PMATH 763: INTRODUCTION TO LIE GROUPS AND LIE ALGEBRAS NOTES

HEESUNG YANG

1. January 05: Introduction

1.1. *Outline*

(1) Weeks 1–3: Matrix Lie groups and Lie algebras (relations between them "Lie correspondence")
(2) Weeks 4–9/10: Representation theory of Lie groups and Lie algebras, both abstractly and via explicit examples (like $\mathrm{SL}(2,\mathbb{C}), \mathrm{SL}(3,\mathbb{C})$)
(3) If time permits: additional topics such as Clifford algebras, exceptional Lie groups, Spin groups, G_2 (octonions), F_4 (Jordan algebras)
(4) Prerequisites: group theory, linear algebra, point-set topology, some analysis. We will focus more on the algebraic aspects of Lie groups and Lie algebras, so not much analysis is needed. However, group theory and linear algebra are very important. Manifold theory is *not* a prerequisite.

1.2. *Moving on to the course itself*

Definition 1.1. A *Lie group* is a group that is also a smooth manifold (i.e. locally Euclidean).

Remark 1.1. We will focus on *matrix Lie groups* (or sometimes called *matrix groups*). All matrix Lie groups are Lie groups. Most, *but not all*, Lie groups are matrix Lie groups. In some sense, all the important ones are. See Appendix C of the textbook for examples of Lie groups that are not matrix groups.

We will study the correspondence between G and $\mathfrak{g}$.

Date: 1 April 2015.

Definition 1.2. We say $\mathrm{GL}(n, \mathbb{C})$ is the set of all invertible $n \times n$ matrices over $\mathbb{C}$ called the *general linear group over* $\mathbb{C}$. We can define a similar notion over $\mathbb{R}$ (write $\mathrm{GL}(n, \mathbb{R})$). We denote $\mathrm{M}_n(\mathbb{C})$ the set of all $n \times n$ matrices over $\mathbb{C}$, and $\mathrm{M}_n(\mathbb{R})$ the set of all $n \times n$ matrices over $\mathbb{R}$.

Remark 1.2. Both are groups, and $\mathrm{GL}(n, \mathbb{R})$ is a subgroup of $\mathrm{GL}(n, \mathbb{C})$. $\mathrm{M}_n(\mathbb{C})$ resp. $\mathrm{M}_n(\mathbb{R})$ are not groups but are algebras (over $\mathbb{C}$ resp. $\mathbb{R}$). Recall that, as vector spaces, we have

$$\mathrm{M}_n(\mathbb{R}) \cong \mathbb{R}^{n \times n} \text{ and } \mathrm{M}_n(\mathbb{C}) \cong \mathbb{C}^{n \times n} \cong \mathbb{R}^{2n^2}.$$

($\mathrm{M}_n(\mathbb{C}) \cong \mathbb{C}^{n \times n}$ as vector spaces over the field $\mathbb{C}$ and $\mathbb{C}^{n \times n} \cong \mathbb{R}^{2n^2}$ as vector spaces over $\mathbb{R}$.)

Hence $\mathrm{M}_n(\mathbb{R})$ and $\mathrm{M}_n(\mathbb{C})$ inherit the natural Euclidean norm

$$A \in \mathrm{M}_n(\mathbb{F}), |A|^2 = \sum_{i,j} A_{ij}^2,$$

inducing the usual topology. There exists other norms, such as operator norms. So a sequence $A_m \in \mathrm{M}_n(\mathbb{F})$ converges to $A \in \mathrm{M}_n(\mathbb{F})$ as $m \to \infty$ if and only if $(A_m)_{ij} \to A_{ij}$.

Definition 1.3. A *matrix Lie group* (or just a *matrix group*) G is a subgroup of $\mathrm{GL}(n, \mathbb{C})$ that is *closed* in $\mathrm{GL}(n, \mathbb{C})$ in the topological sense. That is, if $A_m \in G$ for all m and $A_m \to A \in \mathrm{M}_n(\mathbb{C})$, then either $A \notin \mathrm{GL}(n, \mathbb{C})$ or $A \in G$ (equivalently, a sequence in G that converges in $\mathrm{GL}(n, \mathbb{C})$ has a limit in G).

Remark 1.3. Most interesting and important subgroups of $\mathrm{GL}(n, \mathbb{C})$ are *closed* in $\mathrm{GL}(n, \mathbb{C})$, hence are matrix Lie groups. But not all are (see Assignment #1).

Example 1.4. In "some sense", this list is "almost" exhaustive (we won't make this hand-wavy statement precise now):

(0) $\mathrm{GL}(n, \mathbb{C})$ trivially. $\mathrm{GL}(n, \mathbb{R})$ is a matrix Lie group also, since the sequence in $\mathbb{R}$ converging in $\mathbb{C}$ has a limit in $\mathbb{R}$.

(1) $\mathrm{SL}(n, \mathbb{C}) = \{A \in \mathrm{GL}(n, \mathbb{C}) : \det(A) = 1\}$ and $\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) : \det(A) = 1\}$ are subgroups because $\det(AB) = \det(A)\det(B)$ and $\det(A^{-1}) = (\det(A))^{-1}$. We know they are closed since $\det : \mathrm{M}_n(\mathbb{C}) \to \mathbb{C}$ is continuous: $\lim \det(A_m) = \det(\lim A_m) = 1$.

(2) Consider the following linear form on $\mathbb{R}^n$: let $0 \leq p, q \leq n$ with $p + q = n$. Define the bilinear form

$$\langle \_, \_ \rangle_{p,q} = \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$

by

$$\langle x, y \rangle_{p,q} = \sum_{k=1}^{p} x_k y_k - \sum_{k=p+1}^{n} x_k y_k.$$

If $p = n, q = 0$ then $\langle \_, \_ \rangle_{n,0}$ is the standard Euclidean inner product. We call $\langle \_, \_ \rangle_{p,q}$ the *pseudo-Euclidean inner product of signature* p, q.

Some properties of $\langle \_, \_ \rangle_{p,q}$:

- $\langle x, y \rangle_{p,q} = \langle y, x \rangle_{p,q}$
- bilinear
- non-degenerate, ie. $\langle x, y \rangle_{p,q} = 0$ for all y iff $x = 0$.
- *not positive-definite*, unless $q = 0$.

- $\langle x,y\rangle_{p,q} = x^T I_{p,q} y$, where

$$I_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$$

Definition 1.5. The *orthogonal group* is defined to be

$$\mathrm{O}(p,q) = \{A \in \mathrm{GL}(n,\mathbb{R}) : \langle Ax, Ay\rangle_{p,q} = \langle x,y\rangle_{p,q}\, \forall x,y \in \mathbb{R}^n\},$$

i.e., the group of automorphisms of $\mathbb{R}^n$ preserving $\langle \_,\_ \rangle_{p,q}$. $\mathrm{O}(p,q)$ is also called the (p,q)-*orthogoral group*.

Remark 1.4. $\mathrm{O}(p,q)$ is a subgroup of $\mathrm{GL}(n,\mathbb{R})$. Also, $\mathrm{O}(p,q)$ is closed, since

$$\langle Ax, Ay\rangle_{p,q} = (Ax)^T I_{p,q} Ay = x^T A^T I_{p,q} Ay.$$

We want $x^T A^T I_{p,q} Ay = \langle x,y\rangle_{p,q} = x^T I_{p,q} y$, and this is true if and only if $A^T I_{p,q} A = I_{p,q}$. This is preserved under limits, so $\mathrm{O}(p,q)$ is closed. If $p = n, q = 0$, then $\mathrm{O}(n,0) = O(n0$ the standard orthogonal group, i.e. $\{A \in \mathrm{GL}(n,\mathbb{R}) : A^T A = I\}$. Note that if $A \in O(p,q)$ then $A^T I_{p,q} A = I_{p,q}$, so $\det(A)^2 = 1$. Hence $\det(A) = \pm 1$.

