

$$\frac{1}{n!} \sum_{\sigma \in S_n} \sigma(R(g)v_1 \otimes \cdots \otimes R(g)v_n)$$

or, because  $\sigma$  simply permutes the factors,

$$\frac{1}{n!} \sum_{\sigma \in S_n} R(g)v_{\sigma(1)} \otimes \cdots \otimes R(g)v_{\sigma(n)}.$$

The left-hand side is

$$R^{\otimes n}(g) \frac{1}{n!} \sum_{\sigma \in S_n} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$$

which, by linearity and definition of  $R^{\otimes n}(g)$  equals

$$\frac{1}{n!} \sum_{\sigma \in S_n} R(g)v_{\sigma(1)} \otimes \cdots \otimes R(g)v_{\sigma(n)}.$$

This proves that  $\text{Av}$  is a morphism. □

We introduce some new notation for symmetric powers which will be convenient

$$v_1 \cdots v_n = \text{Av}(v_1 \otimes \cdots \otimes v_n).$$

For example  $xy = \frac{1}{2}(x \otimes y + y \otimes x)$ . This allows us to think of the  $n$ th symmetric power of  $V$  as being homogeneous polynomials in the elements of  $V$ . For example, elements of  $V^*$  are linear functions on  $V$  and elements of  $\text{Sym}^n(V^*)$  are homogeneous polynomial functions on  $V$ .

### 7.3.6 Exterior powers

**Definition 7.19.** Consider the action of  $S_n$  on  $V^{\otimes n}$  which is defined on pure tensors by

$$\sigma(v_1 \otimes \cdots \otimes v_n) = (-1)^{|\sigma|} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}, \quad \sigma \in S_n$$

where  $|\sigma|$  is the sign of the permutation. The  $n$ th exterior power of  $V$  is the subspace  $\Lambda^n(V) \subset V^{\otimes n}$  consisting of tensors which are fixed by every  $\sigma \in S_n$  under this action.

**Example 7.20.** Suppose that  $V = \mathbf{K}\langle e_1, e_2 \rangle$ . The tensor  $e_1 \otimes e_2 - e_2 \otimes e_1$  is an element of  $\Lambda^2(V)$  but the tensor  $e_1 \otimes e_2$  is not.

**Definition 7.21.** We define the *alternating map*

$$\text{Alt}: V^{\otimes n} \rightarrow V^{\otimes n}, \quad \text{Alt}(v) = \frac{1}{n!} \sum_{\sigma \in S_n} (-1)^{|\sigma|} \sigma(v).$$

Note that  $\text{Alt}$  is a projection onto its image  $\Lambda^n(V) \subset V^{\otimes n}$ .

**Example 7.22.** If  $V = \mathbf{K}\langle e_1, e_2 \rangle$ ,  $\text{Alt}(e_1 \otimes e_2) = \frac{1}{2}(e_1 \otimes e_2 - e_2 \otimes e_1)$ .

**Proposition 7.23.** If  $R: G \rightarrow GL(V)$  is a representation then  $\Lambda^n(V)$  is a subrepresentation of  $V^{\otimes n}$ .

The proof is the same as for symmetric powers.

We introduce some new notation for exterior powers which will be convenient

$$v_1 \wedge \cdots \wedge v_n = \text{Alt}(v_1 \otimes \cdots \otimes v_n).$$

For example  $x \wedge y = \frac{1}{2}(x \otimes y - y \otimes x)$ .

**Lemma 7.24.** If  $\dim_{\mathbf{K}} V = n$  then  $\dim_{\mathbf{K}} \Lambda^m V = \binom{n}{m}$ . In particular,  $\Lambda^m V = 0$  if  $m > n$ .

**Proof.** Suppose that  $e_1, \dots, e_n$  is a basis for  $V$ . A basis for  $\Lambda^m V$  is given by

$$e_{i_1} \wedge \cdots \wedge e_{i_m}$$

where  $i_1 < \dots < i_m$ . There are  $\binom{n}{m}$  of these basis vectors. □

*Remark 7.25.* The exterior powers of the dual,  $\Lambda^n V^*$ , are called *alternating forms*. These play a key role in differential geometry and topology. In the same way that vectors are the “infinitesimal version” of vector fields, alternating forms are the “infinitesimal version” of the ubiquitous *differential forms*.

## 7.4 Representations which are not irreducible

The direct sum of two nonzero representations is not irreducible. Is this the only way to construct representations which are not irreducible?

**Definition 7.26.** A representation  $R: G \rightarrow GL(V)$  is *completely reducible* if there exist irreducible subrepresentations  $V_1, \dots, V_k$  such that  $V = \bigoplus_{i=1}^k V_i$ .

There exist groups which admit representations which are not completely reducible.

**Example 7.27.** Consider the group  $\mathbf{C}$  under addition. This admits a representation on  $\mathbf{C}^2$ :

$$z \mapsto \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}.$$

This admits a trivial one-dimensional subrepresentation: the subspace

$$A = \left\{ \begin{pmatrix} a \\ 0 \end{pmatrix} : a \in \mathbf{C} \right\}$$

is preserved and  $\mathbf{C}$  acts trivially on it. If  $B$  is a complementary subrepresentation, spanned by a vector  $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ , ( $b_2 \neq 0$ ) then

$$\begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} b_1 + zb_2 \\ b_2 \end{pmatrix}$$



If this is a subrepresentation then for each  $z \in \mathbf{C}$  there exists a  $\lambda(z)$  such that

$$\begin{pmatrix} b_1 + zb_2 \\ b_2 \end{pmatrix} = \lambda(z) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Since  $b_2 = \lambda(z)b_2$ , this implies  $\lambda(z) = 1$ . Therefore  $b_1 + zb_2 = b_1$  for all  $z$ , which is a contradiction.

In Section 7.5 we will see that representations of *compact* Lie groups are always completely reducible.

## 7.5 Unitarity

### 7.5.1 Unitary representations

Recall that a Hermitian inner product on a complex vector space  $V$  is a map  $\langle, \rangle: V \times V \rightarrow \mathbf{C}$  such that

$$\langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle, \quad \langle w, v \rangle = \overline{\langle v, w \rangle}$$

for all  $\lambda, \mu \in \mathbf{C}$ ,  $u, v, w \in V$ . There is a further requirement of positive-definiteness: note that  $\langle v, v \rangle = \overline{\langle v, v \rangle}$ . Positive-definiteness is the requirement that

$$\langle v, v \rangle \geq 0$$

with equality if and only if  $v = 0$ .

**Definition 7.28.** A unitary representation of a Lie group  $G$  is a homomorphism  $G \rightarrow U(n)$  for some  $n$ . Equivalently, a unitary representation is a representation  $R: G \rightarrow GL(n, \mathbf{C})$  for which there exists a Hermitian inner product  $\langle, \rangle$  such that  $R(g)$  preserves  $\langle, \rangle$  for each  $g \in G$ . This is called an invariant Hermitian inner product.

**Lemma 7.29.** For any representation of a finite group  $R: G \rightarrow GL(n, \mathbf{C})$  there exists an invariant Hermitian inner product.

**Proof.** Let  $\langle, \rangle'$  be a Hermitian inner product on  $\mathbf{C}^n$  (not invariant). Define

$$\langle v, w \rangle = \sum_{g \in G} \langle R(g)v, R(g)w \rangle'$$

It is easy to check that this is a Hermitian inner product. Invariance is proved by considering

$$\langle hv, hw \rangle = \sum_{g \in G} \langle R(gh)v, R(gh)w \rangle'$$

and if we relabel  $gh = k$  then as  $g$  runs over  $G$  so does  $k$  so this sum is equal to

$$\sum_{k \in G} \langle R(k)v, R(k)w \rangle'$$

which equals  $\langle v, w \rangle$ . □

We cannot play such tricks with Lie groups because there are usually infinitely many elements. But for a certain class of Lie groups we can replace the sum by an integral; this is the class of *compact* Lie groups.

**Definition 7.30.** A manifold  $X$  is called *compact* if every collection  $\{U_i\}_{i \in I}$  of open sets which cover  $X$  admits a finite subcollection  $\{U_i\}_{i \in J \subset I}$  which also cover  $X$ .

**Example 7.31.**  $\mathbf{R}$  is not compact. For matrix groups, being compact is the same as being a bounded subset of  $GL(n, \mathbf{R})$ .

**Proposition 7.32** (Weyl's unitarian trick). *For any representation of a compact Lie group  $R: G \rightarrow GL(n, \mathbf{C})$  there exists an invariant Hermitian inner product.*

**Proof.** Starting from an arbitrary Hermitian inner product  $\langle, \rangle'$  we can define

$$\langle v, w \rangle = \frac{\int_{g \in G} \langle R(g)v, R(g)w \rangle' dg}{\int_{g \in G} dg}$$

provided we can define the integral  $\int_{g \in G} f(g) dg$  of a function  $f: G \rightarrow \mathbf{C}$  in such a way that  $\int_{g \in G} dg$  is finite. Moreover, for the relabelling trick to work to prove invariance, we need

$$\int_{g \in G} f(gh) dg = \int_{g \in G} f(g) dg$$

for any  $h \in G$ . An integral with these properties exists and is called the *Haar integral*. We won't prove the existence of the Haar integral. The idea of the proof is to use a *volume form* on the Lie algebra and translate it around the group via left multiplication; for details see Section 4.11 of Warner's *Foundations of differentiable manifolds and Lie groups*.  $\square$

**Remark 7.33.** Actually, one of the most important cases of the Haar integral is also the easiest to define: the Haar integral on  $U(1)$  is just  $\int_0^{2\pi} f(\theta) d\theta$ . Since the action of  $e^{i\phi} \in U(1)$  sends  $e^{i\theta}$  to  $e^{i(\theta+\phi)}$  we just need to observe that  $d(\theta + \phi) = d\theta$  which is clear for fixed  $\phi$  by the usual change of variables formula for integrals. Another important easy case is  $SU(2)$ , where the usual volume element on  $S^3$  is invariant.

## 7.5.2 Complete reducibility

Recall that a representation  $R: G \rightarrow GL(n, \mathbf{C})$  is called *completely reducible* if it can be written as a direct sum of irreducible subrepresentations.

**Proposition 7.34.** *Let  $G$  be a compact Lie group. Then any finite-dimensional complex representation of  $G$  is completely reducible.*

**Proof.** We prove this by induction on dimension. The base case  $n = 1$  is trivial: if  $R$  is a 1-dimensional representation then any subrepresentation has dimension 0 or 1 and is therefore either the zero representation or  $R$ , hence  $R$  is irreducible.

Suppose  $R: G \rightarrow GL(V)$  is an  $n$ -dimensional complex representation and suppose the proposition is true for all representations of dimension strictly less than  $n$ . By Proposition 7.32 we know there exists an invariant Hermitian inner product  $\langle, \rangle$  on  $V$ . If  $V$  is irreducible then we are already done. If not, then there exists a proper subrepresentation  $W \subset V$ . Take the  $\langle, \rangle$ -orthogonal complement  $W^\perp$ . Since  $G$  preserves the inner product, it preserves the subset  $W^\perp$

and therefore this is a complementary subrepresentation. Therefore  $V = W \oplus W^\perp$  and  $W$  and  $W^\perp$  have dimension lower than  $n$  so we are done by induction.  $\square$

# Lecture 8: Representations of Lie algebras

## 8.1 Basics

Let  $V$  be a  $\mathbf{K}$ -vector space. Just as a  $\mathbf{K}$ -representation of a Lie group is a homomorphism  $R: G \rightarrow GL(V)$ , a representation of a Lie algebra is a homomorphism  $\rho: \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ . Remember that a homomorphism of Lie algebras is a linear map of the underlying vector spaces which preserves the Lie bracket, i.e.

$$\rho([X, Y]) = [\rho(X), \rho(Y)].$$

In  $\mathfrak{gl}(V)$  the Lie bracket is just the commutator, so a representation satisfies

$$\rho([X, Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X).$$

**Example 8.1.** Often,  $\mathfrak{g}$  is defined as a subalgebra of  $\mathfrak{gl}(V)$ . For instance the Lie algebras  $\mathfrak{sl}(V)$ ,  $\mathfrak{su}(n)$  or  $\mathfrak{so}(n)$ . In this instance the inclusion of  $\mathfrak{g}$  into  $\mathfrak{gl}(V)$  gives a representation of  $\mathfrak{g}$  on  $V$ , called the standard representation.

**Example 8.2.** The most important example is the adjoint representation of a Lie algebra on itself, given by

$$X \mapsto \text{ad}_X, \quad \text{ad}_X Y = [X, Y].$$

It is an exercise on Sheet 3 to prove that this is a representation, in other words

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

**Example 8.3.** If  $R: G \rightarrow GL(V)$  is a representation of a matrix group  $G$  then its linearisation is a representation  $R_*$  of its Lie algebra  $\mathfrak{g}$ . This was proved in Theorem 5.12. Recall that the linearisation  $F_*$  of  $F$  satisfies  $\exp(F_*X) = F(\exp X)$  so one can compute it using

$$\left. \frac{d}{dt} \right|_{t=0} F(\exp(tX)) = F_*X.$$

Observe that if we differentiate

$$\left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{\exp(tX)} Y = \left. \frac{d}{dt} \right|_{t=0} \exp(tX) Y \exp(-tX)$$

then we get

$$XY - YX = [X, Y] = \text{ad}_X Y.$$

Hence  $\text{ad} = \text{Ad}_*$ .

## 8.2 New representations from old

Just like for Lie groups, we can form direct sums, duals, tensor products and symmetric and exterior powers of representations. If  $R: G \rightarrow GL(V)$  and  $\rho: \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  are related by

$$R(\exp(X)) = \exp(\rho(X))$$

and  $V$  is a (direct sum/dual/tensor product/symmetric power/exterior power) of  $G$ -representations then we would like  $\rho$  to be the corresponding (direct sum/dual/tensor product/symmetric power/exterior power) of  $\mathfrak{g}$ -representations. This suggests the following definitions:

### 8.2.1 Direct sum

Suppose that  $\exp(tX)$  is a one-parameter family in the Lie group  $G$  and

$$R = R_1 \oplus R_2: G \rightarrow GL(V_1 \oplus V_2)$$

is a direct sum. Then

$$R(\exp(tX))(v_1 \oplus v_2) = \exp(t\rho(X))(v_1 \oplus v_2).$$

The action of  $\rho(X)$  can therefore be found by differentiating with respect to  $t$  at  $t = 0$ :

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} R(\exp(tX))(v_1 \oplus v_2) &= \left. \frac{d}{dt} \right|_{t=0} (R_1(\exp(tX))v_1 \oplus R_2(\exp(tX))v_2) \\ &= \rho_1(X)v_1 \oplus \rho_2(X)v_2. \end{aligned}$$

**Definition 8.4.** If  $\rho_1: \mathfrak{g} \rightarrow \mathfrak{gl}(V_1)$  and  $\rho_2: \mathfrak{g} \rightarrow \mathfrak{gl}(V_2)$  are representations then

$$\rho_1 \oplus \rho_2: \mathfrak{g} \rightarrow \mathfrak{gl}(V_1 \oplus V_2)$$

defined by  $(\rho_1 \oplus \rho_2)(X)(v_1 \oplus v_2) = \rho_1(X)v_1 \oplus \rho_2(X)v_2$  is a representation called the direct sum of  $\rho_1$  and  $\rho_2$ .

### 8.2.2 Duals

We have

$$(R^*(\exp(tX))f)(v) = (\exp(\rho^*(tX))f)(v)$$

and

$$(R^*(\exp(tX))f)(v) = f(R(\exp(-tX))v)$$

which (upon differentiating at  $t = 0$ ) gives

$$(\rho^*(X)f)v = -f(\rho(X)v).$$

**Definition 8.5.** If  $\rho: \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  is a representation of  $\mathfrak{g}$  then the representation  $\rho^*: \mathfrak{g} \rightarrow \mathfrak{gl}(V^*)$  defined by

$$(\rho^*(X)f)v = -f(\rho(X)v)$$

is called the *dual representation* to  $\rho$ .

### 8.2.3 Tensor products

Here we must differentiate

$$R(\exp(tX))(v \otimes w) = R_1(\exp(tX))v \otimes R_2(\exp(tX))w$$

at  $t = 0$ , which gives

$$\rho_1(X)v \otimes w + v \otimes \rho_2(X)w$$

by Leibniz's rule.

**Definition 8.6.** If  $\rho_i: \mathfrak{g} \rightarrow \mathfrak{gl}(V_i)$  are representations of  $\mathfrak{g}$  for  $i = 1, 2$  then the representation

$$\rho_1 \otimes \rho_2: \mathfrak{g} \rightarrow \mathfrak{gl}(V_1 \otimes V_2)$$

defined by

$$(\rho_1 \otimes \rho_2)(X)(v \otimes w) = (\rho_1(X)v) \otimes w + v \otimes (\rho_2(X)w)$$

is the *tensor product* of the representations  $\rho_1$  and  $\rho_2$ .

### 8.2.4 Symmetric and exterior powers

The definition is entirely analogous to the case of group representations now that we know the definition of tensor product of representations of Lie algebras. Instead of going into detail, let's do an example.

**Example 8.7.** Consider the standard representation of  $\mathfrak{sl}(2, \mathbb{C})$  on  $\mathbb{C}^2$ . The Lie algebra  $\mathfrak{sl}(2, \mathbb{C})$  has generators  $H, X, Y$  which are the matrices

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

Let  $e_1, e_2$  be a basis of  $\mathbb{C}^2$  (the one we used to write these matrices!), in other words

$$He_1 = e_1, \quad He_2 = -e_2, \quad Xe_1 = 0 = Ye_2, \quad Xe_2 = e_1, \quad Ye_1 = e_2.$$

A basis for the symmetric square  $\text{Sym}^2 \mathbb{C}^2$  is

$$e_1 \otimes e_1, \quad \frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1), \quad e_2 \otimes e_2$$

and the action of  $H$  is

$$\begin{aligned} \text{Sym}^2(H)(e_1 \otimes e_1) &= (He_1) \otimes e_1 + e_1 \otimes (He_1) \\ &= 2e_1 \otimes e_1 \\ \text{Sym}^2(H)(e_2 \otimes e_2) &= -2e_2 \otimes e_2 \\ \text{Sym}^2(H)\frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1) &= \frac{1}{2}((He_1) \otimes e_2 + e_1 \otimes (He_2) + (He_2) \otimes e_1 + e_2 \otimes (He_1)) \\ &= 0 \end{aligned}$$

because  $He_1 = e_1$  and  $He_2 = -e_2$ .

The action of  $\text{Sym}^2(X)$  is

$$\begin{aligned}\text{Sym}^2(X)(e_1 \otimes e_1) &= 0 \\ \text{Sym}^2(X)\frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1) &= \frac{1}{2}(e_1 \otimes e_1 + e_1 \otimes e_1) \\ &= e_1 \otimes e_1 \\ \text{Sym}^2(X)(e_2 \otimes e_2) &= e_1 \otimes e_2 + e_2 \otimes e_1 \\ &= 2\frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1)\end{aligned}$$

and of  $\text{Sym}^2(Y)$  is

$$\begin{aligned}\text{Sym}^2(Y)(e_1 \otimes e_1) &= e_1 \otimes e_2 + e_2 \otimes e_1 \\ &= 2\frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1) \\ \text{Sym}^2(Y)\frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1) &= \frac{1}{2}(e_2 \otimes e_2 + e_2 \otimes e_2) \\ &= e_2 \otimes e_2 \\ \text{Sym}^2(X)(e_2 \otimes e_2) &= 0.\end{aligned}$$

You should do the symmetric cube  $\text{Sym}^3$  just to make sure you understand.

**Example 8.8.** The exterior square of the standard representation of  $\mathfrak{sl}(2, \mathbb{C})$  is spanned by

$$e_1 \wedge e_2 = \frac{1}{2}(e_1 \otimes e_2 - e_2 \otimes e_1).$$

We see that

$$H(e_1 \wedge e_2) = (He_1) \wedge e_2 + e_1 \wedge (He_2) = 0$$

and

$$X(e_1 \wedge e_2) = (Xe_1) \wedge e_2 + e_1 \wedge (Xe_2) = e_1 \wedge e_1 = 0$$

and similarly for  $Y$ .

### 8.3 Complexification

Suppose that  $\mathfrak{g}$  is a Lie algebra over  $\mathbb{R}$ . In particular it is a real vector space, so we can *complexify* it:

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes \mathbb{C}$$

by considering the vector space  $\mathfrak{g} \oplus \mathfrak{g}$  as a vector space over  $\mathbb{C}$  where we think of  $(v, w)$  as  $v + iw$  and let  $\mathbb{C}$  act accordingly. The Lie bracket extends in a complex linear way:

$$[v_1 + iw_1, v_2 + iw_2] = [v_1, v_2] - [w_1, w_2] + i([v_1, w_2] + [w_1, v_2]).$$

**Example 8.9.** We have

$$\mathfrak{u}(n)_{\mathbf{C}} = \mathfrak{gl}(n, \mathbf{C}), \quad \mathfrak{su}(n)_{\mathbf{C}} = \mathfrak{sl}(n, \mathbf{C}).$$

This is because  $i$  times a skew-Hermitian matrix is Hermitian and any matrix  $A$  can be written as a sum

$$\frac{1}{2} (A + A^\dagger) + \frac{1}{2} (A - A^\dagger)$$

of its Hermitian and skew-Hermitian parts. More obviously, we have

$$\mathfrak{gl}(n, \mathbf{R})_{\mathbf{C}} = \mathfrak{gl}(n, \mathbf{C}), \quad \mathfrak{sl}(n, \mathbf{R})_{\mathbf{C}} = \mathfrak{sl}(n, \mathbf{C}).$$

This shows that different Lie algebras can have the same complexification. We say that  $\mathfrak{sl}(n, \mathbf{R})$  and  $\mathfrak{su}(n)$  are different real forms of the complex Lie algebra  $\mathfrak{sl}(n, \mathbf{C})$ . We will see below (Lemma 8.10) that the complex representations of a Lie algebra are determined by the complex representations of its complexification - this means that different real forms have the same complex representation theory. This will prove a useful point of view in the following chapters.

Suppose  $\mathfrak{g}$  is a real Lie algebra. A complex representation of  $\mathfrak{g}$  is a real-linear Lie algebra homomorphism  $\mathfrak{g} \rightarrow \mathfrak{gl}(n, \mathbf{C})$ . A complex linear complex representation of  $\mathfrak{g}_{\mathbf{C}}$  is a complex-linear Lie algebra homomorphism  $\mathfrak{g}_{\mathbf{C}} \rightarrow \mathfrak{gl}(n, \mathbf{C})$ .

**Lemma 8.10.** *There is a one-to-one correspondence between complex representations of  $\mathfrak{g}$  and complex linear complex representations of  $\mathfrak{g}_{\mathbf{C}}$ .*

**Proof.** If  $\rho: \mathfrak{g} \rightarrow \mathfrak{gl}(n, \mathbf{C})$  is a real-linear Lie algebra homomorphism, we can define  $\rho_{\mathbf{C}}: \mathfrak{g}_{\mathbf{C}} \rightarrow \mathfrak{gl}(n, \mathbf{C})$  by

$$\rho_{\mathbf{C}}(X + iY) = \rho(X) + i\rho(Y).$$

Conversely, if  $\sigma: \mathfrak{g}_{\mathbf{C}} \rightarrow \mathfrak{gl}(n, \mathbf{C})$  is a complex linear map then  $\sigma(X + iY) = \sigma(X) + i\sigma(Y)$  so setting  $\rho = \sigma|_{\mathfrak{g}}$  we get  $\rho_{\mathbf{C}} = \sigma$ .  $\square$

For example, this means that the complex representations of  $\mathfrak{su}(n)$  and the complex representations of  $\mathfrak{sl}(n, \mathbf{R})$  are in one-to-one correspondence with one another and also with the complex linear complex representations of  $\mathfrak{sl}(n, \mathbf{C}) \cong \mathfrak{su}(n)_{\mathbf{C}} \cong \mathfrak{sl}(n, \mathbf{R})_{\mathbf{C}}$ .



# Lecture 9: Representations of tori

**Definition 9.1.** The  $n$ -dimensional torus is the group  $U(1)^n$ .

## 9.1 Representations of $U(1)$

Define the homomorphism  $F_n: U(1) \rightarrow U(1)$  by  $F_n(e^{i\theta}) = e^{in\theta}$ . Recall that we showed (Example 6.1) that the only smooth homomorphisms  $U(1) \rightarrow U(1)$  are  $F_n$ . We call  $n$  the weight of the representation.

**Lemma 9.2** (Schur's Lemma). *If  $R: U(1) \rightarrow GL(V)$  is an irreducible finite-dimensional complex representation and  $L \in GL(V)$  is a matrix which commutes with every element in  $R(U(1))$  then  $L$  is just  $\lambda$  times the identity matrix for some  $\lambda \in \mathbb{C}$ .*

**Proof.** Let  $\lambda$  be an eigenvalue of  $L$ . The (nonzero) eigenspace  $E_\lambda = \{v \in V : Lv = \lambda v\}$  is a subrepresentation of  $R$ : to see this, if  $v \in E_\lambda$  then

$$L(R(g)v) = R(g)Lv = R(g)\lambda v = \lambda R(g)v$$

so  $R(g)v \in E_\lambda$ . Therefore  $E_\lambda = V$  by irreducibility of  $V$ , so  $Lv = \lambda v$  for all  $v \in V$ .  $\square$

**Corollary 9.3.** *If  $R: U(1) \rightarrow GL(V)$  is an irreducible complex representation of  $U(1)$  then  $V$  is one-dimensional.*

**Proof.** Consider  $R(e^{i\theta})$  and  $R(e^{i\phi})$ . As  $U(1)$  is abelian and  $R$  is a homomorphism, these commute for all  $\theta, \phi$ . Therefore by Schur's lemma,  $R(e^{i\theta})v = \lambda v$  for some  $\lambda \in \mathbb{C}^*$ . If  $v \in V \setminus \{0\}$  then the subspace  $\mathbb{C}v$  is a one-dimensional subrepresentation and hence, since  $V$  is irreducible,  $V = \mathbb{C}v$ .  $\square$

**Lemma 9.4.** *Let  $R: U(1) \rightarrow GL(1, \mathbb{C})$  be a one-dimensional complex representation of  $U(1)$ . Then  $R(e^{i\theta}) = e^{in\theta}$  for some  $n \in \mathbb{Z}$ .*

**Proof.** Let  $\langle, \rangle$  be an invariant Hermitian form on  $\mathbb{C}$  for the representation  $R$ . The group of  $g \in GL(1, \mathbb{C})$  preserving this form is  $U(1)$ :  $\langle v, w \rangle = \langle gv, gw \rangle = g^\dagger g \langle v, w \rangle$  implies  $g^\dagger = g^{-1}$ . Therefore  $R$  defines a homomorphism  $U(1) \rightarrow U(1)$ , which must be of the form  $F_n$  for some  $n$ .  $\square$

The set of isomorphism classes of irreducible  $U(1)$ -representations is therefore in bijection with the integers ( $\circ$  denotes 0):

...   •   •   •   ◦   •   •   •   ...

Since  $U(1)$  is compact, any finite-dimensional representation is completely reducible and so we can write it as a direct sum of irreducible representations. Let  $R: G \rightarrow GL(V)$  be a representation and suppose that the representation  $F_n$  occurs  $m_n$  times as a summand in  $R$ . We call this the *multiplicity* of  $F_n$  in  $R$  and we can encode  $R$  by writing the number  $m_n$  on the  $n$ th dot;  $R$  can then be reconstructed as  $\bigoplus_{n \in \mathbb{Z}} F_n^{\oplus m_n}$ . For example

...   •<sup>1</sup>   •   •<sup>2</sup>   ◦<sup>6</sup>   •   •<sup>9</sup>   •<sup>1</sup>   ...

corresponds to  $F_{-3} \oplus F_{-1}^{\oplus 2} \oplus F_0^{\oplus 6} \oplus F_2^{\oplus 9} \oplus F_3$

The vector space  $V$  underlying the representation breaks up, correspondingly, into subrepresentations

$$V = \bigoplus_{n \in \mathbf{Z}} V_n, \quad V_n = \{v \in V : R(e^{i\phi})v = e^{in\phi}v\}$$

of dimension  $m_n$ . These are called the weight spaces of the representation. One can also think of them as the  $e^{in\phi}$ -eigenspaces of  $R(e^{i\phi})$ .

## 9.2 Representations of tori

For an  $n$ -tuple of integers  $(k_1, \dots, k_n)$ , consider the representation  $F_{(k_1, \dots, k_n)}: U(1)^n \rightarrow GL(1, \mathbf{C})$  given by  $F_{(k_1, \dots, k_n)}(e^{i\theta_1}, \dots, e^{i\theta_n}) = (e^{ik_1\theta_1}, \dots, e^{ik_n\theta_n})$ .

**Lemma 9.5.** *Any irreducible representation of  $U(1)^n$  is isomorphic to  $F_{(k_1, \dots, k_n)}$  for some  $n$ -tuple  $(k_1, \dots, k_n)$ .*

**Proof.** If  $V$  is an irreducible representation of  $U(1)^n$  then we can consider it as a representation  $R|_{U(1)_k}$  of  $U(1)_k = \{(1, \dots, 1)\} \times U(1) \times \{(1, \dots, 1)\}$  (the  $U(1)$  appearing in the  $k$ th factor). Let  $V = \bigoplus_{j \in \mathbf{Z}} V_j$  denote the weight space decomposition of  $R|_{U(1)_1}$ , i.e. if  $v \in V_j$  and  $h = (e^{i\phi}, 1, \dots, 1) \in U(1)_1$  then  $R(h)v = e^{ij\phi}v$ .

Since  $U(1)^n$  is abelian, elements of  $U(1)_1$  and  $U(1)_k$  commute. Therefore if  $v \in V_j$ ,  $g \in U(1)_k$  and  $h = (e^{i\phi}, 1, \dots, 1) \in U(1)_1$  then we have  $R(h)R(g)v = R(g)R(h)v = R(g)e^{ij\phi}v = e^{ij\phi}R(g)v$  so  $R(g)v \in V_j$ . This means that  $V_j$  is a subrepresentation of  $V$ ; in particular, since  $V$  is irreducible, there is only one nonzero summand in the  $U(1)_1$ -weight space decomposition of  $V$ , say  $j = k_1$ .

Now decompose  $V = V_{k_1}$  into weight spaces of  $R|_{U(1)_2}$ . We have

$$V_{k_1} = \bigoplus_{j \in \mathbf{Z}} V_{k_1, j}$$

where  $V_{k_1, j}$  is the  $e^{ij\phi}$ -eigenspace of  $R(1, e^{i\phi}, 1, \dots, 1)$ . By the same argument, each  $V_{k_1, j}$  is a subrepresentation, so by irreducibility,  $j$  can only take on one value, say  $k_2$ . Continuing in this manner, we see that  $V = V_{k_1, \dots, k_n}$  is simultaneously an eigenspace for

$$(e^{i\phi}, 1, \dots, 1), (1, e^{i\phi}, 1, \dots, 1), \dots, (1, \dots, 1, e^{i\phi}),$$

with eigenvalues  $e^{ik_1\phi}, e^{ik_2\phi}, \dots, e^{ik_n\phi}$  respectively. Any vector  $v$  in this space is just rescaled by  $e^{i(k_1\phi_1 + \dots + k_n\phi_n)}$  and hence spans an invariant subspace (a subrepresentation). This means that  $V$  is one-dimensional and isomorphic to  $F_{k_1, \dots, k_n}$ .  $\square$

The set of isomorphism classes of irreducible  $U(1)^n$ -representations is therefore in bijection with the lattice  $\mathbf{Z}^n$ . Since  $U(1)^n$  is compact, any finite-dimensional representation is completely reducible and so we can write it as a direct sum of irreducible representations. Let  $R: G \rightarrow GL(V)$  be a representation and suppose that the representation  $F_{k_1, \dots, k_n}$  occurs  $m_{k_1, \dots, k_n}$  times as a summand in  $R$ . We call this the *multiplicity* of  $F_{k_1, \dots, k_n}$  in  $R$  and we can encode  $R$  by writing the number  $m_{k_1, \dots, k_n}$  on the  $(k_1, \dots, k_n)$  lattice point;  $R$  can then be reconstructed as  $\bigoplus_{(k_1, \dots, k_n) \in \mathbf{Z}^n} F_{k_1, \dots, k_n}^{\oplus m_{k_1, \dots, k_n}}$ .

The vector space  $V$  underlying the representation breaks up, correspondingly, into subrepresentations

$$V = \bigoplus_{(k_1, \dots, k_n) \in \mathbf{Z}^n} V_{k_1, \dots, k_n}, \quad V_{k_1, \dots, k_n} = \{v \in V : R(e^{i\phi_1}, \dots, e^{i\phi_n})v = e^{i(k_1\phi_1 + \dots + k_n\phi_n)}v\}$$

of dimension  $m_{k_1, \dots, k_n}$ . These are called the weight spaces of the representation. One can also think of them as the simultaneous  $e^{ik_j \phi}$ -eigenspaces of  $R(U(1)_j)$ .

### 9.3 Lattice of weights

Our notation is clearly out of control. Let  $\mathfrak{t}$  denote the (complexified, for simplicity) Lie algebra of  $U(1)^n$  and  $\mathfrak{t}^*$  denote its dual space; an element  $\lambda \in \mathfrak{t}^*$  is a linear map sending Lie algebra elements in  $\mathfrak{t}$  to complex numbers. Consider the vectors  $u_k = (0, \dots, 0, i, 0, \dots, 0)$ , with the  $i$  in the  $k$ th place. We have  $\exp(u_k t) = (1, \dots, 1, e^{it}, 1, \dots, 1)$ . Given a complex representation  $R: U(1)^n \rightarrow GL(V)$ , the weight spaces are the subspaces

$$V_{k_1, \dots, k_n} = \{v \in V : R(\exp(u_1 t_1 + \dots + u_n t_n))v = \exp(ik_1 t_1 + \dots + ik_n t_n)v\}.$$

Equivalently, if  $\rho$  is the corresponding Lie algebra representation (satisfying  $R(\exp(X)) = \exp(\rho(X))$ ) we see that  $\rho(u_1 t_1 + \dots + u_n t_n) = ik_1 t_1 + \dots + ik_n t_n$ . Clearly the weights  $(k_1, \dots, k_n)$  can be encoded as an element  $\lambda \in \mathfrak{t}^*$ , namely the linear map

$$\lambda(u_1 t_1 + \dots + u_n t_n) = ik_1 t_1 + \dots + ik_n t_n.$$

How do we remember the fact that the  $k_j$ s are integers in this new notation? The points  $\sum u_i t_i \in \mathfrak{t}$  with  $t_i \in 2\pi\mathbb{Z}$  are sent to the identity by  $\exp$ . Under  $\lambda$  they are sent to elements of  $2\pi i\mathbb{Z}$ . If we write  $\ker \exp = \{X \in \mathfrak{t} : \exp(X) = 1\}$  then we define the *weight lattice*:

$$\mathfrak{t}_{\mathbb{Z}}^* = \{\lambda \in \mathfrak{t}^* : \lambda(X) \in 2\pi i\mathbb{Z} \forall X \in \ker \exp\}.$$

Our weight space decomposition can now be written much more succinctly as

$$V = \bigoplus_{\lambda \in \mathfrak{t}_{\mathbb{Z}}^*} V_{\lambda}, \quad V_{\lambda} = \{v \in V : R(\exp(X))v = \exp(\lambda(X))v, \forall X \in \mathfrak{t}\}$$

or equivalently

$$V_{\lambda} = \{v \in V : \rho(X)v = \lambda(X)v \forall X \in \mathfrak{t}\}.$$

Usually we will only write  $V = \bigoplus_{\alpha \in A} V_{\alpha}$  where  $A \subset \mathfrak{t}_{\mathbb{Z}}^*$  is the finite subset of weights with nonzero multiplicities in the representation.

### 9.4 Tensor products

To find the weights of a tensor product in general one can use the following lemma.

**Lemma 9.6.** *Let  $T$  be a torus with Lie algebra  $\mathfrak{t}$ . If  $R: T \rightarrow GL(V)$  and  $S: T \rightarrow GL(W)$  are two representations of  $T$  with weight space decompositions  $V = \bigoplus_{\alpha \in A} V_{\alpha}$ ,  $W = \bigoplus_{\beta \in B} W_{\beta}$  for some subsets  $A \subset \mathfrak{t}_{\mathbb{Z}}^*$ ,  $B \subset \mathfrak{t}_{\mathbb{Z}}^*$  then the tensor product representation of  $T$  on  $V \otimes W$  has weight space decomposition  $\bigoplus (V \otimes W)_{\gamma}$  with*

$$(V \otimes W)_{\gamma} = \bigoplus_{\alpha \in A, \beta \in B, \gamma = \alpha + \beta} V_{\alpha} \otimes W_{\beta}.$$

**Proof.** We have

$$V \otimes W = \left( \bigoplus_{\alpha \in A} V_{\alpha} \right) \otimes \left( \bigoplus_{\beta \in B} W_{\beta} \right) = \bigoplus_{(\alpha, \beta) \in A \times B} V_{\alpha} \otimes W_{\beta}.$$

We will show that  $V_{\alpha} \otimes W_{\beta} \subset (V \otimes W)_{\alpha+\beta}$ , which will prove the lemma.

$$\begin{aligned} v \in V_{\alpha} &\Rightarrow R(e^{t\theta})v = e^{\alpha(t)}v \quad \text{for all } t \in \mathfrak{t} \\ w \in W_{\beta} &\Rightarrow S(e^{t\theta})w = e^{\beta(t)}w \quad \text{for all } t \in \mathfrak{t} \\ (R \otimes S)(v \otimes w) &= R(v) \otimes S(w) \\ &= e^{\alpha(t)}e^{\beta(t)}v \otimes w \\ &= e^{(\alpha+\beta)(t)}v \otimes w. \end{aligned}$$

so  $V_{\alpha} \otimes W_{\beta} \subset (V \otimes W)_{\alpha+\beta}$ . □

## Lecture 10: Representations: $SU(2)$ and $SO(3)$

### 10.1 Representations of $SU(2)$

#### 10.1.1 $\mathfrak{su}(2)$ and $\mathfrak{sl}(2, \mathbb{C})$

The Lie algebra  $\mathfrak{su}(2)$  consists of tracefree skew-Hermitian complex two-by-two matrices. A basis is:

$$\sigma_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$

Remember that if we think of  $\sigma_1, \sigma_2, \sigma_3$  as an oriented basis of  $\mathbb{R}^3$  then the Lie bracket is just twice the vector cross-product.

*Remark 10.1.* Note that the complexification of  $\mathfrak{su}(2)$  is  $\mathfrak{sl}(2, \mathbb{C})$ , the Lie algebra of tracefree complex two-by-two matrices. Remember (Lemma 8.10) that complex representations of a Lie algebra and of its complexification are in one-to-one correspondence, so we will switch between the two pictures when convenient. It will be convenient to consider the following basis of  $\mathfrak{sl}(2, \mathbb{C})$ :

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

These basis elements have Lie brackets:

$$[H, X] = 2X, \quad [H, Y] = -2Y, \quad [X, Y] = H.$$

Note that

$$H = -i\sigma_1, \quad X = \frac{1}{2}(\sigma_2 - i\sigma_3), \quad Y = -\frac{1}{2}(\sigma_2 + i\sigma_3).$$

#### 10.1.2 Diagonalising $\sigma_1$ and $H$

Suppose that  $\rho: \mathfrak{su}(2) \rightarrow \mathfrak{gl}(V)$  is a finite-dimensional complex representation of  $\mathfrak{su}(2)$ .

**Lemma 10.2.**  $V$  decomposes as a direct sum  $\bigoplus_{\lambda} V_{\lambda}$  where

$$V_{\lambda} = \{v \in V : \rho(\sigma_1)v = \lambda v\}.$$

The collection of eigenvalues  $\lambda \in \mathbb{C}$  which occur in the decomposition of a given representation  $V$  are called the weights of  $V$  and the corresponding eigenspaces are called weight spaces. In fact, the weights are imaginary integers.