Definition 1.6. The *special orthogonal group of signature* p,q is defined to be

$$\mathrm{SO}(p,q) := O(p,q) \cap \mathrm{SL}(n,\mathbb{R}) = \{A \in \mathrm{GL}(n,\mathbb{R}) : A^T I_{p,q} A = I_{p,q}, \det(A) = 1\}.$$

Also, define $\mathrm{O}(n,\mathbb{C}) = \{A \in \mathrm{GL}(n,\mathbb{C}) : (Ax)\cdot(Ay) = x \cdot y$ for all $x,y \in \mathbb{C}^n\} = \{A \in \mathrm{GL}(n,\mathbb{C}) : A^T A = I\}$.. Recall that

$$x \cdot y = \sum_{k=1}^{n} x_k y_k.$$

This is *not* a positive-definite inner product but is a non-degenerate symmetric bilinear form. $\mathrm{O}(n,\mathbb{C})$ is called *the complex orthogonal group*. There is no notion of signature (p,q) over $\mathbb{C}$.

Definition 1.7. $\mathrm{U}(n)$ is defined to be

$$\mathrm{U}(n) = \{A \in \mathrm{GL}(n,\mathbb{C}), \langle Az, Aw\rangle = \langle z,w\rangle \text{ for all } z,w \in \mathbb{C}^n\} = \{A \in \mathrm{GL}(n,\mathbb{C}) : A^*A = I\},$$

and we call $\mathrm{U}(n)$ the *unitary group of* $\mathbb{C}^n$. Similarly, $\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n,\mathbb{C})$ is said to be *the special unitary group of* $\mathbb{C}^n$.

Remark 1.5. If $A \in \mathrm{U}(n)$ then $A^*A = I$, so $\overline{\det(A)}\det(A) = |\det(A)|^2 = 1$. Thus $\det(A) = e^{i\theta}$ for some θ.

Example 1.8. Continued from Example 1.4:

(2bc) $\mathrm{O}(p,q)$ and $\mathrm{O}(n,\mathbb{C})$ are examples of matrix Lie groups.

(3) Consider the standard inner product on $\mathbb{C}^n$

$$\langle z,w\rangle = \sum_{k=1}^{n} \overline{z_k} w_k = z^* w = \overline{z}^T w$$

($*$ denotes conjugate transpose). The Hermitian inner product is linear in w and conjugate-linear in z, i.e., $\langle \lambda z, w\rangle = \overline{\lambda}\langle z,w\rangle$ and $\langle w,z\rangle = \overline{\langle z,w\rangle}$. Also $\langle z,z\rangle \geq 0$ with equality holding if and only if $z = 0$. Such inner product is known to be *sesquilinear*.

(4) $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are matrix groups.

(5) Symplectic groups are matrix groups. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Define $B = \mathbb{F}^{2n} \times \mathbb{F}^{2n} \to \mathbb{F}$ by

$$B(x,y) = (x_1 y_{n+1} - x_{n+1} y_1) + \cdots + (x_n y_{n+n} - x_{n+n} y_n) = \sum_{k=1}^{n} (x_k y_{n_k} - x_{n+k} y_k).$$

Then B is a skew-symmetric bilinear form. Let

$$J = \begin{bmatrix} 0_{n+n} & -I_{n+n} \\ I_{n+n} & 0 \end{bmatrix}.$$

Then $B(x,y) = (Jx)^T y = (Jx) \cdot y = x^T J^T y = -x^T J y$.

(5a) Over $\mathbb{R}$, the *real symplectic group* $\mathrm{Sp}(n,\mathbb{R})$ is defined to be

$$\begin{aligned} \mathrm{Sp}(n,\mathbb{R}) &= \{A \in \mathrm{GL}(2n,\mathbb{R}) : A^T J A = J\} \\ &= \{A : B(Ax,Ay) = B(x,y) \text{ for all } x,y \in \mathbb{R}^{2n}\}. \end{aligned}$$

(5b) Over $\mathbb{C}$, the *complex symplectic group* $\mathrm{Sp}(n)$ is defined to be $\mathrm{Sp}(n) = \mathrm{Sp}(n,\mathbb{C}) \cap \mathrm{U}(2n)$.

(6) *Generalized Heisenberg groups* H_n are

$$H_n := \left\{ A \in \mathrm{GL}(n,\mathbb{R}), A = \begin{bmatrix} 1 & * & * & \cdots & * \\ & 1 & * & \cdots & * \\ & & 1 & \vdots & \vdots \\ & & & \ddots & * \\ 0 & & & & 1 \end{bmatrix} \right\},$$

i.e. upper triangular matrices with 1 on diagonals. This is clearly closed. To see it is a subgroup, let $E_k = \mathrm{span}\{e_1, e_2, \ldots, e_k\}$, the standard basis of $\mathbb{R}^n$. Then $E_k \subset E_{k+1}$ and $A(E_k) \subset E_k$. More precisely, $A(e_k) = e_k + (\text{stuff in } E_{k_1})$. Note that $A|_{E_k}$ is invertible, so $A : E_k \to E_k$ is an isomorphism. If $A, B \in H_n$, then

$$A(e_k) = e_k + f_{k-1}$$

where $f_{k-1}, \widetilde{f_{k-1}} \in E_{k-1}$, and $(BA)(e_k) = B(A(e_k)) = B(e_k + f_{k_1}) = e_k + B(f_{k-1} + \widetilde{f_{k-1}}$. So $BA \in H_n$, since $B(f_{k-1}) \in E_{k-1}$. Thus $A^{-1}(e_k) = e_k - \underbrace{A^{-1}(f_{k-1})}_{\in E_{k-1}}$, so $A^{-1} \in H_n$. If $n = 3$, then H_n becomes the *classical Heisenberg group*.

(7) $\mathbb{R}^* \cong \mathrm{GL}(1,\mathbb{R}), \mathbb{C}^* \cong \mathrm{GL}(1,\mathbb{C}), S^1 \cong \mathrm{U}(1)$ are all matrix Lie groups.

View $\mathbb{R}^n$ as a subgroup of $\mathrm{GL}(n,\mathbb{R})$ consisting of diagonal matrices with positive entries on diagonals by

$$x := \begin{bmatrix} x_1 \\ x_2 \\ \ddots \\ x_n \end{bmatrix} \leftrightarrow \begin{bmatrix} e^{x_1} & & & 0 \\ & e^{x_2} & & \\ & & \ddots & \\ 0 & & & e^{x_n} \end{bmatrix} =: P(x)$$

$$x + y \leftrightarrow \begin{bmatrix} e^{x_1+y_1} & & & 0 \\ & e^{x_2+y_2} & & \\ & & \ddots & \\ 0 & & & e^{x_n+y_n} \end{bmatrix}.$$

$P(x+y) = P(x)P(y)$ so P is a group isomorphism. Note that P is closed in $\mathrm{GL}(n, \mathbb{C})$ but not in $\mathrm{M}_n(\mathbb{C})$.

2. January 07: Topological properties of matrix groups

Today, we will talk about topological properties of matrix groups, such as

- compactness
- connectedness
- simple-connectedness.

2.1. *On compactness*

Recall that $\mathrm{M}_n(\mathbb{C}) \cong \mathbb{R}^{2n^2}$ is separable, so Heine-Borel applies. Thus, a subset G on $\mathrm{M}_n(\mathbb{C})$ is compact if and only if it is sequentially compact if and only if it is bounded and closed in $\mathrm{M}_n(\mathbb{C})$. Hence, a matrix G is compact if and only if it is bounded and closed in $\mathrm{M}_n(\mathbb{C})$ (Recall that it is already closed in $\mathrm{GL}(n, \mathbb{C})$ by definition.).

Example 2.1. $\mathrm{GL}(n, \mathbb{C})$ and $\mathrm{GL}(n, \mathbb{R})$ are not compact since they are not bounded and not closed in $\mathrm{M}_n(\mathbb{C})$.

Example 2.2. If $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, then $\mathrm{SL}(n, \mathbb{F}), O(p, q), \mathrm{SO}(p, q), \mathrm{U}(n), \mathrm{SU}(n), \mathrm{Sp}(n, \mathbb{F}), \mathrm{Sp}(n), H_n$ are all closed in $\mathrm{M}_n(\mathbb{C})$. Thus it suffices to check if they are bounded.

Suppose $A \in O(n)$, that is, $A^T A = I$, or equivalently, columns of A are orthonormal, hence $|A_{ij}| \leq 1$ for all i, j. Hence $\mathrm{O}(n)$ is compact. It also follows that $\mathrm{SO}(n) = O(n) \cap \mathrm{SL}(n, \mathbb{R})$ is compact also.

As for $\mathrm{U}(n)$, note that $A \in \mathrm{U}(n)$ if and only if $A^*A = I$, i.e., columns of A are orthonormal with respect to Hermitian inner product. Thus $|A_{ij}| \leq 1$ for all i, j hence bounded. Thus $\mathrm{U}(n)$ is compact, and so is $\mathrm{SU}(n)$. Thus, $\mathrm{Sp}(n) := \mathrm{U}(n) \cap \mathrm{Sp}(n, \mathbb{C})$ is compact also. However, $\mathrm{SL}(n, \mathbb{F})$ and $\mathrm{O}(n, \mathbb{C})$ are not compact if $n \geq 2$. $\mathrm{Sp}(n, \mathbb{F})$ is not compact for all $n \geq 1$. If $p, q \geq 1$, then $\mathrm{O}(p, q)$ and $\mathrm{SO}(p, q)$ are not compact. You will prove them in Assignment #1.

2.2. *On connectedness*

Recall that a topological space XR is connected if and only if $U \subseteq X$ open, closed, non-empty implies $U = X$ (i.e., there can be no separation). Every topological space is a disjoint union of its "connected components" $X = \coprod_{\alpha \in A} U_\alpha$ with each U_α connected. Note that all the connected components are closed.

Definition 2.3. A *path* in a topological space X is a continuous map $\alpha : [0, 1] \to X$. A topological space X is *path-connected* if and only if any two points $a, b \in X$ can be joined by a path (i.e., there exists a path such that $\alpha(0) = a, \alpha(1) = b$).

Fact 1. If X is locally path-connected (includes manifold), then connectedness and path-connectedness are equivalent. Thus connected components = path-connected components. Since all matrix Lie groups are manifolds, the two notions are equivalent. Therefore, throughout this course, we are always in the setting where the two definitions are equivalent. So we will say "connected" but will test using the definition of path-connectedness.

Proposition 2.4. *Let G be a matrix group, and G_0 be the connected component containing the identity I. Then G_0 is a matrix group, and is a subgroup of G.*

Proof. Connected components are closed in G, so G_0 is closed in G hence closed in $\mathrm{GL}(n,\mathbb{C})$. We thus only need to show that G_0 is a subgroup. Let $A, B \in G_0$. Since there exist paths $\alpha, beta : [0,1] \to G$ such that $\alpha(0) = \beta(0) = I$ and $\alpha(1) = A, \beta(1) = B$, then $\alpha\beta : [0,1] \to G$ defined by $(\alpha\beta)(t) = \alpha(t)\beta(t)$ is continuous since the matrix multiplication map is continuous. Note that $(\alpha\beta)(0) = I^2 = I$ and $(\alpha\beta)(1) = AB$. Thus $AB \in G_0$. Define $\alpha^{-1} : [0,1] \to G$ to be the matrix inversion map, i.e., $\alpha^{-1}(t) := (\alpha(t))^{-1}$. Then α^{-1} is continuous, so $\alpha^{-1}(0) = I$ and $\alpha^{-1}(1) = A^{-1}$. This completes the proof. ☐

Proposition 2.5. $\mathrm{GL}(n,\mathbb{C})$ *is connected.*

Proof. For this, we prove via induction. If $n = 1$, then $\mathrm{GL}(1,\mathbb{C}) = \mathbb{C}^*$ is connected. Let $n \geq 2$. Recall that every complex square matrix is triangularizable. In other words, for any $A \in \mathrm{GL}(n,\mathbb{C})$, one can find $P \in \mathrm{GL}(n,\mathbb{C})$ such that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}.$$

Since $\det A = \lambda_1 \cdots \lambda_n \neq 0$ so none of λ_i can be zero. Let $B := P^{-1}AP$, and define $\beta : [0,1] \to \mathrm{GL}(n,\mathbb{C})$ as follows:

$$\beta(t) := \begin{bmatrix} \lambda_1 & & & *(1-t) \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}.$$

Then $\beta(0) = B$, and let $\beta(1) =: D$. Define $\alpha(t) = P\beta(t)P^{-1}$. Then we have $\alpha(0) = PBP^{-1} = A$ and $\alpha(1) = PDP^{-1}$. Now apply the $n = 1$ case to derive that there must exist continuous paths $f_i : [0,1] \to \mathbb{C}^* = \mathrm{GL}(1,\mathbb{C})$ such that $f_i(0) = \lambda_i$ and $f_i(1) = 1$. Define

$$\gamma(t) := P \begin{bmatrix} f_1(t) & & & \\ & f_2(t) & & \\ & & \ddots & \\ & & & f_n(t) \end{bmatrix} P^{-1}.$$

Then $\gamma(0) = PDP^{-1} = \alpha(1)$ and $\gamma(1) = PIP^{-1} = I$. Thus there exists a path from A to I. Thus $\mathrm{GL}(n,\mathbb{C})$ is connected, as required. ☐

Proposition 2.6. $\mathrm{SL}(n,\mathbb{C})$ *is connected.*

Proof. Clearly, $\mathrm{SL}(1,\mathbb{C}) = \{1\}$ is connected. For $n \geq 2$, proceed as before, with $\lambda_1\lambda_2 \cdots \lambda_n = 1$. Define α as we did in the previous proof, and for γ take

$$f_n(t) = \frac{1}{f_1(t)f_2(t)\cdots f_{n-1}(t)}.$$

☐

Proposition 2.7. U(n) *and* SU(n) *are connected for all* $n \geq 1$.

Proof. If $A \in \mathrm{U}(n)$, then $A^*A = I$. Thus A is normal and A is unitarily diagonalizable, i.e., there exists $P \in \mathrm{U}(n)$ such that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = \begin{bmatrix} e^{i\theta_1} & & & \\ & e^{i\theta_2} & & \\ & & \ddots & \\ & & & e^{i\theta_n} \end{bmatrix}.$$

Moreover, since $A \in \mathrm{U}(n)$, each $\lambda_k = e^{i\theta_k}$. Let

$$\alpha(t) = P^{-1} \begin{bmatrix} e^{i\theta_1(1-t)} & & & \\ & e^{i\theta_2(1-t)} & & \\ & & \ddots & \\ & & & e^{i\theta_n(1-t)} \end{bmatrix} P.$$

Then $\alpha(0) = A$ and $\alpha(1) = P^{-1}IP = I$. So α is a path in U(n). Hence U(n) is connected. Similarly, we can prove that SU(n) is connected by using the similar argument as in the SL$(n, \mathbb{C})$ case: take

$$\lambda_n(t) = \frac{1}{\lambda_1(t) \cdots \lambda_{n-1}(t)},$$

and the claim follows. □

Proposition 2.8. GL$(n, \mathbb{R})$ *is* not *connected.*

Proof. Suppose that $A, B \in \mathrm{GL}(n, \mathbb{R})$ with $\det(A) > 0$ and $\det(B) < 0$. Suppose that $\alpha(t)$ is a path in GL$(n, \mathbb{R})$ from A to B. Let $f := \det \circ \alpha : [0, 1] \to \mathbb{R}^*$. Then f is continuous. By the intermediate value theorem, since $f(0) > 0$ and $f(1) < 0$, there must exist $t \in [0, 1]$ such that $f(t) = 0$, and this is a contradiction. □

Proposition 2.9. O(n) *is* not *connected but* SO(n) is *connected, for all* $n \geq 1$.

Proof. You will prove this in Assignment #1! □

2.3. *On simple-connectedness*

Definition 2.10. A subset G of $\mathrm{M}_n(\mathbb{C})$ is *simple-connected* if it is connected *and* every closed loop can be continuously deformed to a point while staying in G. In other words, if $\alpha : [0, 1] \to G$ such that $\alpha(0) = \alpha(1) = A$ ("closed loop"). Then there exists $H : [0, 1] \times [0, 1] \to G$ continuous such that $H(t, 0) = \alpha(t)$ for all $t \in [0, 1]$ and $A = H(1, s) = H(0, s)$ for all s, i.e, $H(\_, s)$ is a loop. Also, note that $H(t, 1) = A$ for all t (constant loop).