**Proof.** By Lie's theorem, exponentiating  $\rho$  gives a representation  $R$  of  $SU(2)$  such that

$$R(\exp(t\sigma_1))v = \exp(\rho(t\sigma_1))v.$$

Since the subgroup  $\{\exp(t\sigma_1) : t \in [0, 2\pi)\} \subset SU(2)$  is isomorphic to  $U(1)$  we obtain a representation of  $U(1)$  by restricting to this subgroup. Therefore  $V$  (by complete reducibility and the classification of  $U(1)$ -representations) splits as a direct sum of  $U(1)$ -representations

$$V = \bigoplus_{m \in M} V_m$$

for some subset  $M \subset \mathbf{Z}$ , where  $\exp(t\sigma_1)$  acts as scalar multiplication by  $e^{imt}$  on  $V_m$ .

Since  $R(\exp t\sigma_1)v = \exp(\rho(t\sigma_1))v$ , this means that  $\exp(\rho(t\sigma_1)) = e^{imt}$ , so

$$\rho(\sigma_1)v = imv.$$

□

Note that a complex representation  $\rho$  of  $\mathfrak{su}(2)$  extends (complex linearly) to a complex representation  $\rho$  of  $\mathfrak{sl}(2, \mathbf{C})$ . Since  $H = -i\sigma_1$ , we know that in this complex representation of  $\mathfrak{sl}(2, \mathbf{C})$ ,

$$\rho(H)v = mv$$

for some integer  $m$ . In fact, any complex representation of  $\mathfrak{sl}(2, \mathbf{C})$  restricts to a complex representation to  $\mathfrak{su}(2)$ . By the above argument, this representation splits into eigenspaces of  $\rho(\sigma_1)$  with imaginary integer eigenvalues, so we obtain:

**Corollary 10.3.** *If  $\rho: \mathfrak{sl}(2, \mathbf{C}) \rightarrow \mathfrak{gl}(V)$  is a finite-dimensional complex representation then  $V$  admits a weight space decomposition*

$$V = \bigoplus_{\lambda} V_{\lambda}$$

where the weights  $\lambda$  are integers and

$$V_{\lambda} = \{v \in V : \rho(H)v = \lambda v\}.$$

Notice that we cannot directly apply the argument that proved Lemma 10.2 to obtain this corollary! This is because  $SL(2, \mathbf{C})$  is not compact and  $\{\exp(tH) : t \in \mathbf{R}\}$  is actually a subgroup isomorphic to  $\mathbf{R}$ , not to  $U(1)$ .

### 10.1.3 The action of $X$ and $Y$

We now consider the actions of  $\rho(X)$  and  $\rho(Y)$ . The key bracket relations to remember are:

$$[H, X] = 2X, \quad [H, Y] = -2Y, \quad [X, Y] = H.$$

These imply:

**Lemma 10.4.** *If  $\rho: \mathfrak{sl}(2, \mathbf{C}) \rightarrow \mathfrak{gl}(V)$  is a finite-dimensional complex representation of  $\mathfrak{sl}(2, \mathbf{C})$  and  $V = \bigoplus_{\lambda \in \Lambda} V_{\lambda}$  is the weight space decomposition,*

$$V_{\lambda} = \{v \in V : \rho(H)v = \lambda v\}$$

for some collection  $\Lambda \subset \mathbf{Z}$  of weights then:

- $v \in V_\lambda$  implies  $\rho(X)v \in V_{\lambda+2}$
- $v \in V_\lambda$  implies  $\rho(Y)v \in V_{\lambda-2}$

**Proof.** For ease of notation I will omit the  $\rho$  and simply write the action of  $X$  on  $v$  as  $Xv$ .

If  $v \in V_\lambda$  then  $Hv = \lambda v$ , so we have

$$\begin{aligned} HXv &= [H, X]v + XHv \\ &= 2Xv + X\lambda v \\ &= (\lambda + 2)Xv \end{aligned}$$

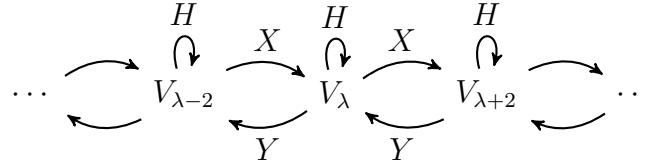
so  $Xv \in V_{\lambda+2}$ .

Similarly

$$\begin{aligned} HYv &= [H, Y]v + YHv \\ &= -2Yv + Y\lambda v \\ &= (\lambda - 2)Yv \end{aligned}$$

so  $Yv \in V_{\lambda-2}$ . □

We can visualise the representation as a chain of weight spaces where  $H$  preserves each weight space,  $X$  shifts to the right (increasing the weight) and  $Y$  shifts to the left (decreasing the weight):



### 10.1.4 Classification of irreducible representations

With the weight space decomposition and the action of  $X$  and  $Y$  in hand, we can now classify all irreducible representations of  $\mathfrak{sl}(2, \mathbb{C})$  (equivalently  $\mathfrak{su}(2)$  or  $SU(2)$ ).

**Theorem 10.5.** *Suppose that  $\rho: \mathfrak{sl}(2, \mathbb{C}) \rightarrow \mathfrak{gl}(V)$  is an irreducible finite-dimensional complex representation. Then the weight spaces  $V_\lambda$  are one-dimensional and the weights live in an uninterrupted chain*

$$-m, -m + 2, \dots, m - 2, m$$

from  $-m$  to  $m$  for some  $m \in \mathbb{Z}$ .

**Proof.** Suppose that  $m$  is the biggest integer for which  $V_m \neq 0$  (this exists because the representation is finite-dimensional). Since  $XV_m \subset V_{m+2}$ , we know that  $Xv = 0$  for all  $v \in V_m$ . We call such a  $v$  a *highest-weight vector*. Pick a highest-weight vector  $v$ . The theorem follows if we can show that  $v, Yv, Y^2v, \dots, Y^m v$  is a basis for the representation.

Consider the sequence  $v, Yv, Y^2v, \dots, Y^k v$  where  $k$  is maximal such that  $Y^{k+1}v = 0$ . The subspace  $W$  spanned by this sequence is clearly preserved by the action of  $Y$  and by the action of  $H$  (because  $Y^n v$  is an  $m - 2n$ -eigenvector of  $H$ ).

The subspace  $W$  is also preserved by the action of  $X$ : indeed I claim that

$$XY^n v = B_n Y^{n-1} v$$

for some numbers  $B_n$ . To see this, note that

$$XYv = [X, Y]v + YXv = Hv + 0 = mv$$

so  $B_1 = m$ , and

$$\begin{aligned} XY^n v &= [X, Y]Y^{n-1}v + YXY^{n-1}v \\ &= HY^{n-1}v + YB_{n-1}Y^{n-2}v \\ &= (m - 2n + 2)Y^{n-1}v + B_{n-1}Y^{n-1}v \end{aligned}$$

We see that

$$B_n = B_{n-1} + m - 2n + 2$$

which (find the  $n$ th term!) gives

$$XY^n v = (m - n + 1)nY^{n-1}v.$$

We can deduce more from this formula:

- We deduce that  $Y^n v \neq 0$  for all  $0 \leq n \leq m$ . This follows by induction: certainly  $Y^0 v = v \neq 0$  and

$$XY^{n+1}v = (m - n)(n + 1)Y^n v$$

which is nonzero by induction for  $n < m$ .

- We deduce that  $Y^{m+1}v = 0$ . To see this, let  $n$  be the smallest integer such that  $Y^n v = 0$  (which exists because the representation is finite-dimensional). We have

$$0 = XY^n v = (m - n + 1)nY^{n-1}v$$

which is nonzero unless  $m - n + 1 = 0$ . This implies  $m + 1 = n$ .

- These two facts imply that the weights in the weight space decomposition of  $W$  occur in an uninterrupted chain  $m, m - 2, \dots, 2 - m, -m$ . Moreover it is clear that the weight spaces of  $W$  are one-dimensional (spanned by vectors of the form  $Y^n v$ ).

It remains to show that  $W = V$ . Since  $W$  is preserved by the generators of the Lie algebra, it is preserved by any linear combination of them and therefore it is a subrepresentation. By irreducibility of  $V$ , we have  $V = W$ .  $\square$

**Theorem 10.6.** *The (isomorphism classes of) finite-dimensional irreducible complex representations of  $\mathfrak{sl}(2, \mathbb{C})$  (equivalently of  $\mathfrak{su}(2)$  or of  $SU(2)$ ) are in bijection with the nonnegative integers. The bijection sends an irreducible representation to its highest weight. The highest weight divided by two is a number in  $\frac{1}{2}\mathbb{Z}$  sometimes called the spin of the representation.*

We have not yet proved the theorem: it remains to construct representations with each nonnegative integer value of spin. We proceed with a couple of examples, then construct the remaining representations by taking symmetric products (Proposition 10.7).



### 10.1.5 Example: the adjoint representation

Consider the adjoint representation  $V = \mathfrak{sl}(2, \mathbb{C})$  of  $\mathfrak{sl}(2, \mathbb{C})$ . This is three-dimensional, spanned by  $H, X, Y$ . Since

$$\mathrm{ad}_H X = [H, X] = 2X, \quad \mathrm{ad}_H Y = [H, Y] = -2Y, \quad \mathrm{ad}_H H = [H, H] = 0$$

we see that the weight spaces are

$$V_{-2} = \mathbb{C}Y, \quad V_0 = \mathbb{C}H, \quad V_2 = \mathbb{C}X$$

so  $X$  is a highest-weight vector (with weight two). Indeed

$$\mathrm{ad}_X Y = [X, Y] = H, \quad \mathrm{ad}_X H = [X, H] = -2X, \quad \mathrm{ad}_X X = [X, X] = 0$$

so we can see that  $X$  shifts weight spaces to the right and

$$\mathrm{ad}_Y Y = [Y, Y] = 0, \quad \mathrm{ad}_Y H = [Y, H] = 2Y, \quad \mathrm{ad}_Y X = [Y, X] = -H$$

so we can see that  $Y$  shifts weight spaces to the left.

### 10.1.6 Example: The standard representation

Consider  $\mathbb{C}^2$  as the standard representation of  $\mathfrak{sl}(2, \mathbb{C})$ , where  $H, X$  and  $Y$  act as the matrices

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

The weight spaces are spanned by

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

with weights  $\pm 1$ .

### 10.1.7 Symmetric powers

**Proposition 10.7.** *The  $n$ th symmetric power  $\mathrm{Sym}^n(\mathbb{C}^2)$  of the standard representation of  $SU(2)$  is irreducible with highest weight  $n$ .*

**Proof.** Let  $e_1$  and  $e_2$  be eigenvectors of  $H$  in the  $\pm 1$  weight-spaces of  $\mathbb{C}^2$ . Then  $e_1^{\otimes n}$  has weight  $n$  and is therefore contained in an irreducible subrepresentation  $W \subset \mathrm{Sym}^n \mathbb{C}^2$  of highest weight at least  $n$ . This subrepresentation has dimension at least  $n + 1$  (by the proof of the classification theorem for irreducible representations) as it contains the nonzero vectors

$$e_1^{\otimes n}, Y e_1^{\otimes n}, Y^2 e_1^{\otimes n}, \dots, Y^n e_1^{\otimes n}.$$

The dimension of  $\mathrm{Sym}^n \mathbb{C}^2$  is the dimension of the space of degree  $n$  homogeneous polynomials in two variables. This is spanned by the  $n + 1$  monomials

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

and hence  $\dim_{\mathbb{C}} \mathrm{Sym}^n V = n + 1 \leq \dim_{\mathbb{C}} W$ , so  $W = \mathrm{Sym}^n V$ . □

In particular, we see that the adjoint representation is isomorphic to the symmetric square of the standard representation:

$$\mathfrak{sl}(2, \mathbb{C}) \cong \text{Sym}^2 \mathbb{C}^2.$$

We have now constructed representations of  $SU(2)$  with every possible highest weight.

## 10.2 Decomposing tensor products

We have seen that tensor products of representations are essentially never irreducible. We know they are completely reducible, so one is led to ask: how do they decompose into irreducible summands? If  $V$  and  $W$  are two representations with irreducible decompositions

$$V_1 \oplus \cdots \oplus V_m, \quad W_1 \oplus \cdots \oplus W_m$$

then  $V \otimes W$  decomposes as

$$\bigoplus V_i \otimes W_j$$

so to decompose  $V \otimes W$  completely it suffices to consider the case when  $V$  and  $W$  are irreducible. In other words, we need to understand the decomposition of

$$\text{Sym}^m \mathbb{C}^2 \otimes \text{Sym}^n \mathbb{C}^2$$

into irreducibles.

We will work out an example which indicates the general pattern.

**Example 10.8.** We have

$$\text{Sym}^2 \mathbb{C}^2 \otimes \text{Sym}^3 \mathbb{C}^2 = \text{Sym}^5 \mathbb{C}^2 \oplus \text{Sym}^3 \mathbb{C}^2 \oplus \text{Sym}^1 \mathbb{C}^2.$$

To see this, let  $v$  and  $w$  be highest weight vectors for  $\text{Sym}^2 \mathbb{C}^2$  and  $\text{Sym}^3 \mathbb{C}^2$ , so

$$Hv = 2v, \quad Hw = 3w.$$

Remember that  $\text{Sym}^2$  is spanned by  $v, Yv, Y^2v$  and  $\text{Sym}^3$  is spanned by  $w, Yw, Y^2w, Y^3w$ . The tensor product is spanned by vectors

$$Y^k v \otimes Y^\ell w$$

which have weight  $2+3-2k-2\ell = 5-2(k+\ell)$ . Therefore the weight space decomposition of the tensor product is

$$Z_{-5} \oplus Z_{-3} \oplus Z_{-1} \oplus Z_1 \oplus Z_3 \oplus Z_5$$

where  $\dim_{\mathbb{C}} Z_{\pm 5} = 1$ ,  $\dim_{\mathbb{C}} Z_{\pm 3} = 2$ ,  $\dim_{\mathbb{C}} Z_{\pm 1} = 3$ .

The vector  $v \otimes w$  generates an irreducible subrepresentation

$$\xi_5 = \{u = A(v \otimes w) : A \in \mathfrak{sl}(2, \mathbb{C})\} \subset \text{Sym}^2 \mathbb{C}^2 \otimes \text{Sym}^3 \mathbb{C}^2$$

of highest weight 5. The weight-space decomposition of this is

$$\xi_{5,-5} \oplus \xi_{5,-3} \oplus \xi_{5,-1} \oplus \xi_{5,1} \oplus \xi_{5,3} \oplus \xi_{5,5}$$

where all weight-spaces are one-dimensional.

The complement  $\zeta$  of  $\xi_5 \subset V \otimes W$  has weights  $-3, -1, 1, 3$  and the corresponding weight spaces have dimensions  $1, 2, 2, 1$ . Picking a vector in  $\zeta$  of weight 3 generates an irreducible subrepresentation  $\xi_3$  in  $\zeta$  of highest weight 3. Similarly the complement  $\zeta'$  of  $\xi_3 \subset \zeta$  has weight-space decomposition with weights  $-1, 1$  and each weight space one-dimensional. A vector of weight one in  $\zeta'$  now generates an irreducible subrepresentation  $\xi_1$  with highest weight one and its complement in  $\zeta'$  is zero.

This implies the decomposition  $\xi_5 \oplus \xi_3 \oplus \xi_1$  as claimed.

The general result is proved in a similar way.

**Theorem 10.9** (Clebsch-Gordan theorem). *The tensor product*

$$\text{Sym}^m \mathbf{C}^2 \otimes \text{Sym}^n \mathbf{C}^2$$

*decomposes into irreducible subrepresentations*

$$\bigoplus_{k=|m-n|}^{m+n} \text{Sym}^k \mathbf{C}^2$$

where the sum is over  $k \equiv m+n \pmod{2}$ .

This theorem is useful for computing the possible angular momenta of composite systems in quantum mechanics. In the next chapter we will see it play a role in the study of energy levels in the hydrogen atom.

### 10.3 Binary quadratic forms

We now analyse the problem of classifying binary quadratic forms from lecture 1 using the representation theory we have learned. A matrix  $g \in SL(2, \mathbf{C})$  acts by  $M \mapsto g^T M g$  on the space  $V$  of symmetric matrices  $M = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$  (this is how coordinate changes act on quadratic forms). This representation is 3-dimensional. The diagonal matrix  $\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$  acts by

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}^T \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} = \begin{pmatrix} ae^{2i\theta} & b/2 \\ b/2 & ce^{-2i\theta} \end{pmatrix}$$

so the weight space decomposition of this representation is  $V_{-2} \oplus V_0 \oplus V_2$  where each summand is one-dimensional. By the classification theorem, this means that it is isomorphic to the adjoint representation!

We are trying to find polynomials in  $a, b, c$  which are invariant under this action. In other words, we consider the matrix entries  $a, b, c$  as linear coordinate functions on  $V$ , that is as elements of the dual representation  $V^*$ . Homogeneous polynomials of degree  $d$  in  $a, b, c$  are now elements of  $\text{Sym}^d(V^*)$ . An invariant polynomial is one which is fixed by the action of  $SL(2, \mathbb{C})$ , in other words it spans a trivial subrepresentation of  $\text{Sym}^d(V^*)$ . So we should be able to detect invariant polynomials by decomposing  $\text{Sym}^d(V^*)$  into irreducible subrepresentations.

**Example 10.10** ( $d = 2$ ).  $V^*$  has weight spaces  $V_{-2}^* \oplus V_0^* \oplus V_2^*$  spanned respectively by  $c, b, a$ . The weight spaces for  $\text{Sym}^2(V^*)$  are

$$\begin{aligned}\text{Sym}^2(V^*)_{-4} &= \mathbb{C}\langle c^2 \rangle \\ \text{Sym}^2(V^*)_{-2} &= \mathbb{C}\langle bc \rangle \\ \text{Sym}^2(V^*)_0 &= \mathbb{C}\langle b^2, ac \rangle \\ \text{Sym}^2(V^*)_2 &= \mathbb{C}\langle ab \rangle \\ \text{Sym}^2(V^*)_4 &= \mathbb{C}\langle a^2 \rangle.\end{aligned}$$

In particular,  $a^2$  is a highest weight vector and spans an irreducible subrepresentation isomorphic to  $\text{Sym}^4(\mathbb{C}^2)$ . Its orthogonal complement (with respect to an invariant Hermitian inner product) is a trivial one-dimensional subrepresentation; this tells us that some linear combination of  $b^2$  and  $ac$  must be invariant!

We could (but won't) figure out precisely which combination is invariant as follows:

1. let  $Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \in \mathfrak{sl}(2, \mathbb{C})$ ,
2. consider  $a^2, Y(a^2), Y^2(a^2), Y^3(a^2), Y^4(a^2)$  - this is a basis for the highest weight space containing  $a^2$ .
3. pick an invariant inner product and let  $\Delta$  be a vector orthogonal to  $Y^2(a^2)$  in the zero weight space.

This  $\Delta$  will be  $b^2 - 4ac$ .

**Example 10.11.** If we decompose  $\text{Sym}^4(V^*)$  then we get

$$\begin{aligned}\text{Sym}^4(V^*)_{-8} &= \mathbb{C}\langle c^4 \rangle \\ \text{Sym}^4(V^*)_{-6} &= \mathbb{C}\langle bc^3 \rangle \\ \text{Sym}^4(V^*)_{-4} &= \mathbb{C}\langle b^2c^2, ac^3 \rangle \\ \text{Sym}^4(V^*)_{-2} &= \mathbb{C}\langle b^3c, abc^2 \rangle \\ \text{Sym}^4(V^*)_0 &= \mathbb{C}\langle b^4, ab^2c, a^2c^2 \rangle \\ \text{Sym}^4(V^*)_2 &= \mathbb{C}\langle ab^3, a^2bc \rangle \\ \text{Sym}^4(V^*)_4 &= \mathbb{C}\langle a^2b^2, a^3c \rangle \\ \text{Sym}^4(V^*)_6 &= \mathbb{C}\langle a^3b \rangle \\ \text{Sym}^4(V^*)_8 &= \mathbb{C}\langle a^4 \rangle\end{aligned}$$

and we see that the decomposition into irreducible representations is

$$\mathrm{Sym}^4(V^*) = \mathrm{Sym}^8 \mathbb{C}^2 \oplus \mathrm{Sym}^4 \mathbb{C}^2 \oplus \mathbb{C}.$$

There is a one-dimensional trivial summand, so there is (up to scale) precisely one quartic invariant. Since  $\Delta^2$  is a quartic invariant, this means any other quartic invariant must be a multiple of  $\Delta^2$ .