Remark 2.1. Checking whether a subset G is simply connected or not is *difficult.* It is crucially important for the "Lie correspondence" between Lie groups and Lie algebras.

Proposition 2.11. *Recall that* SO$(2) \cong$ U(1) *(Assignment #1), and* U(1) *and* S^1 *are homeomorphic. Thus they are* not *simply connected. However,* Sp$(1) \cong$ SU(2)*, and* SU(2) *is homeomorphic to* S^3*, which* is *simply connected. Recall that* S^k *is simply connected for all* $k \geq 2$*. Therefore,* SU(n) *is simply connected but* $U(n)$ *is not.* SO(n) *is also not simply connected for all* $n > 2$*. The spin group* Spin(n) *is simply connected for all* $n > 2$.

2.4. *Polar decomposition of* $\mathrm{SL}(n,\mathbb{R})$

Remark 2.2. Recall that if $A \in \mathrm{M}_n(\mathbb{R})$ is symmetric, then it is orthogonally diagonalizable with real eigenvalues, i.e., there exists $R \in O(n)$ such that

$$R^{-1}AR = R^TAR = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

with $\lambda_i \in \mathbb{R}$.

Definition 2.12. We say that a symmetric matrix is *positive* if

$$(Ax)\cdot x = x^TAx = \sum_{i,j} A_{ij}x_ixg_j \geq 0,$$

where the equality holds if and only if $x = 0$. Recall that A is positive if and only if $\lambda_i > 0$ for all i (exercise). If A is positive-symmetric, then we can define a *square root* $A^{1/2}$ as follows. First, $A^{1/2}$ is also positive-symmetric defined by

$$A^{1/2} := R \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{bmatrix} R^{-1}.$$

Then $(A^{1/2})^2 = A$ and $A^{1/2}$ is positive-symmetric. In fact, this $A^{1/2}$ is the unique matrix in $\mathrm{M}_n(\mathbb{R})$ with these two properties (exercise).

Proposition 2.13. *Let* $A \in \mathrm{SL}(n,\mathbb{R})$. *Then there exists a* unique *pair* (R,P) *where* $R \in \mathrm{SO}(n)$ *and* P *is real and positive-symmetric such that* $\det(P) = 1$, *and* $A = RP$

Proof. $A^TA = PR^TRP = P^2$, and A^TA is positive-symmetric:

$$(A^TAx)\cdot x = (A^TAx)^Tx = x^TA^TAx = |Ax|^2 \geq 0,$$

and the equality holds if and only if $x = 0$ as A is invertible. We need to show that $P = (A^TA)^{1/2}$. So $R = AP^{-1}$, thus only need to show $R \in \mathrm{SO}(n)$.

$$\begin{aligned} RR^T &= (AP^{-1})(AP^{-1})^T = AP^{-1}P^{-1}A^T = A(P^2)^{-1}A^T \\ &= A(A^TA)^{-1}A^T = I. \end{aligned}$$

Thus $R \in O(n)$. On the other hand, $\det(A) = \det(R)\det(P) = 1$ and $\det(P) > 0$, so $\det(R) = 1$. Hence $R \in \mathrm{SO}(n)$ and $\det(P) = 1$, as required. □

Remark 2.3. For $\mathrm{SL}(n,\mathbb{C})$, a self-adjoint matrix A (i.e., $A^* = A$) is called positive if $\langle Ax, x\rangle := x^*A^*x = x^*Ax > 0$ for all $x \neq 0$

Proposition 2.14. *Given* $A \in \mathrm{SL}(n,\mathbb{C})$, *there exists a* unique *pair* (U,P) *with* $U \in \mathrm{SU}(n)$ *with* P *self-adjoint and positive such that* $A = UP$ *and* $\det(P) = 1$.

2.5. *Homomorphisms and Isomorphisms*

Definition 2.15. Let G, H be matrix Lie groups. Let $F : G \to H$ be a map. We say F is a *matrix Lie group homomorphism* if:

- F is a group homomorphism.
- F is continuous.

In addition, if F is bijective and F^{-1} is continuous, then F is called a *matrix Lie group isomorphism* (i.e., it is a group isomorphism and a homeomorphism of topological spaces).

Remark 2.4. In practice, most group homomorphisms between matrix Lie groups will be continuous.

Definition 2.16. Two matrix Lie groups are *isomorphic* (as matrix Lie groups) if there exists a matrix Lie group isomorphisms between them. We will only care about matrix Lie groups up to isomorphism

Example 2.17. $\mathbb{R}^* \cong \mathrm{GL}(1, \mathbb{R})$ and $\mathbb{C}^* \cong \mathrm{GL}(1, \mathbb{C})$. Also, $U(1) \cong \mathrm{SO}(2)$ and $\mathrm{Sp}(1) \cong \mathrm{SU}(2)$ (Assignment #1).

Example 2.18 (Matrix Lie group homomorphisms). The determinant map $\det : \mathrm{GL}(n, \mathbb{F}) \to \mathbb{F}^* \cong \mathrm{GL}(1, \mathbb{F})$ is a Lie group homomorphism. Also, $F : \mathbb{R} \to \mathrm{SO}(2)$ defined as

$$F(t) = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix}$$

is a homomorphism. One can check that $F(0) = I, F(-t) = (F(t))^{-1}$, and $F(t+s) = F(t)F(s)$.

3. January 12

Example 3.1 (One more important example of a matrix Lie group homomorphism). Quick aside: there exists a natural Hermitian inner product on $\mathrm{M}_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$

$$\langle A, B \rangle = \sum_{i,j} \overline{A_{ij}} B_{ij} = \sum_{i,j} (A^*))_{ji} B_{ij} = \sum_{j=1}^{n} (A^* B)_{jj} = \mathrm{tr}(A^* B).$$

We will define a homomorphism from $\mathrm{SU}(2)$ to $\mathrm{SO}(3)$. Define $V = \{A \in \mathrm{M}_2(\mathbb{C}); A^* = A, \mathrm{tr}(A) = 0\}$. Note that V is a real subspace of $\mathrm{M}_2(\mathbb{C})$. Since

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^* = \begin{bmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

we have $\bar{a} = a, \bar{d} = d \in \mathbb{R}$ and $c = \bar{b}$. Therefore,

$$V = \left\{ \begin{bmatrix} t & z \\ \bar{z} & -t \end{bmatrix} : t \in \mathbb{R}, z \in \mathbb{C} \right\} \cong \mathbb{R}^3$$

as real vector spaces; V has

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \right\}$$

as a basis.

Now take a look at $\langle A, B\rangle = \operatorname{tr}(A^*B)$, restricted to V. If

$$A = \begin{bmatrix} t_1 & z_1 \\ \bar{z}_1 & -t_1 \end{bmatrix}, B = \begin{bmatrix} t_2 & z_2 \\ \bar{z}_2 & -t_2 \end{bmatrix}$$

then $\operatorname{tr}(A^*B) = 2(t_1t_2 + x_1x_2 + y_1y_2)$, where $z_k = x_k + iy_k$. Thus $(V, \frac{1}{2}\langle \_, \_\rangle \cong (\mathbb{R}^3, \cdot)$ as inner product spaces.

Now define a map $F : \mathrm{SU}(2) \to \mathrm{GL}(V) \cong \mathrm{GL}(3, \mathbb{R})$ such that $F(U)(A) = UAU^{-1}$ if $A \in V$ (note that $F(U) \in \mathrm{GL}(V)$. Now need to verify that $F(U)(A) \in V$. First, we see that $\operatorname{tr}(UAU^{-1}) = \operatorname{tr}(A) = 0$. Also, $(UAU^{-1})^* = (UAU^*)^* = UAU^*$. But then since $U \in \mathrm{SU}(2)$, $U^* = U^{-1}$. Thus $UAU^* \in V$. Clearly, $F(U) : V \to V$ is linear, and $F(U^{-1}) = (F(U))^{-1}$ since $U^{-1}(UAU^{-1})(U^{-1})^{-1} = A$.

Let $A, B \in V$, and to simplify notation, write $F_U := F(U)$. Consider the inner product

$$\begin{aligned}\langle F_UA, F_UB\rangle &= \operatorname{tr}((UAU^*)(UBU^*)) = \operatorname{tr}(UABU^{-1}) \\ &= \operatorname{tr}(AB) = \langle A, B\rangle.\end{aligned}$$

Thus, for all $U \in \mathrm{SU}(2)$, we have $F_U \in \mathrm{O}(V) \cong \mathrm{O}(3)$, and $F_I = I \in \mathrm{O}(3)$. Since $\det(F_U)$ is continuous in U and $\mathrm{SU}(2)$ is connected, $\det(F_U) = +1$ for all $U \in \mathrm{SU}(2)$. Therefore indeed F sends any matrix in $\mathrm{SU}(2)$ to an element in $\mathrm{SO}(3)$. Note F is continuous, and it is easy to verify that $F_{U_1U_2} = F_{U_1} \circ F_{U_2}$. Thus it is indeed a homomorphism.

Remark 3.1. F is not isomorphic since F is not injective – note that $F_U = F_{-U}$. WE will actually see that F is surjective and $\ker(F) = \{\pm I\}$ (so it is "two-to-one"). Moreover, we will see that induced homomorphism of the Lie algebras is actually an *isomorphism*.

3.1. *Matrix exponential*

Definition 3.2. Let $X \in \mathrm{M}_n(\mathbb{C})$. Then the *matrix exponential* e^X is defined as

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

First, we need to prove that this definition actually makes sense. For this, we need to prove a few claims about norms:

Claim. For any $X, Y \in \mathrm{M}_n(\mathbb{C})$:

(1) $\|X + Y\| \leq \|X\| + \|Y\|$ (triangle inequality)
(2) $\|XY\| \leq \|X\|\|Y\|$

Proof. The first one is clear since $\|\_\|$ is a norm. For the second one,

$$\begin{aligned}\|XY\|^2 = \sum_{i,j=1}^{n} |XY|_{ij}^2 &= \sum_{i,j} \left(\sum_{k=1}^{n} \overline{X_{ik}} Y_{kj} \right)^2 = \sum_{i,j} |\langle X_i, Y_j\rangle|^2 \\ &\leq \sum_{i,j} \|X_i\|^2 \|Y_j\|^2 = \|X\|^2\|Y\|^2,\end{aligned}$$

where X_i is the i-th row of X and Y_j the j-th column of Y. □

Definition 3.3. Let $X_m \in \mathrm{M}_n(\mathbb{C})$ for all m. We say that $\sum X_m$ converges absolutely if $\sum \|X_m\| < \infty$.

Claim. If $\sum X_m$ converges absolutely, then it converges.

Proof. Let

$$S_N := \sum_{m=0}^{N} X_m.$$

We have

$$\|S_{N'} - S_N\| = \left\| \sum_{m=N+1}^{N'} X_m \right\| \le \sum_{m=N}^{N'} \|X_m\| \to 0$$

as $N, N' \to \infty$, since $\sum \|X_m\| < \infty$. □

Claim. $e^X := \sum \frac{1}{m!} X_m$ converges for all $X \in \mathrm{M}_n(\mathbb{C})$.

Proof. It converges absolutely, since

$$\left\| \frac{1}{m!} X^m \right\| \le \frac{1}{m!} \|X\|^m.$$

So

$$\sum_m \left\| \frac{1}{m!} X^m \right\| \le \sum_m \frac{1}{m!} \|X\|^m = e^{\|X\|} < \infty.$$

Moreover, the partial sums are continuous in X and the convergence is uniform on $\{X : \|X\| \le R\}$ for all $R > 0$. Thus e^X is continuous in X. □

Proposition 3.4 (Properties of matrix exponential)**.** *For any $X, Y \in \mathrm{M}_n(\mathbb{C})$:*

(1) $e^0 = I$.
(2) $(e^X)^* = e^{X^*}$.
(3) If $XY = YX$, then $e^X e^Y = e^{X+Y} = e^Y e^X$.
(4) e^X is always invertible, with $(e^X)^{-1} = e^{-X}$.
(5) $e^{(s+t)X} = e^{sX} e^{tX}$ for all $s, t \in \mathbb{C}$.
(6) If P is invertible, then $e^{PXP^{-1}} = P e^X P^{-1}$.
(7) $\|e^X\| \le e^{\|X\|}$.

Proof. (Property (2)) Recall that adjoint A^* of a matrix A is defined as the unique $n \times n$ matrix such that

$$\langle Aa, b \rangle = \langle a, A^* b \rangle$$

for all $a, b \in \mathbb{C}^n$. If

$$S_N = \sum_{m=0}^{N} \sum_{m=0}^{N} \frac{X^m}{m!},$$

then $\langle s_N a, b \rangle = \langle a, (s_N)^* b \rangle$ for all $a, b \in \mathbb{C}^n$. Note that

$$(S_N)^* = \sum_{m=0}^{N} \frac{(X^*)^m}{m!}.$$

Take $N \to \infty$ to get $\langle e^X a, b \rangle = \langle a, e^{X^*} b \rangle$ for all a, b. Hence $e^{X^*} = (e^X)^*$.

(Property (3)) Since the series converges absolutely, we are free to rearrange the terms:

$$\begin{aligned} e^X e^Y &= \left(I + X + \frac{X^2}{2} + \frac{X^3}{6} + \cdots\right)\left(I + Y + \frac{Y^2}{2} + \frac{Y^3}{6} + \cdots\right) \\ &= \sum_{m=0}^{\infty}\left(\sum_{k=0}^{m} \frac{X^k}{k!}\frac{Y^{m-k}}{(m-k)!}\right) = \sum_{m=0}^{\infty} \frac{1}{m}! \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} X^k Y^{m-k} \\ &\overset{!}{=} \sum_{m=0}^{\infty} \frac{1}{m!}(X+Y)^m = e^{X+Y}. \end{aligned}$$

Note that $\overset{!}{=}$ follows since $XY = YX$.

(Property (4)) Let $Y = -X$ in (3). And $X, -X$ commute. Hence $e^X e^{-X} = e^{-X} e^X = e^{X+(-X)} = e^0 = I$. Thus $e^{-X} = (e^X)^{-1}$.

(Property (6)) Notice

$$P\left(\sum_{m=0}^{N} \frac{1}{m!} X^M\right) P^{-1} = \sum_{m=0}^{N} \frac{1}{m!}(PXP^{-1})^m.$$

Take $N \to \infty$ to get $Pe^X P^{-1} = e^{PXP^{-1}}$. □

Proposition 3.5. *Let $\alpha(t) = e^{tX}$ for $t \in \mathbb{R}$. Then $\alpha : \mathbb{R} \to \mathrm{M}_n(\mathbb{C})$ is smooth, and that*

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

for all t. Therefore

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

Proof. e^{tX} is a convergent power series in t, so can be differentiated term-by-term. So we have

$$\frac{d}{dt} e^{tX} = \frac{d}{dt}\left(\sum_{m=0}^{\infty} \frac{t^m X^m}{m!}\right) = \sum_{m=1}^{\infty} \frac{m t^{m-1} X^m}{m!} = X\left(\sum_{m=1}^{\infty} \frac{(tX)^{m-1}}{(m-1)!}\right) = X e^{tX}.$$ □

How can we compute e^X in practice? A theorem of linear algebra (see Appendix B of the theft or Hoffman-Kunze) says that every matrix X can be written uniquely in the form $X = S + N$ where S is diagonalizable and N is nilpotent (i.e., there exists $k \in \mathbb{Z}_+$ such that $N^k = O$) and $SN = NS$. Then

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

Since S is diagonalizable one can find P such that

$$P^{-1} S P = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} = D$$

and $PDP^{-1} = S$. So $e^S = e^{PDP^{-1}} = Pe^D P^{-1}$, so

$$e^D = \begin{bmatrix} e^{\lambda_1} & & & 0 \\ & e^{\lambda_2} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_n} \end{bmatrix}.$$

Also, since N is nilpotent,

$$e^N = \sum_{m=0}^{\infty} \frac{1}{m!} N^m$$

reduces to a finite sum.