## 10.4 \* Representations of $SO(3)$

The homomorphism  $\mathrm{Ad}: SU(2) \rightarrow SO(3)$  comes from a map of Lie algebras

$$\mathfrak{su}(2) \rightarrow \mathfrak{so}(3).$$

Explicitly this map sends  $\sigma_i$  to  $2K_i$  where  $tK_i$  is the matrix which exponentiates to a rotation by an angle  $t$  around the  $x_i$ -axis. For example,

$$K_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

This map is an isomorphism of Lie algebras, so  $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ . However,  $SO(3)$  is not simply-connected, so  $SO(3) \not\cong SU(2)$  and it is not true that every representation of  $\mathfrak{so}(3)$  comes from a representation of  $SO(3)$ .

Given a complex representation  $R: SO(3) \rightarrow GL(V)$  we get a complex representation  $\tilde{R}: SU(2) \rightarrow GL(V)$  by precomposing with the double cover  $SU(2) \rightarrow SO(3)$ . We say  $\tilde{R}$  is a lift of  $R$ . The next lemma tells us which representations of  $SU(2)$  arise as lifts.

**Lemma 10.12.** *The finite-dimensional complex representations of  $SU(2)$  which arise as lifts of representations of  $SO(3)$  are precisely those whose highest weight is even (or integer spin).*

**Proof.** Since  $SU(2) \rightarrow SO(3)$  has kernel  $\{\pm 1\}$ , the representations  $\tilde{R}$  which arise as lifts all satisfy

$$\tilde{\rho}(-1) = 1.$$

Conversely, if  $\kappa$  is a representation of  $SU(2)$  with  $\kappa(-1) = 1$  then we can define a representation  $R$  of  $SO(3)$  as follows. For each  $A \in SO(3)$  pick a preimage  $\tilde{A} \in SU(2)$ . The two choices of preimage differ by a sign, so  $R(A) = \kappa(\tilde{A}) = \kappa(-\tilde{A})$  is well-defined. Since  $-1 \in SU(2)$  is central (commutes with everything), it's easy to check that  $R$  is a representation.

This means that the representations of  $SU(2)$  which are lifts of representations of  $SO(3)$  are those in which  $-1$  acts as the identity.

Remember from Lemma 10.2 that the element  $e^{t\sigma_1}$  in  $SU(2)$  acts as scalar multiplication by  $e^{imt}$  on a weight space with weight  $m$ . The element  $-1 \in SU(2)$  is  $e^{\pi\sigma_1}$ , so acts as  $e^{i\pi m}$  on a weight space with weight  $m$ . Therefore  $-1$  acts as the identity if and only if all weights are even. This proves the lemma.  $\square$

## 10.5 Outlook

The weights and weight spaces of the adjoint representation are particularly special and are called the *roots* and *root spaces*. They are special for the following reason.

Suppose that we didn't have the clever idea of picking  $H$ ,  $X$  and  $Y$  to begin with. Actually the choice of  $H$  was unimportant: we just picked an element  $H$  such that  $iH$  exponentiated to give a subgroup  $U(1) \subset SU(2)$ , which meant that its action diagonalised and we obtained a weight-space decomposition. Any other  $H$  would have done just as well, provided it exponentiated to give a subgroup  $U(1)$ .

Having picked this  $H$ , the roots will turn out to be 0 and  $\pm\lambda \in \mathbf{Z}$ . It's possible at this stage that  $\lambda \neq 2$ , for instance we might pick twice the usual  $H$ , so let's rescale our choice to ensure that  $\lambda = 2$ . Pick a generator of the root space corresponding to  $\lambda$ ; call it  $X$ . Pick a generator of the other root space and call it  $Y$ . We know

$$[H, X] = \lambda X, \quad [H, Y] = -\lambda Y$$

and, using these relations and the Jacobi identity, we get

$$\begin{aligned} \text{ad}_H[X, Y] &= [H, [X, Y]] \\ &= [X, [H, Y]] - [Y, [H, X]] \\ &= -\lambda[X, Y] - \lambda[Y, X] \\ &= 0 \end{aligned}$$

so  $[X, Y]$  is in the zero root space spanned by  $H$ , in other words

$$[X, Y] = \mu H$$

for some  $\mu$ . We can rescale  $X$  to make sure  $\mu = 1$  and in this way we can recover the Lie algebra (and this nice basis  $H, X, Y$ ) completely from a knowledge of its roots and root spaces.

In summary: the root space decomposition of the adjoint representation is exactly what we need to read off the key Lie bracket relations

$$[H, X] = 2X, \quad [H, Y] = 2Y, \quad [X, Y] = H$$

which allowed us to perform the above analysis of representations.

# Lecture 11: \* The hydrogen atom

## 11.1 Motivation

The hydrogen atom consists of a single electron sitting in the potential well of a proton. The potential of the proton is the Coulomb potential

$$V = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2}$$

where  $r$  is the radial distance of the electron from the proton and where  $\epsilon_0$  and  $q$  are constants (respectively the permittivity of free space and the electric charge of the proton). In Schrödinger's quantum mechanical model of the atom, the state of the electron is represented by a *wavefunction*: a complex-valued function  $\psi: \mathbf{R}^3 \rightarrow \mathbf{C}$  whose magnitude  $|\psi(\mathbf{x})|^2$  is the probability density that the electron is found at  $\mathbf{x}$  (to get actual probabilities one must integrate this over open sets).

To extract physical information from a wavefunction one has to apply a linear operator. For example, if one wants to know the energy of a state, one applies the *Hamiltonian operator*

$$\Delta - V$$

to the wavefunction (where  $\Delta$  is the Laplacian). In the rare case that  $\psi$  is an eigenfunction of  $\Delta - V$ ,

$$(\Delta - V)\psi = \lambda\psi,$$

we say that  $\psi$  has energy  $\lambda$ . With suitable (e.g. periodic) boundary conditions imposed, the linear span of the eigenfunctions is dense in the space of all functions: in the same way that any periodic (square-integrable) function on  $\mathbf{R}$  has a (Fourier) expansion as a convergent sum of eigenfunctions of  $d^2/dx^2$  (sines and cosines), any 2-periodic (square-integrable) function on  $\mathbf{R}^2$  has an expansion as a convergent sum of eigenfunctions of  $\Delta - V$ . Therefore most states are in a *superposition* of states of definite energy. It is not possible to assign a value of energy to these states directly in quantum mechanics and it is one of the biggest mysteries in the foundations of quantum mechanics that we can measure a definite value of energy at all. Something happens during the measurement process to extract a particular eigenstate whose energy is measured, and the system subsequently evolves from that eigenstate. The nature of this "wavefunction collapse" is, to my knowledge, not well-understood.

Nonetheless, it is clearly important to understand the eigenfunctions of the Hamiltonian operator if one wants to understand the orbitals of the hydrogen atom.

## 11.2 The spherical Laplacian

The Laplacian is the operator

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

We will now write this in a slightly different form which is more amenable to the rotational symmetry of the situation.

**Lemma 11.1.** *In coordinates  $x, y, z$  with  $r^2 = x^2 + y^2 + z^2$  we can write*

$$\Delta = \left( \partial_r^2 + \frac{2}{r} \partial_r \right) + \frac{1}{r^2} (J_x^2 + J_y^2 + J_z^2)$$

where  $J_x = y\partial_z - z\partial_y$  and  $J_y, J_z$  are defined by cyclically permuting  $x, y, z$ . We write

$$\Delta_S = J_x^2 + J_y^2 + J_z^2$$

and call it the spherical Laplacian. We have

$$\Delta_S(F_1 F_2) = F_1 \Delta_S(F_2)$$

if  $F_1$  is a function of  $r$  alone.

If we separate variables in the equation  $\Delta\psi - V\psi = \lambda\psi$  by assuming  $\psi$  has the form  $R(r)Y(\mathbf{u})$  where  $\mathbf{u} \in S^2$  is a unit vector, then

$$0 = \frac{\Delta(RY) - VRY - \lambda RY}{RY} = \frac{(\partial_r^2 + \frac{2}{r}\partial_r) R}{R} + \frac{\Delta_S Y}{r^2 Y} - V(r) - \lambda$$

i.e.

$$r^2 \frac{(\partial_r^2 + \frac{2}{r}\partial_r) R}{R} - r^2 V(r) - r^2 \lambda = \frac{\Delta_S Y}{Y}$$

where both sides must be constant because the left depends only on  $r$  and the right only on  $\mathbf{u}$ . Thus

$$\Delta_S Y = \kappa Y \text{ and } r^2 R'' + 2rR' - (V + \lambda)r^2 R = \kappa R$$

for some  $\kappa$  (where  $R'$  denotes  $dR/dr$ ).

Much of the information is therefore contained in the eigenfunctions of the spherical Laplacian.

### 11.3 Rotational symmetry

The rotation group  $SO(3)$  acts on  $S^2 \subset \mathbf{R}^3$  and hence on the space of functions  $F: S^2 \rightarrow \mathbf{C}$  by

$$g(F)(x) = F(g^{-1}(x)).$$

We need the inverse of  $g$  so that

$$(gh)(F)(x) = F(h^{-1}g^{-1}(x))$$

equals

$$g(h(F))(x) = h(F)(g^{-1}(x)) = F(h^{-1}g^{-1}(x))$$

giving a representation.



Moreover the rotation group acts on *operators*. Operators are linear transformations  $H: \mathcal{F} \rightarrow \mathcal{F}$  on the space  $\mathcal{F}$  of functions, so the action is given by

$$g(H)(F) = (H(g^{-1}(F)))$$

**Lemma 11.2.** *The spherical Laplacian is invariant under rotations, in other words*

$$\Delta_S(g^{-1}F) = \Delta_S(F).$$

**Proof.** This is certainly true of the usual Laplacian, and the spherical Laplacian differs from  $\Delta$  only by differential operators in  $r$  which are also invariant.  $\square$

We would like to understand  $\mathcal{F}$  as a representation of  $SO(3)$ , but it is an infinite-dimensional space so we restrict to the finite-dimensional subspace  $P_n$  of polynomials of degree  $n$ . Here we mean complex polynomial functions of  $x, y, z$  restricted to  $S^2$ .

The rotation action of  $SO(3)$  on  $\mathbf{R}^3$  is just the standard representation. The functions  $x, y, z$  are linear functions on  $\mathbf{R}^3$  and so can be considered as elements of the dual of the standard representation. Homogeneous polynomials of degree  $n$  are then elements of the symmetric  $n$ -th power of this dual representation. Thus

$$P_n = \text{Sym}^n(\mathbf{R}^3)^* \oplus \text{Sym}^{n-1}(\mathbf{R}^3)^* \oplus \cdots \oplus \text{Sym}^1(\mathbf{R}^3)^* \oplus \mathbf{R}$$

where  $\mathbf{R}$  is the trivial one-dimensional representation corresponding to constant polynomials. This is the decomposition of  $P_n$  into irreducibles.

## 11.4 The Lie algebra action

Consider the action of  $\mathfrak{so}(3)$  on  $\mathbf{R}^3$  which exponentiates to the standard representation of  $SO(3)$ . Remember that  $\mathfrak{so}(3)$  is the Lie algebra consisting of antisymmetric matrices, for instance, we have

$$\exp t \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}.$$

Let  $tK_x, tK_y, tK_z$  be the matrices which exponentiate to give rotations by an angle  $t$  around the  $x, y, z$ -axes respectively. For instance,

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

The action of this on polynomials is just the dual (or transpose) action, so on  $P_1$ :

$$K_x(x) = 0, K_x(y) = -z, K_x(z) = y.$$

It is easy to check that the action of  $K_x$  agrees with the action of  $y\partial_z - z\partial_y = J_x$  on  $P_1$ . In fact:

**Lemma 11.3.** *We have*

$$K_{x_i}P = J_{x_i}P$$

where  $x_1, x_2, x_3 = x, y, z$  and  $P$  is any smooth function of  $x, y, z$ , in particular any polynomial.

**Proof.** Consider the space  $V$  of complex functions on  $\mathbf{R}^3$ . A linear map  $d: V \rightarrow V$  satisfying Leibniz's rule  $d(xy) = (dx)y + xdy$  is called a *derivation*. The commutator of two derivations is again a derivation, so derivations form a Lie algebra and  $V$  is a representation of that algebra. Note that if  $d$  is a derivation then its action on any function is determined by its action on  $x, y$  and  $z$ : for example  $d(x^2) = 2xdx$ .

The operators  $J_x, J_y, J_z$  obey Leibniz's rule and satisfy  $J_{x_i}x_j = K_{x_i}x_j$  so if we can show that each  $K_{x_i}$  satisfies Leibniz's rule we are done. This will follow from a general fact about Lie group actions.

Suppose that a Lie group  $G$  acts smoothly on  $\mathbf{R}^n$  (in other words, each transformation  $g: \mathbf{R}^n \rightarrow \mathbf{R}^n$  is a diffeomorphism). Then it acts on the space  $V$  of smooth functions on  $\mathbf{R}^n$ :

$$R(g)F(x) = F(g^{-1}x)$$

This defines an infinite-dimensional representation of  $G$ . One can take the derivative  $\frac{d}{dt}\big|_{t=0} R(e^{tX})F$  to find the corresponding Lie algebra representation. We claim that elements of the Lie algebra obey Leibniz's rule.

To prove this, consider the action of  $G$  on a product of functions:

$$\exp(tX)(F_1F_2)(x) = (F_1F_2)(\exp(-tX)x)$$

and differentiate with respect to  $t$  at  $t = 0$  to find the Lie algebra action of  $X$ . The chain rule now implies Leibniz's formula for this Lie algebra representation, hence  $X$  acts as a derivation.  $\square$

Remember that  $SU(2) \rightarrow SO(3)$  is a double covering homomorphism and so it induces a homomorphism of Lie algebras. Since  $\exp(t\sigma_i)$  is sent to the rotation by an angle  $2t$  around the  $x_i$ -axis, the Pauli matrix  $\sigma_i$  goes to  $2K_{x_i}$ . This is actually an isomorphism of Lie algebras (not an isomorphism of Lie groups because  $SO(3)$  is not simply-connected). In particular, the complexification of  $\mathfrak{so}(3)$  is  $\mathfrak{sl}(2, \mathbf{C})$  and we pick a standard basis

$$H = -2iK_z, \quad X = K_x + iK_y, \quad Y = K_x - iK_y$$

of  $\mathfrak{sl}(2, \mathbf{C})$ .

In our representation,  $J_{x_i} = K_{x_i}$ . The spherical Laplacian is

$$J_x^2 + J_y^2 + J_z^2.$$

If  $\rho$  is the representation of  $\mathfrak{sl}(2, \mathbf{C})$  associated to the representation  $P_n$  of  $\mathfrak{so}(3)$  then

$$-\frac{1}{4}\rho(H^2) + \frac{1}{2}(\rho(XY) + \rho(YX)) = J_x^2 + J_y^2 + J_z^2 = \Delta_S$$

so we can compute the eigenvalues of  $\Delta_S$  using the fact that

$$\rho(X)\rho(Y) + \rho(Y)\rho(X) = 2\rho(X)\rho(Y) - \rho(H)$$

and the computation from Theorem 10.5 which showed:

$$\rho(X)\rho(Y)^kv = (m - k + 1)k\rho(Y)^{k-1}v$$

for a highest weight vector  $v$  of weight  $m$ . The result is:

**Theorem 11.4.** *The eigenvalues of  $\Delta_S$  on  $P_n$  are  $m(m+1)$ ,  $0 \leq m \leq n$ . The  $m(m+1)$ -eigenspace has dimension  $2m + 1$ .*

**Proof.** Omitting  $\rho$  from the notation, we have

$$\begin{aligned}\Delta_S &= \frac{1}{4}H^2 + \frac{1}{2}(XY + YX) \\ &= \frac{1}{4}H^2 - \frac{1}{2}H + XY.\end{aligned}$$

If  $v_m$  is a highest weight vector (with weight<sup>2</sup>  $2m$ ) for the subrepresentation  $\text{Sym}^m(\mathbf{R}^3)^*$  then this subrepresentation is spanned by vectors

$$Y^k v_m, \quad k = 0, \dots, 2m.$$

From the computations we did while classifying representations of  $\mathfrak{sl}(2, \mathbf{C})$ , we know that

$$HY^k v_m = (2m - 2k)Y^k v_m, \quad XY(Y^k v_m) = (2m - k)(k + 1)Y^k v_m.$$

Therefore

$$\left( \frac{1}{4}H^2 - \frac{1}{2}H + XY \right) Y^k v_m = (m - k)^2 - (m - k) + (2m - k)(k + 1)$$

or

$$m(m + 1).$$

In other words, the subspace  $\text{Sym}^m(\mathbf{R}^3)^*$  is a  $\Delta_S$ -eigenspace with eigenvalue  $m(m + 1)$ .  $\square$

*Remark 11.5.* Note that the number  $2(m - k)$  is itself an eigenvalue: the weight of  $H$  acting on the corresponding weight-space. Since  $H = -2iJ_z$ , the number  $i(m - k)$  is interpreted as an eigenvalue of the *angular momentum operator*  $J_z$ . Thus states in the hydrogen atom (with polynomial wavefunction) can be specified by giving the energy  $m(m + 1)$  where  $m \in \mathbf{Z}$  and the  $z$ -component of angular momentum, which is one of the  $2m + 1$  even numbers between  $-2m$  and  $2m$ .

If you want some explicit eigenfunctions, see [?, Section 4.9].

---

<sup>2</sup>Note that this is a representation of  $SO(3)$  and therefore has even weight, or integer spin.

## Lecture 12: Representations of $SU(3)$ , I

Having classified the irreducible representations of  $SU(2)$ , we now move on to the more complicated case of  $SU(3)$ . This case presents essentially all of the issues and complications that arise in the general theory of semisimple Lie groups.

### 12.1 Strategy

In the case of  $SU(2)$ , given a complex representation  $R: SU(2) \rightarrow GL(V)$ :

1. We took the subgroup  $T = U(1) \rightarrow SU(2)$  of diagonal matrices  $\begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}$  and considered the composition  $R|_T: U(1) \rightarrow SU(2) \rightarrow GL(V)$ .
2. We decomposed  $V$  into weight spaces  $V_n = \left\{ v : R \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix} v = e^{in\phi} v \right\}$ .
3. We then considered the corresponding complexified Lie algebra representation  $\rho: \mathfrak{sl}(2, \mathbb{C}) \rightarrow \mathfrak{gl}(V)$ , so that we would make sense of the action of the matrix  $H \in \mathfrak{sl}(2, \mathbb{C})$ , and we get  $V_n = \{v : \rho(H)v = nv\}$  (saves writing  $i$  everywhere; we should technically write  $\rho_{\mathbb{C}}$ ).
4. We took the complex basis  $X, Y, H$  of  $\mathfrak{sl}(2, \mathbb{C})$  and analysed how the elements  $\rho(X)$  and  $\rho(Y)$  acted on the weight spaces:

$$\rho(X)V_n \subset V_{n+2}, \quad \rho(Y)V_n \subset V_{n-2}.$$

The reason this worked was because of the commutation relations  $[H, X] = 2X$  and  $[H, Y] = -2Y$ , i.e. because  $X$  and  $Y$  are weight vectors in the adjoint representation.

5. By considering the commutation relations between  $X, Y, H$  we showed that a highest weight vector  $v \in V_m$  ( $m$  maximal) generated an irreducible subrepresentation with one-dimensional weight spaces  $V_{-m} \oplus V_{-m+2} \oplus \cdots \oplus V_{m-2} \oplus V_m$ .

Given a complex representation  $R: SU(3) \rightarrow GL(V)$ :

1. We will take the subgroup  $T = U(1)^2 \rightarrow SU(3)$  of diagonal matrices  $\begin{pmatrix} e^{i\phi_1} & 0 & 0 \\ 0 & e^{i\phi_2} & 0 \\ 0 & 0 & e^{-i(\phi_1+\phi_2)} \end{pmatrix}$ , let  $\mathfrak{t}$  denote its Lie algebra, and consider the composition  $R|_T: T \rightarrow SU(3) \rightarrow GL(V)$ .
2. We will decompose  $V$  into weight spaces  $V_\lambda = \{v : R(\exp(t))v = e^{i\lambda(t)}v \forall t \in \mathfrak{t}\}$ .
3. We will take the complexified Lie algebra representation  $\rho: \mathfrak{sl}(3, \mathbb{C}) \rightarrow \mathfrak{gl}(V)$  so that we can make sense of  $\rho_{\mathbb{C}}(H_{12})$  and  $\rho_{\mathbb{C}}(H_{23})$  where  $H_{12} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  and  $H_{23} =$

$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$  are elements of  $\mathfrak{t}_{\mathbb{C}}$ . Then  $V_{\lambda} = \{v : \rho_{\mathbb{C}}(H)v = \lambda(H)v \forall H \in \mathfrak{t}_{\mathbb{C}}\}$ .