Example 3.6. Write

$$A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} = \underbrace{\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}}_{=S} + \underbrace{\begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}}_{=N}.$$

Since $N^2 = O$, we have $e^N = 1 + N$. Thus

$$e^A = e^S e^N = \begin{bmatrix} e^a & e^a b \\ 0 & e^a \end{bmatrix}.$$

Example 3.7. Let

$$X = \begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix} = \theta \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{=J}. \; (\theta \in \mathbb{R})$$

It is easy to check that $J^2 = -I$, and that $J^{2l} = (-1)^l I$ and $J^{2l+1} = (-1)^l J$. So

$$\begin{aligned} e^X = e^{\theta J} &= \sum_{m=0}^{\infty} \frac{(\theta J)^m}{m!} \\ &= \sum_{l=0}^{\infty} \frac{\theta^{2l} J^{2l}}{(2l)!} + \sum_{l=0}^{\infty} \frac{\theta^{2l+1} J^{2l+1}}{(2l+1)!} \\ &= \left(\sum_{l=0}^{\infty} \frac{(-1)^0 \theta^{2l}}{(2l)!} I \right) + \left(\sum_{l=0}^{\infty} \frac{(-1)^l \theta^{2l+1}}{(2l+1)!} \right) J \\ &= \cos\theta I + \sin\theta J. \end{aligned}$$

Therefore,

$$e^{\theta J} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in \mathrm{SO}(2).$$

3.2. *Matrix logarithm*

Proposition 3.8. *Define* $\log(z)$ *by*

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

This is defined and is analytic in $B_1(1) := \{w \in \mathbb{C} : |w-1| < 1\}$*. Moreover, for all* $z \in \mathbb{C}$ *such that* $|z-1| < 1$*, we have* $e^{\log(z)} = z$ *for all* $u \in \mathbb{C}$ *with* $|u-0| < \log 2$*, we have* $|e^u - 1| < 1$ *and* $\log(e^u) = u$*.*

Proof. Let $x > 0$ and $|x| < 1$. Then the identity

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

Integrate term by term to see

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

Let $z = 1 - x$ hence $x = 1 - z$. So

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

This has radius of convergence $|z-1| < 1$, as expected.

Also, it coincides with the usual logarithm on $(0,2) \in \mathbb{R}$, i.e., $e^{\log(x)} = x$ for all $x \in (0,2)$. By analyticity, $e^{\log(z)} = z$ for all $z \in B_1(1)$. If $|u| < \log 2$ then

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

So $\log(e^u) = u$ for all real u with $|u| < \log 2$. By analyticity, we have $\log(e^w) = w$ for all $w \in \mathbb{C}$ and $|w| < \log 2$. ☐

Definition 3.9. For $A \in \mathrm{M}_n(\mathbb{C})$, define

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

whenever this series converges. We ono for sure this converges when $\|A - I\| < q$ because $\|(A-I)^m\| \leq \|A - I\|^m$ and

$$\sum_m (-1)^{m+1} \frac{(z-1)^m}{m!}$$

converges.

Remark 3.2. In general, it may converge on a bigger set, but we need not care about this for the purpose of this course. For instance, if $A - I$ is nilpotent, then $\log(A)$ converges regardless of $\|A - I\|$.

So, in summary, the log function

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

is defined (and continuous, by the uniform convergence of compact subsets) on $\{A \in \mathrm{M}_n(\mathbb{C}) : \|A - I\| < 1\}$. If $\|A - I\| < 1$, then $e^{\log A} = A$. Also, if $\|X\| < \log 2$, then $\|e^X - I\| < 1$ and $\log(e^X) = X$. Hence, e and log are continuous and are inverses of each other near O and I, respectively. We will start proving this fact in the next lecture. We will also discuss more properties of e and log. We will also talk about the "Lie product formula", and Lie algebras of G.

4. January 14: Lie product formula and Lie algebras

Proposition 4.1. *The function*

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

is defined and continuous *on* $B_1(I) := \{A : \mathrm{M}_n(\mathbb{C}) : \|A - I\| < 1\}$. *Also, if* $\|A - I\| < 1$, *then* $e^{\log A} = A$ *and if* $\|x\| < \log 2$, *then* $\|e^X - I\| < 1$ *and* $\log e^X = X$.

Proof. Note that $\|(A-I)^m\| \leq \|A - I\|^m < 1$ for all m. Hence

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

converges. So the tseries for log converges absolutely for $A \in B_1(I)$ and uniformly on compact subsets. Hence log is continuous on $B_1(I)$.

To show that $e^{\log A} = A$ for all A such that $\|A - I\| < 1$, consider the following two cases:

(1) A is diagonalizable

If so, then there exists some P such that

$$P^{-1}AP = D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

So we have

$$A - I = PDP^{-1} - I = P(D-I)P^{-1} = P \begin{bmatrix} \lambda_1 - 1 & & & \\ & \lambda_2 - 1 & & \\ & & \ddots & \\ & & & \lambda_n - 1 \end{bmatrix} P^{-1},$$

and $\|A - I\| = \|P(D-I)P^{-1}\| \leq \|P\|\|D - I\|\|P^{-1}\|$. Now we need the following claim:

Claim. If $\|B\| < 1$ then all eigenvalues of B has norm < 1. (Assignment #2!)

Let $B = D - I$. Note that $\|B\| < 1$ (Spiro couldn't think of why this is the case – we need to check this ourselves.). Hence, by the claim above, we have $|\lambda_i - 1| < 1$ for all i. Hence

$$\sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m} = P \left(\sum_{m=1}^{\infty} \frac{(-1)^{m+1}(D-I)^m}{m} \right) P^{-1},$$

so equivalently,

$$\log A = P \begin{bmatrix} \log \lambda_1 & & & \\ & \log \lambda_2 & & \\ & & \ddots & \\ & & & \log \lambda_n \end{bmatrix} P^{-1}.$$

Similarly, it follows

$$e^{\log A} = P \begin{bmatrix} e^{\log \lambda_1} & & & \\ & e^{\log \lambda_2} & & \\ & & \ddots & \\ & & & e^{\log \lambda_n} \end{bmatrix} P^{-1} = A.$$

(2) General case We need to use the fact that the diagonalizable matrices are dense in $\mathrm{M}_n(\mathbb{C})$, which you will prove in Assignment #2. We have $\mathbf{e}^{\log A_m} = A_m$ for all diagonalizable matrices A_m such that $\|A_m - I\| < 1$. So by continuity, we have $e^{\log A} = A$ as $m \to \infty$. One can derive $\log e^X = X$ in a similar manner.

This completes the proof. □

Proposition 4.2. *There exists $c > 0$ such that for all $B \in \mathrm{M}_n(\mathbb{C})$ with $\|B\| < \frac{1}{2}$ such that*

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

Note that c is independent of B.

Proof. $\log(I+B)$ is defined since $\|(I+B) - I\| = \|B\| < \frac{1}{2} < 1$. Hence

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

So we have

$$\|\log(I+B) - B\| \leq \|B\|^2 \sum_{m \geq 2} \frac{(1/2)^{m-2}}{m}.$$

Now choose $c = \sum_{m\geq 2} \frac{(1/2)^{m-2}}{m}$, and we are done. Note that this shows that $\log(I+B) = B + O(\|B\|^2)$, i.e., for $\|B\|^2$ sufficiently small, we have the $O(\|B\|^2)$ portion $\leq c\|B\|^2$ with the choice of c independent of B. □

Theorem 4.3 (Lie product formula)**.** *Let $X, Y \in \mathrm{M}_n(\mathbb{C})$. Then*

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

That is, this identity holds even when X and Y do not *commute.*

Proof. Start with

$$\begin{aligned} e^{X/m} &= I + \frac{X}{m} + \frac{1}{m^2}\frac{X^2}{2} + \frac{1}{m^3}\frac{X^3}{6} + \cdots \\ &= I + \frac{X}{m} + \frac{1}{m^2} \underbrace{\left(\frac{X^2}{2} + \frac{1}{m}\frac{X^3}{6} + \cdots \right)}_{(*)}. \end{aligned}$$