4. We will take basis elements  $E_{ij}$ ,  $i < j$ , and analyse how the elements  $\rho(E_{ij})$  act on the weight spaces  $V_{\lambda}$ . This will work because we will pick  $E_{ij}$  to be weight vectors for the adjoint representation.
5. By considering the commutation relations between  $E_{ij}, H_{jk}$  we will show that a “highest weight vector”  $v \in V_{\lambda}$  generates an irreducible subrepresentation, in a suitable sense.

## 12.2 The Lie algebra

The Lie algebra of  $\mathfrak{su}(3)$  is the space of tracefree skew-Hermitian complex three-by-three matrices. Its complexification is  $\mathfrak{sl}(3, \mathbb{C})$ , the Lie algebra of tracefree complex three-by-three matrices and a basis for  $\mathfrak{sl}(3, \mathbb{C})$  is

$$\begin{aligned} H_{12} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & H_{23} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \\ E_{12} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & E_{23} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, & E_{13} &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ E_{21} &= E_{12}^T, & E_{32} &= E_{23}^T, & E_{31} &= E_{13}^T \end{aligned}$$

(so  $\dim_{\mathbb{C}} \mathfrak{sl}(3, \mathbb{C}) = 8$ ).

We will write  $\mathfrak{h} := \mathfrak{t}_{\mathbb{C}} = \mathfrak{t} \oplus i\mathfrak{t}$ , the abelian Lie algebra of diagonal complex matrices with trace zero. Note that  $i\mathfrak{t}$  consists of the diagonal matrices in  $\mathfrak{sl}(3, \mathbb{R})$ , so we will write this as  $\mathfrak{h}_{\mathbb{R}}$ .

A general element of  $\mathfrak{h}$  is a  $\mathbb{C}$ -linear combination of  $H_{12}$  and  $H_{23}$ . Another way of writing  $\mathfrak{h}$  is as:

$$\mathfrak{h} = \left\{ \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} : a_1 + a_2 + a_3 = 0 \right\}.$$

## 12.3 The weight lattice

Recall that we want to think of weights as elements of  $\mathfrak{h}^*$ . The weight lattice is the set of  $\lambda \in \mathfrak{h}^*$  such that  $\lambda(X) \in 2\pi i\mathbb{Z}$  for all  $X \in \ker \exp$ .

What is  $\ker \exp$ ? We have

$$\exp \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} = \begin{pmatrix} e^{a_1} & 0 & 0 \\ 0 & e^{a_2} & 0 \\ 0 & 0 & e^{a_3} \end{pmatrix}$$

which equals the identity if and only if  $a_1, a_2, a_3 \in 2\pi i\mathbf{Z}$ .

Let  $L_i \in \mathfrak{h}^*$  denote the coordinate function

$$L_i \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} = a_i.$$

Then  $\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \in \ker \exp$  if and only if  $L_k \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \in 2\pi i\mathbf{Z}$  for  $k = 1, 2, 3$ .

Therefore these elements span the weight lattice. They satisfy  $L_1 + L_2 + L_3 = 0$ .

How should we represent the weight lattice? It should be a lattice in  $\mathbf{R}^2$  (thinking of  $\mathbf{R}^2 = \mathfrak{h}_{\mathbf{R}}^*$ ) and the elements  $L_1, L_2, L_3$  should have centre of mass at the origin and sit in a symmetric way, so we will put them at the vertices of an equilateral triangle (say  $L_1 = (1, 0)$ ,  $L_2 = (\cos(2\pi/3), \sin(2\pi/3))$ ,  $L_3 = (\cos(4\pi/3), \sin(4\pi/3))$ ). We will be able to justify this geometrical choice later when we discuss the Killing form (a natural inner product on the Lie algebra of any semisimple Lie group).

If  $R: SU(3) \rightarrow GL(V)$  is a representation we can now consider it as a representation of the diagonal subgroup  $T$  and hence decompose it as a direct sum of weight spaces

$$V = \bigoplus_{\lambda \in \Lambda \subset \mathfrak{t}_{\mathbf{Z}}^*} V_{\lambda}, \quad V_{\lambda} = \{v : \rho(H)v = \lambda(H)v\}.$$

Here  $H = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}$  and if  $\lambda = A_1 L_1 + A_2 L_2 + A_3 L_3$  then  $\lambda(H) = \sum A_i a_i$ . In other words

$$V_{\lambda} = \{v : R(\exp(iH))v = \exp(i\lambda(H))v\}.$$

## 12.4 Root spaces

To continue our analysis, we want to find the elements of  $\mathfrak{sl}(3, \mathbf{C})$  analogous to  $X$  and  $Y$  in  $\mathfrak{sl}(2, \mathbf{C})$ . As the example of  $\mathfrak{sl}(2, \mathbf{C})$  suggested, we need to look at the weight space decomposition of the adjoint representation to find them.

**Definition 12.1.** The weight spaces of the adjoint representation (for any Lie group) are called root spaces.

Suppose that

$$H = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \in \mathfrak{h}$$

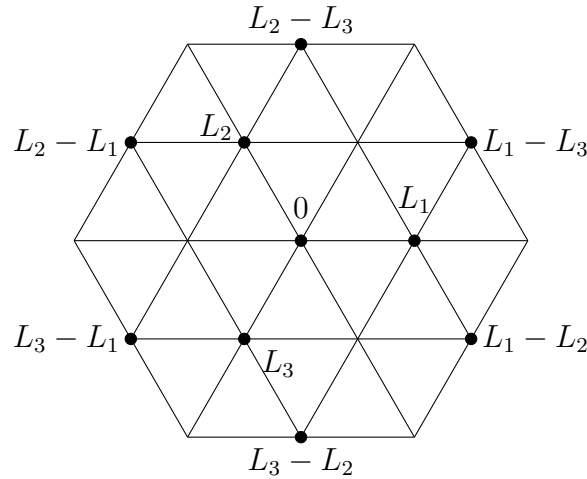
(so  $a_1 + a_2 + a_3 = 0$ ) and that  $E_{ij}$  ( $i \neq j$ ) is the matrix with 1 in the  $(i, j)$ th entry and zeros elsewhere. Then we have

$$\text{ad}_H E_{ij} = (a_i - a_j)E_{ij}.$$

Therefore,

**Lemma 12.2.** *The root spaces of  $\mathfrak{sl}(3, \mathbb{C})$  are  $\mathfrak{h}$  with root 0 and (for each pair  $1 \leq i, j \leq 3, i \neq j$ )  $\mathbb{C}E_{ij}$  with root  $L_i - L_j$ .*

Here  $L_i$  is the functional on  $\mathfrak{h}$  which sends  $\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}$  to  $a_i$ . We have already decided to represent  $L_1, L_2, L_3$  as the vertices of an equilateral triangle, so the vectors  $\pm(L_1 - L_2), \pm(L_2 - L_3), \pm(L_3 - L_1)$  sit at the vertices of a hexagon



**Definition 12.3.** We write

$$R = \{L_i - L_j\}_{i \neq j}$$

and we call  $R$  the *root system* of  $\mathfrak{sl}(3, \mathbb{C})$ . The *root diagram* is the picture of the roots as vectors in  $\mathfrak{h}_{\mathbb{R}}^*$ . For each root  $\alpha \in R$  we write  $\mathfrak{sl}(2, \mathbb{C})_{\alpha}$  for the corresponding *root space*

$$\mathfrak{sl}(3, \mathbb{C})_{\alpha} = \{v \in \mathfrak{sl}(3, \mathbb{C}) : \text{ad}_H v = \alpha(H)v \ \forall H \in \mathfrak{h}\}$$

**Lemma 12.4.** *If  $v \in \mathfrak{sl}(3, \mathbb{C})_{\alpha}$  and  $w \in \mathfrak{sl}(3, \mathbb{C})_{\beta}$  then*

$$[v, w] \in \mathfrak{sl}(3, \mathbb{C})_{\alpha+\beta}$$

**Proof.** If  $H \in \mathfrak{h}$  then, by the Jacobi identity,

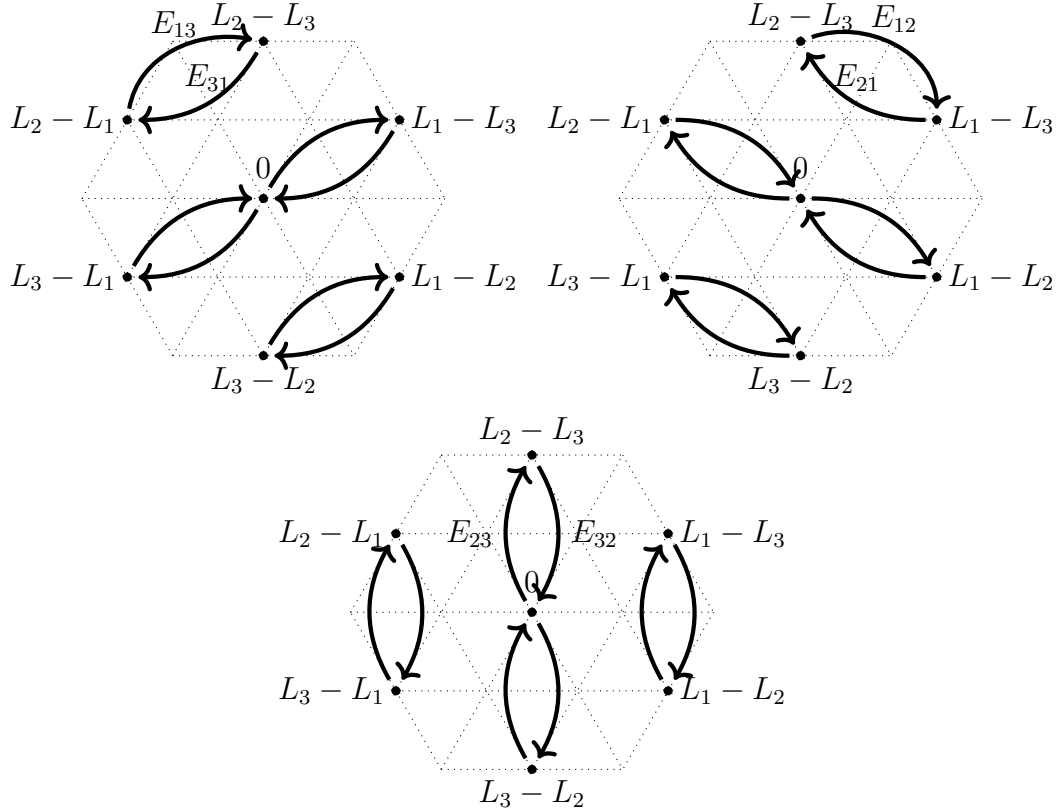
$$\begin{aligned} \text{ad}_H[v, w] &= [H, [v, w]] \\ &= [v, [H, w]] - [w, [H, v]] \\ &= (\beta(H) + \alpha(H))[v, w] \end{aligned}$$

so

$$[v, w] \in \mathfrak{sl}(3, \mathbb{C})_{\alpha+\beta}.$$

□

We illustrate this lemma in the following sequence of diagrams.



## 12.5 $\mathfrak{sl}(2, \mathbb{C})$ -subalgebras

**Lemma 12.5.** *For each root  $\alpha \in R$ , the subspace*

$$\mathfrak{s}_\alpha = \mathfrak{sl}(3, \mathbb{C})_{-\alpha} \oplus [\mathfrak{sl}(3, \mathbb{C})_\alpha, \mathfrak{sl}(3, \mathbb{C})_{-\alpha}] \oplus \mathfrak{sl}(3, \mathbb{C})_\alpha \subset \mathfrak{sl}(3, \mathbb{C})$$

*is a Lie subalgebra isomorphic to  $\mathfrak{sl}(2, \mathbb{C})$ .*

**Proof.** Consider the root  $\alpha = L_i - L_j$ . Note that the root spaces  $\mathfrak{sl}(3, \mathbb{C})_\alpha$ ,  $\alpha \neq 0$  are one-dimensional. Pick generators  $E_{ij} = X_\alpha \in V_\alpha$  and  $E_{ji} = Y_\alpha \in \mathfrak{sl}(3, \mathbb{C})_{-\alpha}$ . One can check  $0 \neq H_{12} = H_\alpha = [X_\alpha, Y_\alpha] = E_{ii} - E_{jj}$ . This element generates  $[\mathfrak{sl}(3, \mathbb{C})_\alpha, \mathfrak{sl}(3, \mathbb{C})_{-\alpha}]$ . Moreover,

$$[H_\alpha, X_\alpha] = \alpha(H_\alpha)X_\alpha, \quad [H_\alpha, Y_\alpha] = -\alpha(H_\alpha)Y_\alpha$$

and it's easy to check that  $\alpha(H_\alpha) = 2$ . Provided  $\alpha(H_\alpha) \neq 0$  we could always achieve  $\alpha(H_\alpha) = 2$  by rescaling  $X_\alpha$ .

We see explicitly that this subalgebra is isomorphic to  $\mathfrak{sl}(2, \mathbb{C})$  by identifying this newly constructed basis with the standard basis of  $\mathfrak{sl}(2, \mathbb{C})$ . □



*Remark 12.6.* It is clear from the proof that the key fact we used was

$$\alpha([X_\alpha, Y_\alpha]) \neq 0.$$

*Remark 12.7.* Each line through zero in the root system defines us such a subalgebra (since  $\mathfrak{sl}(3, \mathbb{C})_\alpha = \mathfrak{sl}(3, \mathbb{C})_{-\alpha}$ ) so there are three distinguished subalgebras. When I say distinguished, remember that all this depended on the choice of  $\mathfrak{h}$ : we could have picked a different two-dimensional abelian subalgebra.

## 12.6 Highest weight vectors

Now suppose that  $\rho: \mathfrak{sl}(3, \mathbb{C}) \rightarrow \mathfrak{gl}(V)$  is a finite-dimensional irreducible complex representation and let  $V = \bigoplus_{\alpha \in \Lambda} V_\alpha$  be the weight-space decomposition. We have three distinguished subalgebras  $\mathfrak{s}_\alpha$  and we study how they act on  $V$ .

**Lemma 12.8.** *If  $X \in \mathfrak{sl}(3, \mathbb{C})_\alpha$  and  $v \in V_\beta$  then*

$$\rho(X)v \in V_{\alpha+\beta}.$$

**Proof.** If  $H \in \mathfrak{h}$  then

$$\begin{aligned} \rho(H)(\rho(X)v) &= \rho([H, X])v + \rho(X)(\rho(H)v) \\ &= \alpha(H)\rho(X)v + \beta(H)\rho(X)v \\ &= (\alpha(H) + \beta(H))\rho(X)v \end{aligned}$$

so

$$\rho(X)v \in V_{\alpha+\beta}.$$

□

To make full use of this observation, we need the analogue of a highest-weight vector. This is not as straightforward as before, because the plane has no canonical ordering.

**Definition 12.9.** A linear functional  $\pi: \mathfrak{h}_{\mathbb{R}}^* \rightarrow \mathbb{R}$  is *irrational with respect to a lattice  $L$*  if for any  $\alpha, \beta \in L$  we have  $\pi(\alpha) = \pi(\beta)$  if and only if  $\alpha = \beta$ .

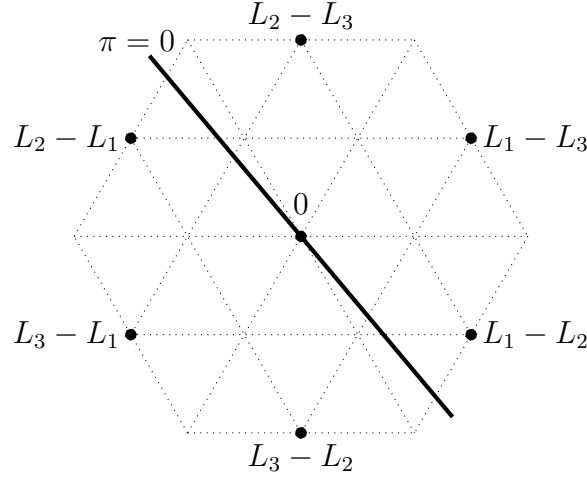
**Example 12.10.** If  $e_1, \dots, e_k$  is an integral basis for the lattice  $L$  and  $\mu_1, \dots, \mu_k \in \mathbb{R}$  are linearly independent over  $\mathbb{Q}$  then the functional  $\pi(\sum a_i e_i) = \sum a_i \mu_i$  is irrational with respect to  $L$ . This is because if  $\alpha = \sum a_i e_i$  and  $\beta = \sum b_i e_i$  then  $\pi(\alpha) = \pi(\beta)$  if and only if  $\sum (a_i - b_i) \mu_i = 0$ . If this vanishes then  $a_i - b_i = 0$  because otherwise we would get a linear dependence between  $\{\mu_i\}_{i=1}^k$  over  $\mathbb{Q}$ .

**Definition 12.11** (Highest weight vector). Pick a linear functional  $\pi$  on  $\mathfrak{h}_{\mathbb{R}}^*$  which is irrational with respect to  $\mathfrak{h}_{\mathbb{Z}}$ . The weight  $\alpha$  such that

$$\pi(\alpha) = \max_{\beta \in \Lambda} \pi(\beta)$$

is called the highest weight. Such a weight exists (because there are finitely many weights) and is unique by irrationality of  $\pi$  with respect to  $\mathfrak{h}_{\mathbb{Z}}$  (which contains  $\Lambda$ ). If  $\alpha$  is the highest weight then any nonzero  $v \in V_\alpha$  is called a *highest weight vector*.

**Definition 12.12** (Positive and negative roots). Since  $\pi$  is irrational with respect to  $\mathfrak{h}_\mathbb{Z}$  which contains the roots, none of the nonzero roots  $\gamma \in R$  has  $\pi(\gamma) = 0$ . Since  $\gamma \in R$  implies  $-\gamma \in R$  we see that precisely half of the roots are positive (i.e.  $\pi(\gamma) > 0$ ) and half are negative ( $\pi(\gamma) < 0$ ). We write this decomposition  $R = R_+ \cup R_-$ . For convenience, we will pick  $\pi$  so that  $R_+ = \{L_1 - L_2, L_1 - L_3, L_2 - L_3\}$ .



**Definition 12.13.** If a vector  $u$  lies in some weight-space  $V_\delta$  for some representation  $\rho$  then it will be convenient to write  $w_\rho(u) = \delta$  for the weight of  $u$ . It will be even more convenient to drop the  $\rho$  and hope that it is clear which representation we mean. If  $\rho$  is the adjoint representation, if  $u$  is in the root space for the root  $\alpha$  we will write  $r(u) = \alpha$ .

**Corollary 12.14.** If  $v \in V_\alpha$  is a highest weight vector for a representation  $\rho$ ,  $X \in \mathfrak{sl}(3, \mathbb{C})$  and  $r(X) \in R_+$ , then  $\rho(X)v = 0$ .

**Proof.** The lemma implies that  $\rho(X)v \in V_{\alpha+r(X)}$  but  $\pi(\alpha + r(X)) > \pi(\alpha)$  and  $\pi(\alpha)$  is maximal because  $v$  is a highest weight vector, so  $\rho(X)v = 0$ .  $\square$

In the case of  $\mathfrak{sl}(2, \mathbb{C})$  we showed that an irreducible representation was spanned by elements of the form  $v, Yv, Y^2v, \dots, Y^m v$ . The analogue for  $\mathfrak{sl}(3, \mathbb{C})$  is:

**Proposition 12.15.** If  $V$  is a complex linear complex representation of  $\mathfrak{sl}(3, \mathbb{C})$  and  $v$  is a highest weight vector with weight  $\alpha$  then  $v$  is contained in a unique subrepresentation spanned by all elements of the form

$$\rho(X_1)\rho(X_2) \cdots \rho(X_n)v$$

where  $X_i \in \mathfrak{sl}(3, \mathbb{C})_{\alpha_i}$  for some sequence of negative roots  $\alpha_i \in R_-$ . With our choice of  $\pi$  we may assume that  $X_i \in \{E_{21}, E_{31}, E_{32}\}$ . As a consequence, if  $V$  is irreducible, then it is equal to this subrepresentation and hence spanned by the vectors  $\rho(X_1) \cdots \rho(X_n)v$ , taken over all sequences of negative root vectors  $X_i$ .

**Proof.** To prove the proposition it suffices to show that the subspace  $W \subset V$  spanned by these elements is preserved by the action of  $\mathfrak{sl}(3, \mathbb{C})$  because then it is a subrepresentation and irreducibility implies it must agree with  $V$ . Certainly by construction  $W$  is preserved by the action of  $E_{21}, E_{31}, E_{32}$ .

Next we check that if  $H \in \mathfrak{h}$  then  $\rho(H)W \subset W$ :

$$\begin{aligned} \rho(H)\rho(X_1)\cdots\rho(X_n)v &= \rho([H, X_1])\rho(X_2)\cdots\rho(X_n)v + \rho(X_1)\rho(H)\rho(X_2)\cdots\rho(X_n)v \\ &= \alpha_1(H)\rho(X_1)\cdots\rho(X_n)v + \rho(X_1)\rho(H)\rho(X_2)\cdots\rho(X_n)v \\ &= \cdots \\ &= \left( \alpha(H) + \sum_{i=1}^n \alpha_i(H) \right) \rho(X_1)\cdots\rho(X_n)v \in W \end{aligned}$$

using the fact that  $[H, X_i] = \alpha_i(H)X_i$  and that  $Hv = \alpha(H)v$ .

Finally we need to check that  $E_{12}, E_{13}, E_{23}$  preserve  $W$ . This is an excellent inductive exercise in Lie algebra. If you can't do it yourself, see the proof of [?, Claim 12.10].  $\square$

## 12.7 Applying $\mathfrak{s}_\alpha$

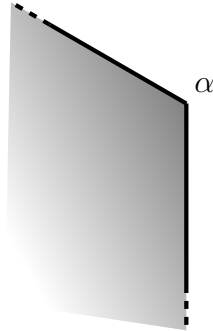
Together, Lemma 12.8 and Proposition 12.15 imply:

- the highest weight space  $V_\alpha$  is one-dimensional (spanned by  $v$ ) because all other vectors

$$u = \rho(X_1)\rho(X_2)\cdots\rho(X_n)v, \quad w(X_i) \in R_-, \quad n \geq 1$$

have  $\pi(w(u)) < \pi(\alpha)$ .

- Similarly, the weight spaces  $V_{\alpha+k(L_2-L_1)}$  and  $V_{\alpha+k(L_3-L_2)}$  are at most one-dimensional (spanned by  $\rho(E_{21})^k v$  and  $\rho(E_{32})^k v$  respectively).
- Moreover, all weights which occur in the representation  $V$  are contained in the shaded subspace of  $\mathfrak{h}_\mathbb{R}^*$ :



To get more information we apply the subalgebras  $\mathfrak{s}_\beta \cong \mathfrak{sl}(2, \mathbb{C})$  to  $v$  for  $\beta = L_2 - L_1$  and  $\beta = L_3 - L_2$ . The space  $V$  splits as a direct sum of  $\mathfrak{s}_\beta$ -representations.

### 12.7.1 The upper edge

Along the upper edge of the shaded region we see that the direct sum of weight spaces  $V_{\alpha+k(L_2-L_1)}$  living here gives a representation of  $\mathfrak{s}_{L_2-L_1}$ . This is the irreducible representation of  $\mathfrak{sl}(2, \mathbb{C})$  generated by the highest weight vector  $v$ . The weights of these  $V_{\alpha+k(L_2-L_1)}$

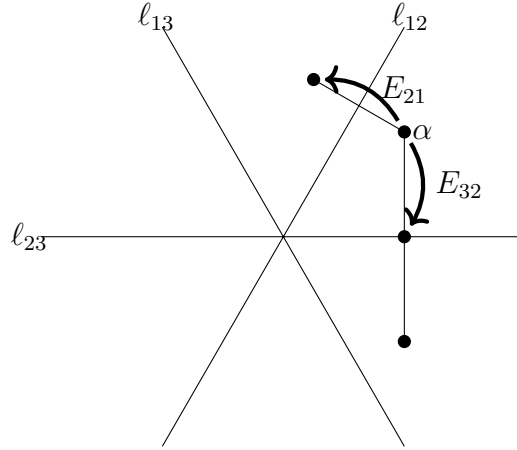


Figure 1: The edges of the weight diagram are symmetric about the lines  $\ell_{ij}$ . Fill in the rest of the diagram!

as  $\mathfrak{sl}_{L_2-L_1}$ -weight spaces are  $\alpha(H_{12}) - 2k$ : remember that the diagonal element in  $\mathfrak{sl}_{L_2-L_1}$  is  $H_{12}$  and  $(L_2 - L_1)(H_{12}) = 2$ .

Using our knowledge of irreducible representations of  $\mathfrak{sl}(2, \mathbb{C})$ , this confirms that all weight spaces along the upper edge are at most one-dimensional. In fact it tells us much more: they have to come in an unbroken sequence and they terminate when  $k$  reaches  $\alpha(H_{12})$ .

Geometrically, the edge is parallel to  $L_2 - L_1$  and therefore it is orthogonal to the line  $\ell_{12} = \{\beta : \beta(H_{12}) = 0\}$  which is parallel to  $L_3$  (remember that  $L_1 \cdot L_3 = L_2 \cdot L_3 = 1$  and that  $L_3(H_{12}) = 0$ ). The weights of the  $\mathfrak{sl}_{L_2-L_1}$ -representation along the edge are the values  $\beta(H_{12})$  and they are symmetric about zero, by the classification theorem for representations of  $\mathfrak{sl}(2, \mathbb{C})$ . This means that  $\ell_{12}$  intersects the edge in its midpoint, so the edge is symmetric under reflection in  $\ell_{12}$ .

### 12.7.2 The right edge

Similarly we can analyse the action of  $\mathfrak{sl}_{L_3-L_2}$  on the weight spaces  $V_{\alpha+k(L_3-L_2)}$  sitting along the right edge of the shaded region. These come in an unbroken sequence of one-dimensional spaces which terminate at  $k = \alpha(H_{23})$ .

### 12.7.3 Rotating $\pi$

We can now carry out the same analysis but using  $\alpha_1 = \alpha + \alpha(H_{12})(L_2 - L_1)$  as a highest weight, simply by rotating our irrational linear functional  $\pi$ . There are six possible ways of dividing  $R$  into  $R_-$  and  $R_+$  using an irrational functional  $\pi$ : for  $\alpha_1$  to be a highest weight we will need  $R_+ = \{L_2 - L_3, L_1 - L_3, L_2 - L_1\}$ .

After doing this process six times we see that the weights must lie in a hexagon whose edges are parallel to the root vectors. Moreover, by analysing the action of the distinguished  $\mathfrak{sl}(2, \mathbb{C})$ -subalgebras we see that:

- the weight spaces along the edges of this hexagon are each one-dimensional
- the hexagon is symmetric under the reflections in the lines  $\ell_{ij} = \{\beta : \beta(H_{ij}) = 0\}$ ,
- points  $\beta$  contained in the intersection of the weight lattice  $\mathfrak{h}_{\mathbb{Z}}^*$  with the interior of the hexagon occur as weights of the representation: each weight space on an edge generates an irreducible representation of one of the distinguished  $\mathfrak{sl}(2, \mathbb{C})$ -subalgebras, the weights of which lie on a line inside the hexagon parallel to one of its edges. Each  $\beta$  in the interior of the hexagon lives on one of these lines.

**Definition 12.16.** The group of reflections in the lines  $\ell_{ij}$  is isomorphic to the permutation group  $S_3$  and is called the *Weyl group* of  $\mathfrak{sl}(3, \mathbb{C})$ .

*Remark 12.17.* Note that the hexagon corresponding to a representation might actually be a triangle if some  $\alpha(H_{ij}) = 0$ . In this case we say the hexagon is *degenerate*.

## 12.8 Uniqueness

We want to prove that there is a unique irreducible representation with a given highest weight. We need the following observation.

**Lemma 12.18.** *Suppose that  $V$  and  $W$  are irreducible representations of a Lie algebra  $\mathfrak{g}$ . A homomorphism  $f : V \rightarrow W$  of representations is either zero or an isomorphism.*

**Proof.** The kernel of  $f$  is a subrepresentation of  $V$  which is either 0 or  $V$  by irreducibility. The image is a subrepresentation of  $W$  which is either 0 or  $W$  by irreducibility.  $\square$

**Lemma 12.19 (Uniqueness).** *There is a unique (up to isomorphism) irreducible representation of  $\mathfrak{sl}(3, \mathbb{C})$  with a given highest weight.*

**Proof.** Exercise!  $\square$

In fact, any highest weight occurs:

**Theorem 12.20** (Classification of irreducible representations of  $SU(3)$ ). *The irreducible representations of  $SU(3)$  are as follows. Take a weight  $\alpha = aL_1 - bL_3$ , with  $a$  and  $b$  nonnegative integers, and consider its reflections under the Weyl group. The convex hull of these is a hexagon  $X$ , possibly degenerate. There is a unique irreducible representation  $\Gamma_{a,b}$  whose weight diagram is supported on this hexagon. The weights which occur are the translates of  $\alpha$  by roots.*

Write  $W(\Lambda) \subset \mathfrak{h}_{\mathbb{Z}}^*$  for the  $\mathbb{Z}$ -lattice spanned by the weights  $\Lambda$  and call  $X \cap W(\Lambda)$  the hexagon of weights. The hexagon of weights is layered by concentric hexagons:

$$X \cap W(\Lambda) = X_0 \cap W(\Lambda) \supset X_1 \cap W(\Lambda) \supset \cdots \supset X_m \cap W(\Lambda)$$

The dimension of the weight spaces is constant on each layer and is given as follows. For a weight on the boundary of the hexagon the dimension is one. On each subsequent layer the dimension increases by one if the hexagon is nondegenerate and does not change if the hexagon is degenerate.

We will prove the first part of this theorem in the next chapter by first constructing many examples of representations. In particular we will prove existence of  $\Gamma_{a,b}$ . We will not prove the formula for the dimensions (“multiplicities”) of weight spaces, though we will work out some special cases. When we discuss representations in general we will give a much more general theorem called the Freudenthal multiplicity formula which will subsume this description and whose proof we may or may not have time for.

## Lecture 13: Representations of $SU(3)$ , II

### 13.1 The standard representation

The simplest representation of  $SU(3)$  is its standard representation on  $\mathbb{C}^3$ . Pick a basis  $e_1, e_2, e_3$  with respect to which we will write our matrices. Now  $e^{isH_{12}}$  and  $e^{itH_{23}}$  act as

$$\begin{pmatrix} e^{is} & 0 & 0 \\ 0 & e^{-is} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{it} & 0 \\ 0 & 0 & e^{-it} \end{pmatrix}$$

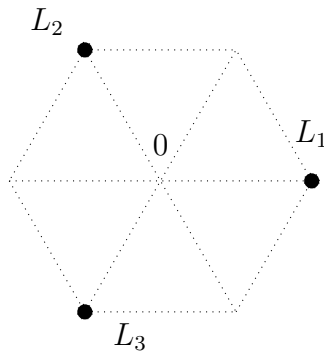
so we see that the simultaneous eigenspaces are just  $\mathbb{C}e_1$ ,  $\mathbb{C}e_2$  and  $\mathbb{C}e_3$ . The corresponding weights are

$$\alpha_1 = L_1, \alpha_2 = L_2, \alpha_3 = L_3$$

that is to say

$$e^{itH}e_k = e^{itL_k(H)}e_k$$

where we recall that  $L_k(H) = a_k$  for  $H = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}$ . Therefore the weight diagram looks like:



This consists of a single degenerate hexagon and corresponds to  $\Gamma_{1,0}$  in the notation of Theorem 12.20.

### 13.2 The dual standard representation

The dual of the space of column vectors  $\mathbb{C}^3$  is the space of row vectors (which we will write  $(\mathbb{C}^3)^*$  to distinguish it as a representation). The dual representation of  $SU(3)$  is just the right action of matrices on row vectors. In order for this to be a representation the action is

$$\rho(A) \begin{pmatrix} a & b & c \end{pmatrix} = \begin{pmatrix} a & b & c \end{pmatrix} A^{-1}, \quad A \in SU(3).$$

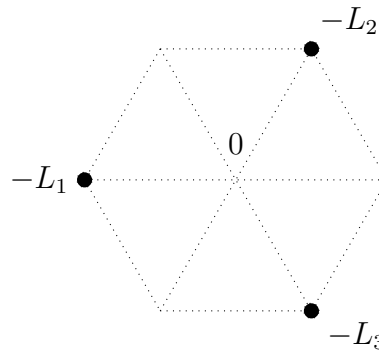
This implies that with respect to the dual basis  $e_1^*, e_2^*, e_3^*$  (defined by  $e_i^*(e_j) = \delta_{ij}$ ) the action of  $e^{isH_{12}}$  and  $e^{itH_{23}}$  is

$$\begin{pmatrix} e^{-is} & 0 & 0 \\ 0 & e^{is} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-it} & 0 \\ 0 & 0 & e^{it} \end{pmatrix}$$

with simultaneous eigenspaces  $\{\mathbb{C}e_k^*\}_{k=1}^3$ . The corresponding weights are

$$\alpha_1 = -L_1, \alpha_2 = -L_2, \alpha_3 = -L_3$$

and the weight diagram is



which corresponds to  $\Gamma_{0,1}$ . Note that this is not the same diagram as the standard representation, so they are not isomorphic (by contrast, the standard and standard dual representations of  $\mathfrak{sl}(2, \mathbb{C})$  are isomorphic because the weight diagrams agree).

*Remark 13.1.* It is not hard to see that given a representation  $\rho$ , the weight diagram of  $\rho^*$  can be obtained by applying the transformation

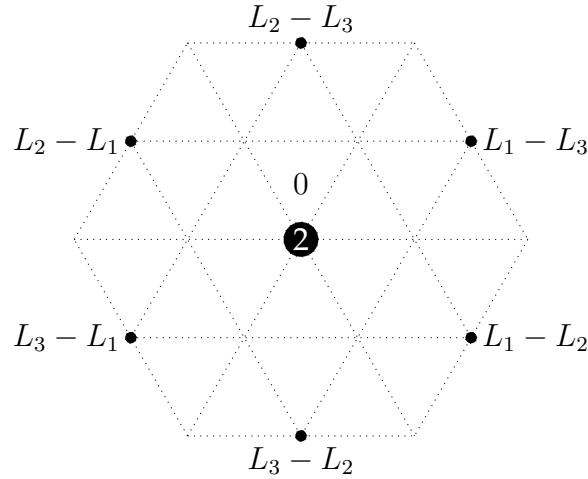
$$-1: \mathfrak{h}_{\mathbb{R}}^* \rightarrow \mathfrak{h}_{\mathbb{R}}^*$$

to the weight diagram of  $\rho$ .

### 13.3 The adjoint representation

We have already seen the weight diagram of this representation





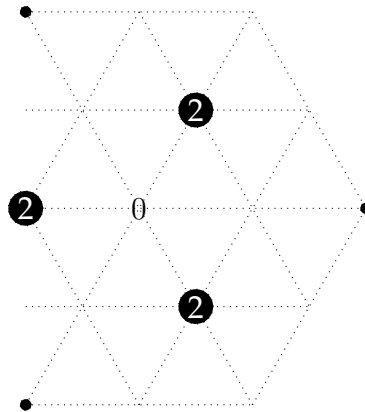
which corresponds to  $\Gamma_{1,1}$ . Note that the weight space with weight zero is  $\mathfrak{h}$  so it is two-dimensional (indicated by the numbered blob). All other weight spaces are one-dimensional. Since this diagram is symmetric under  $x \mapsto -x$ , the adjoint representation is self-dual.

### 13.4 The tensor square $(\mathbb{C}^3)^{\otimes 2}$

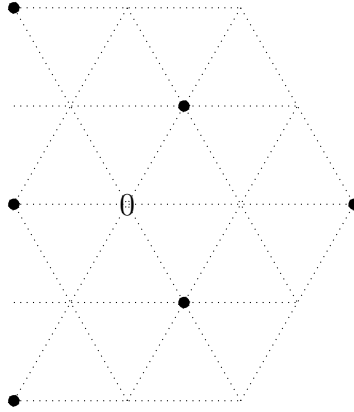
If we take the tensor square of the standard representation then the new weight spaces are

$$\mathbb{C}(e_i \otimes e_j) \text{ with weight } L_i + L_j$$

so the weight diagram is



where we have drawn numbered blobs to denote weight spaces of dimension two. We see that this is not irreducible because the dimensions of the weight spaces around the edge of the diagram are not all one. Indeed the tensor square of a nontrivial representation is never irreducible: it always contains the symmetric and exterior squares. Taking the highest weight vector furthest right in the diagram we generate an irreducible representation with weight diagram



or  $\Gamma_{2,0}$ , which leaves a copy of the weight diagram  $\Gamma_{0,1}$  of  $(\mathbb{C}^3)^*$  behind. In other words

$$\mathbb{C}^3 \otimes \mathbb{C}^3 = \Gamma_{2,0} \oplus \Gamma_{0,1}$$

A quick dimension check (or thinking a little harder) shows that

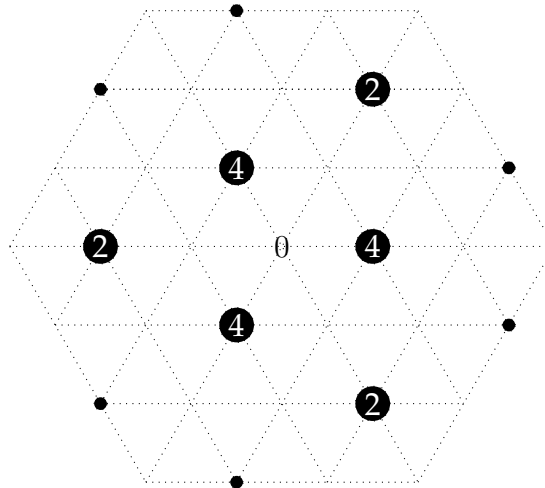
$$\text{Sym}^2 \mathbb{C}^3 = \Gamma_{2,0} \text{ and } \Lambda^2 \mathbb{C}^3 = \Gamma_{0,1} \cong (\mathbb{C}^3)^*$$

In general, it's easy to check that

$$\Gamma_{n,0} = \text{Sym}^n \Gamma_{1,0}, \quad \Gamma_{0,n} = \text{Sym}^n \Gamma_{0,1}.$$

### 13.5 Another tensor product

Take the tensor product  $\Gamma_{1,0} \otimes \Gamma_{1,1}$  of the standard and adjoint representations. The weight diagram is:



The labelled blobs denote weight spaces of dimension given by the label.