We claim that $(*)$ is bounded, and that the bound depends on X but *not* on m. Note that:

$$\|(*)\| \leq \frac{\|X\|^2}{2} + \frac{1}{m}\frac{\|X\|^3}{6} + \cdots \leq e^{\|X\|} - \|I\| - \|X\| \leq e^{\|X\|}.$$

Note that for m^{-2} sufficiently small (in fact, for all $m \geq 1$) we have $e^{X/m} = I + \frac{X}{m} + O(\frac{1}{m^2})$ and the similar claim can be made for $e^{Y/m}$. Hence,

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

Since $e^{X/m}e^{Y/m} \to I$ as $m \to \infty$, it will be in the domain of log as long as m is sufficiently large. Apply Proposition 4.2 after taking logs on both sides:

$$\begin{aligned}\log(e^{X/m}e^{Y/m}) &= \log\left(I + \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right) \\ &= \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right) + \underbrace{O\left(\left\|\frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right\|\right)}_{\leq \widetilde{c}m^{-2}},\end{aligned}$$

where $\widetilde{c}$ does not depend on m but on X and Y. Hence

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

Exponentiate both sides and take m-th powers, i.e.,

$$e^{X/m}e^{Y/m} = e^{X/m+Y/m+O(m^{-2})},$$

hence

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

Since e is continuous, upon making $m \to \infty$, we have

$$\lim_{m\to\infty}(e^{X/m}e^{Y/m})^m = e^{X+Y},$$

as required. □

Theorem 4.4. *For all* $X \in \mathrm{M}_n(\mathbb{C})$*, we have* $\det(e^X) = e^{\mathrm{tr}(X)}$*.*

Proof. Suppose when X is diagonalizable. Then we have, for some P,

$$PXP^{-1} = P\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}P^{-1} = D,$$

or $X = PDP^{-1}$. So we have

$$\det(e^X) = \det(Pe^DP^{-1}) = \det(P)\det(e^D)\det(P^{-1}) = e^{\lambda_1+\cdots+\lambda_n} = e^{\mathrm{tr}(D)} = e^{\mathrm{tr}(X)}. \qquad \square$$

For the general case, use the fact that the set of diagonalizable matrices is dense in $\mathrm{M}_n(\mathbb{C})$ and that the exponential map and the determinant map are continuous.

Definition 4.5. A map $\alpha : \mathbb{R} \to \mathrm{GL}(n, \mathbb{C})$ is called a *1-parameter subgroup of* $\mathrm{GL}(n, \mathbb{C})$ if:

(1) α is continuous; and
(2) $\alpha(s + t) = \alpha(s)\alpha(t)$ for all $s, t \in \mathbb{R}$.

In other words, α is a continuous group homomorphism.

Remark 4.1. By property (2), we have $\alpha(0) = I$ and $\alpha(-t) = (\alpha(t))^{-1}$.

Remark 4.2. If $x \in \mathrm{M}_n(\mathbb{C})$, then $\alpha(t) = e^{tX}$ is an example of a 1-parameter subgroup.

Theorem 4.6. *If α is a 1-parameter subgroup, then there exists a* unique *$X \in \mathrm{M}_n(\mathbb{C})$ such that $\alpha(t) = e^{tX}$ for all $t \in \mathbb{R}$.*

Proof. Suppose that there exists X, Y such that $\alpha(t) = e^{tX} = e^{tY}$. Then

$$\left.\frac{d}{dt}\right|_{t=0} \alpha(t) = X = Y$$

so the uniqueness is proved. Now we need to show existence.

Let $B_\varepsilon(0) := \{Y \in \mathrm{M}_n(\mathbb{C}) : \|Y\| < \varepsilon\}$, and let $\varepsilon < \log 2$. We have shown that exponentiation takes $B_\varepsilon(0)$ bijectively onto $\exp(B_\varepsilon(0))$ with continuous inverse log map from $\exp(B_\varepsilon(0))$ to $B_\varepsilon(0)$. Now let $U := \exp(B_{\varepsilon/2}(0)) = \log^{-1}(B_{\varepsilon/2}(0))$. Then U is open.

We claim that every $B \in U$ has a unique square root in U, given by $\exp(\frac{1}{2}\log B)$. Let $Y := \log(B)$, with $\|B\| < \varepsilon/2$ so that $\exp(\frac{1}{2}\log B) \in U$. Then $Z := \exp(\frac{1}{2}Y) = \exp(\frac{1}{2}\log B)$. Then clearly we have $Z^2 = (\exp(\frac{1}{2}B))^2 = \exp(Y) = B$. Hence Z is indeed a square root of B. For uniqueness, suppose that Z' is another square root of B, with $Z' \in U$. Since $(Z')^2 = B$ and $\exp(\log Z') = Z'$, we have

$$\exp(2\log Z') = (\exp(\log Z'))^2 = (Z')^2 = B = \exp(Y).$$

Since $Z' \in U$, we have $\log Z' \in B_{\varepsilon/2}(0)$, and $2\log Z' \in B_\varepsilon(0)$. Also, $Y \in B_{\varepsilon/2}(0) \subseteq B_\varepsilon(0)$ and exp is injective on $B_\varepsilon(0)$. Hence $Y = 2\log Z'$ so $Z' = \exp(\frac{1}{2}Y) = \exp(\frac{1}{2}\log B) = Z$, as required.

Now, since α is continuo and B is in the neighbourhood of I, there exists some $t_0 > 0$ such that $\alpha(t) \in U$ for all $|t| \leq t_0$ and $\alpha(0) = I$. Let $X = \frac{1}{t_0}\log(\alpha(t_0))$ so $t_0 X = \log(\alpha(t_0))$. So $t_0 X \in B_{\varepsilon/2}(0)$ and $\alpha(t_0) = e^{t_0 X}$. Note that $\alpha(t_0/2) \in U$ and $(\alpha(t_0/2))^2 = \alpha(t_0/2)\alpha(t_0/2) = \alpha(t_0) = e^{t_0 X}$, by property (2) of an 1-parameter subgroup. So by the claim we just proved, $\alpha(t_0/2) = e^{\frac{t_0}{2}X}$. Repeat this operation to see that

$$\alpha\left(\frac{t_0}{2^k}\right) = e^{\frac{t_0}{2^k}X}$$

for all $k \in \mathbb{N}$. Take m-th powers and use property (2) to get

$$\alpha\left(\frac{m}{2^k}t_0\right) = e^{\frac{mt_0}{2^k}X}$$

for all $k, m \in \mathbb{N}$. Now $\{\frac{m}{2^k}t_0 : m, k \in \mathbb{N}\}$ is dense in $\mathbb{R}$ and α and exp are continuous, so $\alpha(t) = e^{tX}$ for all t as desired. □

Remark 4.3. We needed the following ingredients to prove the previous theorem:

(1) The fact that α is continuous *and* is a homomorphism
(2) the fact that exp and log are continuous inverses of each other near 0 and I respectively.

Now that we identified all the 1-parameter subgroups, we are ready to define Lie algebras.

4.1. *Lie algebra of a matrix Lie group*

Definition 4.7. Let G be a matrix Lie group. Then the *Lie algebra of* G, denoted $\mathfrak{g}$, is the set

$$\mathfrak{g} = \{X \in \mathrm{M}_n(\mathbb{C}) : e^{tX} \in G \text{ for all } t \in \mathbb{R}\},$$

i.e., $X \in \mathfrak{g} \Leftrightarrow$ the 1-parameter subgroup $\{e^{tX} : t \in \mathbb{R}\}$ generated by X lies in G.

Remark 4.4. We cannot say that $X \in \mathfrak{g} \Leftrightarrow e^X \in G$. One can find examples where $e^X \in G$ but $e^{tX} \notin G$ for some $t \in \mathbb{R}$. Those X are *not* in $\mathfrak{g}$.

Also, we cannot say that $X \in \mathfrak{g} \Leftrightarrow e^{tX} \in G$ for all $t \in \mathbb{C}$. A Lie algebra for which this is true is called a *complex Lie algebra.*

Remark 4.5. Every matrix lie group G is an embedded submanifold of $\mathrm{GL}(n, \mathbb{C})$ and it follows that it Lie algebra $\mathfrak{g}$ is isomorphic to the tangent space to G at I, i.e., $T_I G \cong \mathfrak{g}$.