It is clear that this representation contains an irreducible summand  $\Gamma_{2,1}$ , generated by  $e_1 \otimes E_{13}$ . There are three ways of applying  $E_{21}$ ,  $E_{31}$  and  $E_{32}$  to  $e_1 \otimes E_{13}$  to obtain a vector

with weight  $L_1$ :

$$\rho(E_{21})\rho(E_{32})(e_1 \otimes E_{13}), \rho(E_{23})\rho(E_{21})(e_1 \otimes E_{13}), \rho(E_{31})(e_1 \otimes E_{13})$$

where  $\rho$  is the tensor representation. Therefore the weight-space of  $\Gamma_{2,1}$  with weight  $L_1$  is spanned by these three vectors. Clearly these have a linear dependence, since

$$\rho(E_{21})\rho(E_{32}) - \rho(E_{32})\rho(E_{21}) = \rho([E_{21}, E_{32}]) = -\rho(E_{31})$$

so the  $L_1$  weight-space of  $\Gamma_{2,1}$  is at most two-dimensional. We know it is at least one-dimensional. We have

$$E_{32}e_1 = 0, \text{ ad}_{E_{32}} E_{13} = -E_{12}$$

and

$$E_{21}e_1 = e_2, \text{ ad}_{E_{21}} E_{12} = -H_{12}$$

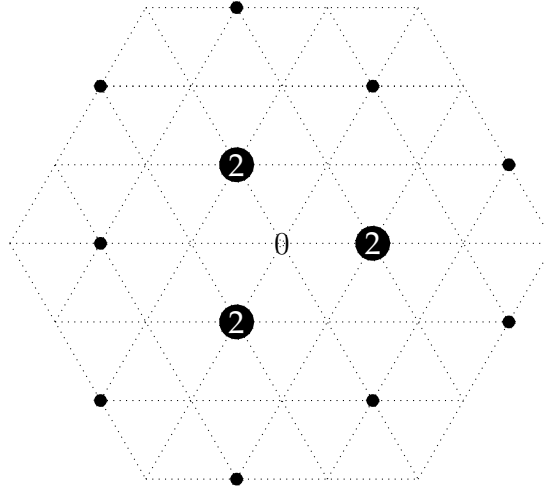
so

$$\rho(E_{21})\rho(E_{32})(e_1 \otimes E_{13}) = e_2 \otimes E_{12} - e_1 \otimes H_{12}.$$

Similarly

$$\rho(E_{32})\rho(E_{21})(e_1 \otimes E_{13}) = e_3 \otimes E_{13} - e_2 \otimes E_{12} - e_1 \otimes H_{23}.$$

Since these two vectors are linearly independent, we see that the  $L_1$  weight-space is two-dimensional. By Weyl symmetry, the  $L_2$  and  $L_3$  weight-spaces are also two-dimensional, so the weight diagram of  $\Gamma_{2,1}$  is:



Further decomposing  $\Gamma_{1,0} \otimes \Gamma_{1,1}$  we get

$$\Gamma_{1,0} \otimes \Gamma_{1,1} \cong \Gamma_{2,1} \oplus \Gamma_{0,2} \oplus \Gamma_{1,0}.$$

## Lecture 14: The eightfold way

The representation theory of  $SU(3)$  and other more complicated groups is important in particle physics. The goal of this more informal chapter is to explain why.

### 14.1 $SU(2)$ : The Yang-Mills model

The proton and the neutron look pretty similar, except that one has electric charge and the other is electrically neutral and the neutron is slightly heavier. When studying the strong nuclear interactions between these particles, which bind them into nuclei, the electric charge is not so important (how else could two protons get as close together as they do in the nucleus of an atom?) and to an extent one can forget about the difference between them. The theory is approximately symmetric under the interchange of proton and neutron.

In quantum mechanics one models particles by specifying the *quantum state* they inhabit. One can take complex linear combinations (superpositions) of states in the same way one can superimpose water waves in a bath by exciting the water in two different ways simultaneously (the only difference being that one usually takes real coefficients in the bath). This means that the state space of quantum mechanics is a complex vector space. The quantities one can measure (usually) only depend on the norm of vectors in this space, so one can rescale all vectors by a complex scalar. In other words, the physics is described by a complex line in a complex vector space, or a point of  $\mathbb{CP}^n$  (the projectivisation)<sup>3</sup>.

Imagine now that we have a state  $P$  corresponding to a particle being a proton and a state  $N$  corresponding to a particle being a neutron. These span a copy of  $\mathbb{C}^2$  which is the space of states and the projectivisation is  $\mathbb{CP}^1 = SU(2)/U(1)_\Delta$ . If your theory really is symmetric under interchanging protons and neutrons then (by linearity) it's also symmetric under the action of  $SU(2)$  which *rotates* amongst all possible linear combinations of the state  $P$  and the state  $N$ . This approximate "isospin symmetry" led Yang and Mills to postulate a model for strong interactions based on the group  $SU(2)$ . This original Yang-Mills model, while not a good description for the strong interaction, developed into a formalism called gauge theory which is now used to describe all the interactions whose quantum behaviour we understand. Gauge theory is possibly the most important subject in which Lie groups appear in a useful way outside pure mathematics.

### 14.2 $SU(3)$ : The quark model

In the early 1960s particle accelerators were producing many different particles and it became a challenge to classify them. Inspired by the idea of generalising  $SU(2)$ -isospin

---

<sup>3</sup>...though often the vector space of states is infinite-dimensional.

symmetry to  $SU(3)$ , Gell-Mann and Ne'eman were led to such a classification. The picture that gradually emerged (through further ideas of Gell-Mann and Zweig) was called the quark model.

The idea is as follows. There are three<sup>4</sup> fundamental particles called quarks: the up quark  $u$ , the down quark  $d$  and the strange quark  $s$ . These are called the different flavours of quarks. The space of flavour states of a single quark is therefore  $\mathbf{C}^3$  (or its projectivisation  $\mathbf{CP}^2$ , but let's ignore that subtlety). Quarks  $q$  (and their antiparticles  $\bar{q}$ ) combine in pairs  $q_1\bar{q}_2$  or triples  $q_1q_2q_3$  to form bigger particles like pions, kaons, protons and neutrons. A quark-antiquark pair is called a meson; a quark triple is called a baryon. To get the space of states of pair we take the tensor square of  $\mathbf{C}^3$ , to get it for a triple we take the tensor cube of  $\mathbf{C}^3$ . To get the state space of an antiparticle we take the dual representation.

We know that the tensor powers of a representation are not irreducible. This means that there are combinations of quarks which can never be transformed into one another, for instance  $uds - usd + dsu - dus + sud - sdu$  spans the one-dimensional subrepresentation  $\Lambda^3\mathbf{C}^3$  (the exterior cube) and so could never be made into any other baryon. Therefore we classify particles into the irreducible  $SU(3)$ -representations to which they belong.

Particles which belong to a one-dimensional irreducible representation in the above sense are said to form a singlet. Those in a two-dimensional representation form a doublet, and so on. The word multiplet is used for the collection of all particles in a given irreducible representation.

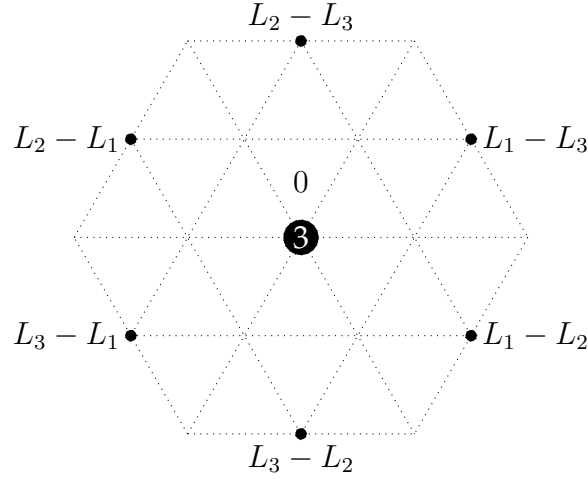
The eight-dimensional adjoint representation  $\mathfrak{su}(3)$  (or octet) was called the "eightfold way" by (an overexcited?) Gell-Mann. We will see that it plays a key role for both mesons and baryons.

### 14.2.1 Mesons

For mesons we start with the representation  $\mathbf{C}^3 \otimes (\mathbf{C}^3)^*$  and decompose it into irreducibles. The weight diagram for this representation is

---

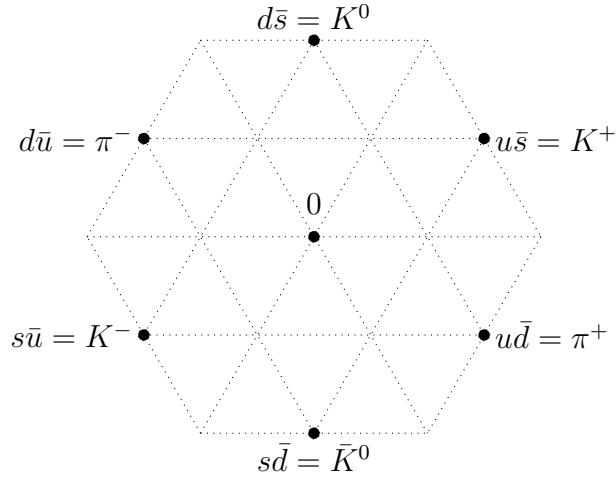
<sup>4</sup>...actually there are three others called top, bottom and charm, but for the  $SU(3)$  model we only need up, down and strange.



where the dimension of the zero weight-space is three. Therefore

$$\mathbf{C}^3 \otimes (\mathbf{C}^3)^* \cong \Gamma_{1,1} \oplus \mathbf{C}$$

where  $\mathbf{C}$  is the trivial one-dimensional representation (“singlet”) corresponding to the invariant subspace spanned by  $\eta_1 = \frac{u\bar{u} + d\bar{d} + s\bar{s}}{\sqrt{3}}$ . We call  $\Gamma_{1,1}$  the *meson octet*. If  $u, d, s$  are the 1, 2, 3-directions in  $\mathbf{C}^3$  then  $u\bar{d}$  corresponds to the matrix  $E_{12}$  (where we are thinking of  $V \otimes V^*$  as the space of matrices on  $V$ ),  $u\bar{s}$  corresponds to  $E_{13}$ , etc. so that the weight spaces correspond to quark combinations as follows



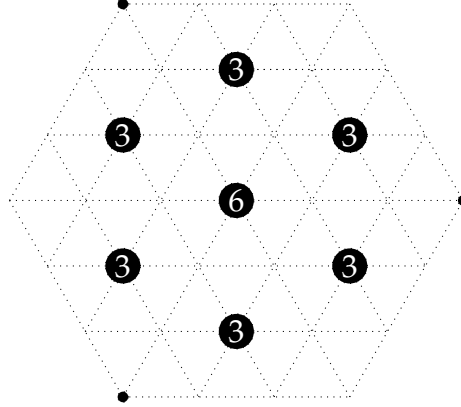
where we have written on the particle names which have these quark decompositions. Note that 0 corresponds to all combinations of  $u\bar{u}, d\bar{d}, s\bar{s}$  and this space is spanned by the  $\eta, \eta'$  and  $\pi^0$  mesons. Note that these latter particles are superpositions of different quark-antiquark states. Also, the observed particles  $\eta$  and  $\eta'$  do not actually live in the octet or the singlet: if  $\pi_0$  and  $\eta_8$  form a basis for the zero weight-space of the octet and  $\eta_1$  spans the singlet then  $\eta$  and  $\eta'$  are a linear combination of  $\eta_1$  and  $\eta_8$ . This happens because in reality the  $SU(3)$ -symmetry is only approximate and the weak nuclear interaction leads to “Cabbibo mixing” between down and strange quarks.

### 14.2.2 Baryons

For the tensor cube, we get

$$(\mathbf{C}^3)^{\otimes 3} \cong \Gamma_{3,0} \oplus \Gamma_{1,1} \oplus \Gamma_{1,1} \oplus \mathbf{C}$$

where  $\Gamma_{3,0}$  is the symmetric cube and  $\mathbf{C}$  is the exterior cube. This is clear from the weight diagram:



This should lead us to a classification of baryons into a decuplet:

$$\begin{aligned}
 &\bullet \Omega^- = sss \\
 &\bullet \Delta^0 = udd \\
 &\bullet \Sigma^{*-} = dds \quad \bullet \Delta^+ = uud \\
 &\bullet \Sigma^{*0} = uds \quad \bullet \Delta^{++} = uuu \\
 &\bullet \Xi^{*-} = dss \quad \bullet \Sigma^{*+} = ddd \\
 &\bullet \Xi^{*0} = uss \\
 &\bullet \Delta^- = ddd
 \end{aligned}$$

an octet (including  $p^+$  the proton and  $n^0$  the neutron):

$$\begin{aligned}
 &n^0 = udd \\
 &\Sigma^- = dds \bullet \quad \Sigma^0, \Lambda = uds \quad \bullet p^+ = uud \\
 &\Xi^- = dss \bullet \quad \bullet \Sigma^+ = uus \\
 &\Xi^0 = uss
 \end{aligned}$$

and a singlet:  $\Lambda_1 = uds$ .

*Remark 14.1.* Confusingly, it seems that there should be two octets. However this is an oversimplified picture. Really we should remember that each quark is a spin-1/2 particle, meaning that it has two states:  $\uparrow$  and  $\downarrow$ . These span a copy of  $\mathbb{C}^2$  in the space of quantum states. The true state space for a single quark is therefore  $\mathbb{C}^3 \otimes \mathbb{C}^2 \cong \mathbb{C}^6$  and the relevant group is really  $SU(6)$ . The tensor cube  $(\mathbb{C}^6)^{\otimes 3}$  decomposes into irreducible subrepresentations one of which is the 56-dimensional  $\text{Sym}^3 \mathbb{C}^6$ . Under the subgroup  $SU(3) \times SU(2) \subset SU(6)$  this decomposes into two pieces: a 40-dimensional piece and a 16-dimensional piece. The 40-dimensional piece is the tensor product of the decuplet with the standard 4-dimensional (spin-3/2) representation of  $SU(2)$ ; the 16-dimensional piece is the tensor product of the octet with the standard 2-dimensional (spin-1/2) representation of  $SU(2)$ . These are the actual baryons that we see: the baryons in the decuplet above all have spin-3/2 and those in the octet have spin-1/2.

*Remark 14.2.* The true wavefunction of (for example) the proton is therefore a linear combination of wavefunctions from the two  $SU(3)$ -octets combined with different spin wavefunctions. The naive picture we have of a proton being “two up quarks and one down quark” is misleadingly simple. Indeed the  $\Delta^+$  is also made up of two up quarks and one down quark. They are distinguished by the way they transform under  $SU(3)$  (or  $SU(6)$ ).

### 14.2.3 The point of all this

This is not just abstract nonsense:

- If one assigns charges  $2/3, -1/3, -1/3$  to  $u, d, s$  then they add correctly to give the charges of the composite particles.
- Similarly (to a good approximation) with masses  $2.4, 4.8, 104 \text{ MeV}/c^2$  and other quantum numbers like spin, baryon number and strangeness which are preserved by strong nuclear interactions.
- When the classification was originally made, the  $\Omega^-$  had not been observed. Its existence, mass, charge and other basic properties were predicted on the basis of this classification. It was discovered [?] in 1964.

The group  $SU(3)$  plays an even more central (but essentially different) role in the theory strong nuclear interactions as the gauge group of quantum chromodynamics. In other words, quarks have a charge-like property called colour which has three linearly independent states that can be acted on by  $SU(3)$ . This is not the same as the flavour symmetry we have been discussing!



## Lecture 15: Strategy for the general case

Much of the theory we have outlined in the foregoing lectures generalises to other groups. The most general setting in which it works is that of *compact semisimple Lie groups*.

### 15.1 Abelian Lie groups

**Proposition 15.1.** *Any compact abelian Lie group  $G$  is a torus.*

**Proof.** Since  $G$  is abelian the Lie bracket is identically zero on the Lie algebra, so the exponential map is a homomorphism (i.e. the usual law of logarithms holds) from  $\mathfrak{g} \cong \mathbb{R}^n$  to  $G$ . The kernel of this is a lattice in  $\mathbb{R}^n$ : that is, it is a subgroup where every point  $p \in \ker \exp$  is surrounded by a ball  $B_p$  such that  $B_p \cap \ker \exp = \{p\}$ . The balls are obtained by looking at the preimages of a small exponential chart. This lattice is finitely generated and hence isomorphic to  $\mathbb{Z}^n$  by the classification of finitely generated abelian groups (without torsion because  $\mathbb{R}^n$  has no torsion). If  $a_1, \dots, a_n$  are generators for this group then they form an  $\mathbb{R}$ -basis for  $\mathbb{R}^n$  (otherwise the quotient is not a compact group) and when written in that basis (so  $t \in \mathfrak{g}$  is  $t = \sum t_i a_i$ ) the group is just  $(\exp(t_1), \dots, \exp(t_n))$  which is clearly isomorphic to  $U(1)^n$ .  $\square$

### 15.2 Maximal tori

The key to the structure of  $SU(2)$ ,  $SU(3)$  and their representations was the existence of a torus  $H$  of diagonal matrices, isomorphic to  $U(1)$  and  $U(1) \times U(1)$  respectively. We decomposed representations of  $SU(2)$  and  $SU(3)$  into weight spaces under the action of this torus.

**Definition 15.2.** A torus in a Lie group is a subgroup isomorphic to  $U(1)^n$  for some  $n \geq 1$ . A maximal torus is a torus which is not contained in any other torus.

**Proposition 15.3.** *If  $G$  is a compact group then there exists a torus in  $G$ . Moreover there exists a maximal torus.*

**Proof.** Let  $X \in \mathfrak{g}$  and take the one-parameter subgroup  $\exp(tX)$ . The image of this is an abelian subgroup of  $G$ . Its closure (topological!) is again an abelian matrix group (one of the first things we proved about matrix groups) and because a closed subset of a compact set is compact, it must be a torus. Its Lie algebra contains  $X$  and is therefore at least one-dimensional.

Now consider the partially-ordered set of abelian subalgebras of  $\mathfrak{g}$ . Since subalgebras are in particular subspaces there are maximal elements in the poset (as dimension is finite). We claim that the exponential image of a maximal abelian subalgebra  $\mathfrak{t}$  is a maximal torus. To see this, consider the Lie group  $T = \overline{\exp(\mathfrak{t})}$ . It is a torus and its Lie algebra is an abelian subalgebra containing  $\mathfrak{t}$ , hence its Lie algebra equals  $\mathfrak{t}$ . If  $T \subset T'$  for some abelian group  $T'$  then  $\mathfrak{t}'$  is an abelian subalgebra containing  $\mathfrak{t}$  and hence  $\mathfrak{t} = \mathfrak{t}'$  and the two subgroups are equal.  $\square$

Henceforth we will *choose* a particular maximal torus. Changing this choice won't affect our analysis because:

**Theorem 15.4.** *Any two maximal tori  $T, T'$  in a compact finite-dimensional Lie group are conjugate to one another, that is  $T = gT'g^{-1}$  for some  $g \in G$ .*

**Proof.** The proof is an application of the Lefschetz fixed point formula from algebraic topology and is beyond the scope of this course. For those who are interested, see [?, 4.21–4.23].  $\square$

## 15.3 The strategy

Given a compact group  $G$  and a complex representation  $R: G \rightarrow GL(V)$ :

- Pick a maximal torus  $T \subset G$  and restrict the representation to  $T$  to obtain a representation  $R|_T$  of  $T$ .
- Decompose  $R|_T: T \rightarrow GL(V)$  into weight spaces

$$V = \bigoplus_{\lambda \in \mathfrak{t}_\mathbb{Z}^*} V_\lambda.$$

- Let  $\rho: \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  be the associated Lie algebra representation. We now know that the elements  $X \in \mathfrak{t}$  act by

$$\rho(X)v = \lambda(X)v$$

on  $v \in V_\lambda$ . The final step is to figure out how the other elements of  $\mathfrak{g}$  (not in  $\mathfrak{t}$ ) act on the  $V_\lambda$ s. This will involve careful analysis of the adjoint representation.

## 15.4 Killing form

For  $SU(3)$ , we drew a picture of the roots (a hexagon in  $\mathbb{R}^2$ ) and were able to do geometry in  $\mathfrak{h}_\mathbb{R}^*$ . At the time this involved an arbitrary choice of metric by declaring  $\mathfrak{h}_\mathbb{R}^*$  to be a quotient of the inner product space  $\mathbb{R}\check{L}_1 \oplus \mathbb{R}\check{L}_2 \oplus \mathbb{R}\check{L}_3$  with a basis given by  $\{\check{L}_i\}_{i=1}^3$ . For the general story we need a canonical prescription.

**Lemma 15.5.** *There is a natural symmetric bilinear form  $K: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$  which is invariant in the sense that*

$$K([X, Y], Z) = K(Y, [X, Z]).$$

**Proof.** This form is called the Killing form and is defined as

$$K(X, Y) = \text{Tr}(\text{ad}_X \text{ad}_Y).$$

Here  $\text{ad}_X$  and  $\text{ad}_Y$  are matrices, we take their product and then take the trace. Invariance follows from the Jacobi identity.  $\square$

We will assume in what follows that  $-K$  is positive-definite. Of course this is not true for all Lie algebras and we must further restrict our attention. The resulting class of Lie algebras are the Lie algebras of *compact semisimple Lie groups*.

**Definition 15.6.** A Lie algebra over  $\mathbb{C}$  is called semisimple if its Killing form is nondegenerate (this is not the right definition for Lie algebras over fields of characteristic  $p \neq 0$ ). A Lie group  $G$  is called semisimple if  $\mathfrak{g}_\mathbb{C}$  is semisimple.

**Lemma 15.7.** *The Killing form of a compact semisimple Lie group is negative definite.*

**Proof.** For a compact group we know that any representation is unitary for some choice of Hermitian inner product. This means that the adjoint representation can be thought of as a map  $\text{ad}: \mathfrak{g} \rightarrow \mathfrak{u}(n)$  (where  $n = \dim G$ ). Therefore  $\text{ad}_X$  is a skew-Hermitian matrix. Now

$$\text{Tr}(\text{ad}_X \text{ad}_X) = -\text{Tr}(\text{ad}_X \text{ad}_X^\dagger) = -\sum_{i,j} |x_{ij}|^2 \leq 0$$

where  $x_{ij}$  are the entries of  $\text{ad}_X$ . Equality occurs if and only if  $\text{ad}_X = 0$ .

We are also assuming semisimplicity, i.e. nondegeneracy of the Killing form. If  $\text{ad}_X = 0$  then  $K(X, Y) = 0$  for all  $Y$ , so this cannot occur. Therefore the Killing form is negative definite.  $\square$

*Remark 15.8.* Henceforth we will work only with compact semisimple Lie groups so that we can assume the existence of a maximal torus and the negative definiteness of the Killing form.

**Example 15.9.** As an example, let's consider the Killing form on  $\mathfrak{su}(2)$ . With respect to the basis  $\sigma_1, \sigma_2, \sigma_3$  of  $\mathfrak{su}(2)$ , the adjoint action of  $\sigma_1$  is

$$\text{ad}_{\sigma_1} \sigma_1 = 0, \text{ad}_{\sigma_1} \sigma_2 = 2\sigma_3, \text{ad}_{\sigma_1} \sigma_3 = -2\sigma_2$$

so  $\text{ad}_{\sigma_1}^2$  is the diagonal matrix  $\text{diag}(0, -4, -4)$  with trace  $-8$ . Similarly the trace of  $\text{ad}_{\sigma_k}^2$  is  $-8$  for  $k = 2, 3$ . We can also check that  $\sigma_k$  are orthogonal: for example  $\text{ad}_{\sigma_1} \text{ad}_{\sigma_3}$  is a matrix with a four in the bottom right corner, so its trace vanishes. When we look at  $i\mathfrak{su}(2) = \mathfrak{sl}(2, \mathbb{R})$ , we get a minus sign. It is convenient to use the form  $K/8$  so that the standard basis is orthonormal.

An inner product on  $\mathfrak{g}$  gives a vector-space isomorphism<sup>5</sup>  $\flat: \mathfrak{g} \rightarrow \mathfrak{g}^*$ :

$$\flat(v)(w) = \langle v, w \rangle.$$

So we can also think of the inner product as defined on the dual space. We will now check that (up to an overall rescaling) the inner product  $-K$  is the inner product we used when analysing  $\mathfrak{h}_{\mathbb{R}}^*$  for  $\mathfrak{su}(3)$ .

**Example 15.10.** Let's use the basis given by  $E_{ij}, i < j$  and  $H_{ij} = E_{ii} - E_{jj}, i < j$ . It's easy to compute the (eight-by-eight) matrix of  $\text{ad}_{H_{12}}$  with respect to the basis  $H_{12}, H_{23}, S_{12}, A_{12}, S_{13}, A_{13}, S_{23}, A_{23}$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}$$

whose square is  $\text{diag}(0, 0, -4, -4, -1, -1, -1, -1)$  with trace  $K(H_{12}, H_{12}) = -12$ .

---

<sup>5</sup>The inverse of  $\flat$  is  $\sharp$ .

Similarly  $K(H_{23}, H_{23}) = -12$  and  $K(H_{12}, H_{23}) = 6$ . Let's use the inner product  $-K/12$  (or work on  $\mathfrak{isu}(3) = \mathfrak{sl}(3, \mathbb{R})$ ). With respect to this,  $I = H_{12}, J = \frac{2}{\sqrt{3}}(H_{23} + \frac{1}{2}H_{12})$  form an orthonormal basis so  $I^\flat, J^\flat$  form a basis for the dual space. We would like to use this to compute the inner products of  $L_1, L_2$  and  $L_3$ . Recall that  $\langle H_{jk}^\flat, L_i \rangle = L_1(H_{jk}) = \delta_{ij} - \delta_{jk}$ . Therefore

$$\begin{aligned} L_1 \cdot H_{12}^\flat &= 1 & L_1 \cdot H_{23}^\flat &= 0 \\ L_2 \cdot H_{12}^\flat &= -1 & L_2 \cdot H_{23}^\flat &= 1 \end{aligned}$$

so

$$L_1 = I^\flat + \frac{1}{\sqrt{3}}J^\flat, \quad L_2 = -I^\flat + \frac{1}{\sqrt{3}}J^\flat.$$

Therefore  $L_1 \cdot L_1 = 4/3 = L_2 \cdot L_2$  and  $L_1 \cdot L_2 = -2/3$ . Rescaling again by  $3/2$  (so we are now using  $-K/8$ ) we get the dot products

$$L_i \cdot L_j = -1, \quad i \neq j \quad L_i \cdot L_i = 2$$

as before.

## 15.5 Geometry of roots

Now, for a compact semisimple Lie group  $G$  (with associated Lie algebra  $\mathfrak{g}$ ) we have (a) a maximal torus  $H$  and the accompanying root decomposition

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_{\alpha}$$

of the adjoint representation, and (b) the definite Killing form on  $\mathfrak{h}_{\mathbb{R}}^*$ . Given this we can start proving things about root systems of compact semisimple Lie groups.

**Lemma 15.11.** *If  $X \in \mathfrak{g}_{\alpha}$  and  $Y \in \mathfrak{g}_{\beta}$  then*

$$[X, Y] \in \mathfrak{g}_{\alpha+\beta}.$$

**Proof.** We proved this before for  $\mathfrak{su}(3)$  and the proof goes through unchanged.  $\square$

**Corollary 15.12.** *If  $X \in \mathfrak{g}_{\alpha}, Y \in \mathfrak{g}_{\beta}$  and  $\alpha + \beta \neq 0$  then  $K(X, Y) = 0$ .*

**Proof.** By the lemma, the matrix  $(\text{ad}_X \text{ad}_Y)^n$  applied to a vector  $Z \in \mathfrak{g}_{\gamma}$  gives

$$(\text{ad}_X \text{ad}_Y)^n Z \in \mathfrak{g}_{\gamma+n(\alpha+\beta)}.$$

Since  $\alpha + \beta \neq 0$  and since there are only finitely many nonzero root spaces, this is eventually zero (for large enough  $n$ ). This implies that  $\text{ad}_X \text{ad}_Y$  is nilpotent: when raised to some power it vanishes. The trace of a nilpotent matrix over  $\mathbb{C}$  is zero: this is easiest to see in Jordan normal form, where nilpotence means that the diagonal entries are all zero.  $\square$

**Remark 15.13.** In other words, the only way for two root vectors to pair nontrivially under the Killing form is for them to lie in opposite root spaces  $\mathfrak{g}_{\alpha}$  and  $\mathfrak{g}_{-\alpha}$ .

Semisimplicity means nondegeneracy of the Killing form, which implies that

**Corollary 15.14.** *Suppose that  $G$  is a compact semisimple Lie group and  $R$  is the root system of its Lie algebra. If  $\alpha \in R$  then  $-\alpha \in R$ .*

**Lemma 15.15.** *If  $X \in \mathfrak{g}_\alpha$  and  $Y \in \mathfrak{g}_{-\alpha}$  then*

$$[X, Y] = K(X, Y)\alpha^\sharp.$$

**Proof.** Translating into a Killing-dual statement, we need to prove that

$$K(H, [X, Y]) = \alpha(H)K(X, Y)$$

for all  $H \in \mathfrak{h}_\mathbb{R}$ . Using invariance of the Killing form we get

$$K(H, [X, Y]) = K([H, X], Y) = \alpha(H)K(X, Y)$$

as required. □

## 15.6 $\mathfrak{sl}(2, \mathbb{C})$ -subalgebras

We can now find the all-important  $\mathfrak{sl}(2, \mathbb{C})$ -subalgebra associated to a root  $\alpha$ . Pick  $X_\alpha \in \mathfrak{g}_\alpha$ . By Remark 15.13 and nondegeneracy of the Killing form, we know there exists  $Y_\alpha \in \mathfrak{g}_\alpha$  such that  $K(X_\alpha, Y_\alpha) \neq 0$ . Lemma 15.15 tells us that

$$H_\alpha := [X_\alpha, Y_\alpha] = K(X_\alpha, Y_\alpha)\alpha^\sharp.$$

Rescaling, we can pick  $X_\alpha$  and  $Y_\alpha$  so that  $K(X_\alpha, Y_\alpha) = 2/|\alpha|^2$  (which makes sense because we are assuming positive-definiteness of the Killing form, so  $|\alpha|^2 \neq 0$ ). Then

$$\alpha(H_\alpha) = 2$$

and it is easy to check that

**Lemma 15.16.**

$$\mathfrak{s}_\alpha = \mathbb{C}X_\alpha \oplus \mathbb{C}Y_\alpha \oplus \mathbb{C}H_\alpha \subset \mathfrak{g}_\mathbb{C}$$

*is a Lie subalgebra isomorphic to  $\mathfrak{sl}(2, \mathbb{C})$ .*

It remains to show that  $\mathfrak{g}_\alpha$  is one-dimensional, so that there is really no choice in constructing this  $\mathfrak{sl}(2, \mathbb{C})$ -subalgebra.

**Lemma 15.17.** *The subspace*

$$V = \mathbb{C}H_\alpha \oplus \bigoplus_{k \in \mathbb{Z} \setminus \{0\}} \mathfrak{g}_{k\alpha} \subset \mathfrak{g}$$

*is an irreducible representation of  $\mathfrak{s}_\alpha$ . In particular,  $\mathfrak{g}_{\pm\alpha}$  are one-dimensional.*

**Proof.** We must check that it is a representation, i.e. that it is preserved by the adjoint action of  $\mathfrak{s}_\alpha$ . Certainly

$$\mathrm{ad}_{X_\alpha} \mathfrak{g}_{k\alpha} \subset \mathfrak{g}_{(k+1)\alpha}, \quad \mathrm{ad}_{Y_\alpha} \mathfrak{g}_{-k\alpha} \subset \mathfrak{g}_{-(k+1)\alpha}$$

when  $k \neq -1$ . Moreover, Lemma 15.15 implies that

$$\mathrm{ad}_{X_\alpha} \mathfrak{g}_{-\alpha} \subset \mathbf{C}H_\alpha, \quad \mathrm{ad}_{Y_\alpha} \mathfrak{g}_\alpha \subset \mathbf{C}H_\alpha.$$

We also have

$$v \in \mathfrak{g}_{k\alpha} \Rightarrow \mathrm{ad}_{H_\alpha} v = k\alpha(H_\alpha)v = 2kv$$

so  $V$  is preserved by the action of the generators  $X_\alpha, Y_\alpha, H_\alpha$  of  $\mathfrak{s}_\alpha$ . Moreover we see that the weight decomposition of this  $\mathfrak{sl}(2, \mathbf{C})$ -representation has  $\mathfrak{g}_{k\alpha}$  as the  $2k$  weight-space and  $\mathbf{C}H_\alpha$  as the zero weight-space.

Now  $V$  decomposes into irreducible subrepresentations, each with even weight. Each subrepresentation has a zero weight summand. The direct sum of these zero weight summands is the zero weight-space of  $V$ , which is one-dimensional. Hence there is only one irreducible summand and  $V$  is irreducible.

From the classification of irreducible representations of  $\mathfrak{sl}(2, \mathbf{C})$ , this immediately implies that  $\mathfrak{g}_\alpha$  and  $\mathfrak{g}_{-\alpha}$  are one-dimensional.  $\square$

We can push this a little further.

**Corollary 15.18.** *On the line connecting  $-\alpha$  to  $\alpha$  there are no roots other than  $\pm\alpha$  and 0.*

**Proof.** We know that  $\mathfrak{s}_\alpha$  preserves the direct sum of root spaces along this line:  $X_\alpha$  and  $Y_\alpha$  translate in either direction. If  $\beta = b\alpha$  is a root on this line then the weight of  $H_\alpha$  acting on  $\mathfrak{g}_\beta$  is  $2K(\alpha, \beta)/K(\alpha, \alpha) = 2b$  and this must be an integer, so  $b$  is a half-integer. By reversing the roles of  $\alpha$  and  $\beta$  we also see that  $1/b$  is a half-integer, so  $b \in \{\pm 1, \pm 2, \pm 1/2\}$ . Without loss of generality (swapping  $\alpha$  and  $\beta$  and changing  $\beta$  to  $-\beta$ ) assume that  $b \in \{1, 2\}$ . We want to rule out the case  $b = 2$ .

The previous lemma showed that  $V = \mathbf{C}H_\alpha \oplus \bigoplus_{k \in \mathbf{Z} \setminus \{0\}} \mathfrak{g}_{k\alpha}$  was an irreducible  $\mathfrak{s}_\alpha$ -representation. If  $b = 2$  were possible then this representation would be

$$\mathfrak{g}_{-2\alpha} \oplus \mathfrak{g}_\alpha \oplus \mathbf{C}H_\alpha \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{2\alpha}.$$

However, we know that  $X_\alpha$  generates  $\mathfrak{g}_\alpha$  and hence  $\mathrm{ad}_{X_\alpha} \mathfrak{g}_\alpha = 0$ . In an irreducible representation, the weight space  $\mathfrak{g}_{2\alpha}$  would be generated by  $\mathrm{ad}_{X_\alpha}$  applied to the  $\alpha$ -weight space. Since this vanishes, we see that  $b \neq 0$ . This proves the corollary.  $\square$

We can further decompose  $\mathfrak{g}$  as a representation of  $\mathfrak{s}_\alpha$ :

**Lemma 15.19.** *Suppose that  $\beta$  is a root which is linearly independent from  $\alpha$ . Then*

$$\bigoplus_{k \in \mathbf{Z}} \mathfrak{g}_{\beta+k\alpha}$$

*is an irreducible representation of  $\mathfrak{s}_\alpha$ .*

**Proof.** Certainly this subspace is preserved by  $\mathfrak{s}_\alpha$ . It decomposes into weight spaces  $\mathfrak{g}_{\beta+k\alpha}$  with weight  $\beta(H_\alpha) + 2k$ . Each of these weight spaces is one-dimensional (as it is a nonzero root space). Therefore the representation is irreducible.  $\square$

There are two important facts about irreducible representations of  $\mathfrak{sl}(2, \mathbf{C})$  which give us yet more structure:

- The weights are integers.

Since  $H_\alpha$  acts with weight  $\beta(H_\alpha) = 2K(\alpha, \beta)/K(\alpha, \alpha)$  on  $\mathfrak{g}_\beta$  this means that:

**Corollary 15.20.** *For any nonzero roots  $\alpha, \beta$  we have*

$$\frac{2K(\alpha, \beta)}{K(\alpha, \alpha)} \in \mathbf{Z}.$$

- The weights are distributed symmetrically about the origin.

This means that

**Corollary 15.21.** *The reflection operator*

$$s_\alpha(\beta) = \beta - \frac{2K(\alpha, \beta)}{K(\alpha, \alpha)}\alpha$$

*which reflects in the hyperplane*

$$\Omega_\alpha = \{\beta : K(\alpha, \beta) = 0\}$$

*preserves the set of roots in  $\mathfrak{h}_\mathbf{R}^*$ . The group generated by these reflection operators is called the Weyl group.*

## 15.7 Summary

**Theorem 15.22.** *If  $G$  is a compact semisimple Lie group with Lie algebra  $\mathfrak{g}$  then:*

- *The Killing form*

$$K(X, Y) = \text{Tr}(\text{ad}_X \text{ad}_Y)$$

*on  $\mathfrak{g}$  is negative-definite (positive-definite on  $i\mathfrak{g}$ ).*

- *There exists a maximal torus  $H \subset G$  (unique up to conjugation) with Lie algebra  $\mathfrak{h} \subset \mathfrak{g}$ .*
- *Under the adjoint action of the maximal torus, the complexified Lie algebra  $\mathfrak{g}_\mathbf{C}$  decomposes into root spaces*

$$\mathfrak{g}_\mathbf{C} = \mathfrak{h}_\mathbf{C} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha$$

*where  $\mathfrak{h}_\mathbf{C}$  is the zero root-space.*

- *If  $\alpha \in R$  then  $-\alpha \in R$ . Moreover the only roots on the line through  $\pm\alpha$  are  $\pm\alpha$  and 0.*
- *Each root-space  $\mathfrak{g}_\alpha$  is one-dimensional.*
- *For each pair of roots,  $\alpha, \beta$  we have*

$$\frac{2K(\alpha, \beta)}{K(\alpha, \alpha)} \in \mathbf{Z}.$$

- The reflection operator

$$s_\alpha(\beta) = \beta - \frac{2K(\alpha, \beta)}{K(\alpha, \alpha)}\alpha$$

which reflects in the hyperplane

$$\Omega_\alpha = \{\beta : K(\alpha, \beta) = 0\}$$

preserves  $R \subset \mathfrak{h}_\mathbf{R}^*$ .

- The group of reflections in  $\Omega_\alpha$  is called the Weyl group.
- If  $\beta$  is a root which is linearly independent from  $\alpha$ . Then

$$V_\beta = \bigoplus_{k \in \mathbf{Z}} \mathfrak{g}_{\beta+k\alpha}$$

is an irreducible representation of  $\mathfrak{s}_\alpha$ .

- As a representation of  $\mathfrak{s}_\alpha$ ,  $\mathfrak{g}_\mathbf{C}$  decomposes into irreducibles

$$\mathfrak{s}_\alpha \oplus \bigoplus_{\beta} V_\beta \oplus \mathfrak{h}_\mathbf{C}/CH_\alpha.$$

## 15.8 Irreducible representations

We can analyse representations of compact semisimple Lie groups as we analysed representations of  $SU(2)$  or  $SU(3)$ . We have complete reducibility into irreducible subrepresentations. We can take a weight space decomposition with respect to a maximal torus (whose complexified Lie algebra is  $\mathfrak{t}$ ) and we get a collection of vertices in the weight lattice

$$\mathfrak{t}_\mathbf{Z}^* = \{f \in \mathfrak{t}^* : f(v) \in 2\pi\mathbf{Z} \text{ for all } v \in \ker \exp\}.$$

A highest weight vector (with respect to an irrational linear function for the weight lattice) generates an irreducible subrepresentation by acting using negative roots. For each weight  $\alpha$  there is a unique irreducible representation containing a highest weight vector with weight  $\alpha$  and the weight diagram for this irreducible representation:

- is symmetric under the Weyl group
- is obtained by:
  - reflecting the highest weight under all the elements of the Weyl group,
  - taking the convex hull of the images,
  - and looking at all lattice points in this convex hull which can be obtained from the highest weight by translating along a root.

The only conundrum is what multiplicities to put on the weights. There are a number of beautiful approaches (Weyl's character formula, the Littlemann path model, Freudenthal's multiplicity formula) and we will use one of them.



## 15.9 Multiplicities

**Theorem 15.23** (Freudenthal's multiplicity formula). *Suppose that  $G$  is a Lie group and  $G \rightarrow GL(V)$  is a finite-dimensional complex irreducible representation of highest weight  $\lambda$ . Let  $R^+$  denote the set of positive roots of  $G$ . Then the multiplicity of the weight  $\mu$  is*

$$\dim V_\mu = \frac{2 \sum_{\alpha \in R^+} \sum_{j \geq 1} \langle \mu + j\alpha, \alpha \rangle \dim V_{\mu+j\alpha}}{\|\lambda + \rho\|^2 - \|\mu + \rho\|^2}$$

where  $\rho = \frac{1}{2} \sum_{\alpha \in R^+} \alpha$ ,  $\|X\|^2 = \langle X, X \rangle$  and  $\langle, \rangle$  means the Killing form.

**Proof.** The proof involves studying the action of a Casimir operator in the Lie algebra  $\mathfrak{g}$  - this is an element  $C$  which has the property that, for any representation  $\rho$ ,  $\rho(C)$  commutes with  $\rho(X)$  for all  $X \in \mathfrak{g}$ . We met the Casimir for  $\mathfrak{sl}(2, \mathbb{C})$  on Sheet 6. The proof is beyond the timescale of this course and can be found in [?, Chapter 25.1].  $\square$