Example 4.8. Some examples of Lie algebras:

(1) $\mathrm{GL}(n, \mathbb{C})$
If $X \in \mathrm{M}_n(\mathbb{C})$, then $e^{tX} \in \mathrm{GL}(n, \mathbb{C})$ for all $t \in \mathbb{R}$. Hence the Lie algebra of $\mathrm{GL}(n, \mathbb{C})$ is $\mathfrak{gl}(n, \mathbb{C}) = \mathrm{M}_n(\mathbb{C})$.

(2) $\mathrm{GL}(n, \mathbb{R})$
If $X \in \mathrm{M}_n(\mathbb{R})$, then $e^{tX} \in \mathrm{GL}(n, \mathbb{R})$ for all $t \in \mathbb{R}$. Conversely, if $e^{tX} \in \mathrm{GL}(n, \mathbb{R})$ for all $t \in \mathbb{R}$, then

$$X = \left.\frac{d}{dt}\right|_{t=0} e^{tX} \in \mathrm{M}_n(\mathbb{R}).$$

Hence we have $\mathfrak{gl}(n, \mathbb{R}) = \mathrm{M}_n(\mathbb{R})$. In particular, we have shown that if G is a subgroup of $\mathrm{GL}(n, \mathbb{R})$, then its Lie algebra $\mathfrak{g}$ consists of real matrices.

(3) $\mathrm{SL}(n, \mathbb{C})$
Recall that $\det(e^X) = e^{\mathrm{tr}(X)}$. Hence $\det(e^{tX}) = e^{t \cdot \mathrm{tr}(X)}$ for all $t \in \mathbb{R}$ since tr is linear. So if $e^{tX} \in \mathrm{SL}(n, \mathbb{C})$ for all $t \in \mathbb{R}$ then $e^{t \cdot \mathrm{tr}(X)} = 1$ for all $t \in \mathbb{R}$. Hence $\mathrm{tr}(X) = 0$. Conversely, if $\mathrm{tr}(X) = 0$, then $\det(e^{tX}) = e^{t \cdot \mathrm{tr}(X)} = e^0 = 1$ for all t. Thus $e^{tX} \in \mathrm{SL}(n, \mathbb{C})$ for all $t \in \mathbb{R}$. It follows that

$$\mathfrak{sl}(n, \mathbb{C}) = \{X \in \mathfrak{gl}(n, \mathbb{C}) : \mathrm{tr}(X) = 0\}.$$

Similarly, in the real case, we have

$$\mathfrak{sl}(n, \mathbb{R}) = \{X \in \mathfrak{gl}(n, \mathbb{R}) : \mathrm{tr}(X) = 0\}.$$

(4) $\mathrm{U}(n) = \{A \in \mathrm{GL}(n, \mathbb{C}) : A^* = A^{-1}\}$
Let $e^{tX} \in \mathrm{U}(n)$ for all $t \in \mathbb{R}$, i.e.

$$(e^{tX})^* = e^{tX^*} = e^{-tX} = (e^{tX})^{-1}$$

for all $t \in \mathbb{R}$. So

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

Conversely, if $X^* = -X$ then $(e^{tX})^* = (e^{tX})^{-1}$ for all $t \in \mathbb{R}$. Hence

$$\mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) : X^* = -X\}.$$

Thus $\mathfrak{u}(n)$ consists of *skew-Hermitian matrices*, i.e., the matrices with diagonal entries purely imaginary. Simiarly, in the case of $\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n, \mathbb{C})$, we have

$$\mathfrak{su}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) : X^* = -X \text{ and } \operatorname{tr}(X) = 0\},$$

so $\mathfrak{su}(n)$ consists of traceless skew-Hermitian matrices.

5. January 19

Recall the definition of Lie algebra first:

Definition 5.1. $\mathfrak{g}$ is the *Lie algebra* of a matrix Lie group $G \subset \mathrm{GL}(n, \mathbb{C})$ if

$$\mathfrak{g} = \{X \in \mathrm{GL}(n, \mathbb{C}) : e^{tX} \in G \,\forall t \in \mathbb{R}\}.$$

Remark 5.1. If $X \in \mathfrak{g}$, then $e^X \in G_0$, where G_0 denotes a connected component of I in G, since $e^{tX} \in G$ for all t, and we have a path from I to e^X. Moreover, $X \in \mathfrak{g} \Leftrightarrow e^{tX} \in G$ for all $t \in \mathbb{R} \Leftrightarrow e^{tX} \in G_0$ for all $t \in \mathbb{R}$. Therefore, the Lie algebra $\mathfrak{g}$ of G is the same as the Lie algebra of G_0.

Example 5.2. If $G = \mathrm{O}(n) = \{A \in \mathrm{GL}(n, \mathbb{R}) : A^T A = I\}$, or $A^T = A^{-1}$. Suppose that $e^{tX} \in \mathrm{O}(n)$ for all $t \in \mathbb{R}$. So it follows that $e^{-tX} = e^{tX^T} = (e^{tX})^T = (e^{tX})^{-1}$. Therefore, the Lie algebra of G is

$$\mathfrak{o}(n) = \{X \in \mathfrak{gl}(n, \mathbb{R}) : X^T = -X\}.$$

that is, $\mathfrak{o}(n)$ consists of skew-symmetric matrices. Therefore, necessarily $\operatorname{tr}(X) = 0$. Therefore if $X \in \mathfrak{o}(n)$, then $\det(e^{tX}) = e^{t \operatorname{tr}(X)} = e^{t \cdot 0} = 1$ for all t. Hence $e^{tX} \in \mathrm{SO}(n)$ for all $t \in \mathbb{R}$, so $\mathfrak{so}(n) = \mathfrak{o}(n) = \{X \in \mathfrak{gl}(n, \mathbb{R}) : X^T = -X\}$. Therefore $\mathfrak{so}(n) = \mathfrak{o}(n)$ (in fact, $\mathrm{SO}(n) = \mathrm{O}(n)$.).

Example 5.3. We also claim that $\mathrm{Lie}(\mathrm{SO}(p, q)) = \mathrm{Lie}(\mathrm{O}(p, q))$ where $\mathrm{Lie}(A)$ denotes the Lie algebra of A. If $A \in \mathrm{O}(p, q)$, then $A^T I_{p,q} A = I_{p,q}$, or $I_{p,q} A^T I_{p,q}^{-1} = A^{-1}$. So if $e^{tX} \in \mathrm{O}(p, q)$, then $I_{p,q} (e^{tX})^T I_{p,q}^{-1} = (e^{tX})^{-1} = e^{t(I_{p,q} X^T I_{p,q}^{-1})} = e^{-tX}$. Hence $\mathfrak{so}(p, q) = \mathfrak{o}(p, q) = \{X \in \mathfrak{gl}(p + q, \mathbb{R}) : I_{p,q} X^T I_{p,q} = -X\}$.

Example 5.4. Recall that $\mathrm{Sp}(n, \mathbb{F}) : \{A \in \mathrm{GL}(n, \mathbb{F}) : A^T J A = J\}$, where

$$J := \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}.$$

So $A \in \mathrm{Sp}(n, \mathbb{F}) \Leftrightarrow J A^T J^{-1} = -A^{-1}$. So if $e^{tX} \in \mathrm{Sp}(n, \mathbb{F})$ then $J (e^{tX})^T J = e^{tJX^t J} = -(e^{tX})^{-1} = -e^{-tX}$ for all $t \in \mathbb{F}$. That is, $X \in \mathfrak{sp}(n, \mathbb{F}) \Leftrightarrow J X^T J = X$. Thus $\mathfrak{sp}(n, \mathbb{F}) = \{X \in \mathfrak{gl}(2n, \mathbb{F}) : J X^T J = X\}$. More generally,

$$\mathfrak{sp}(n) = \mathfrak{sp}(n, \mathbb{C}) \cap \mathfrak{u}(2n) = \{X \in \mathfrak{gl}(2n, \mathbb{C}) : J X^T J = X, X^* = -X\}.$$

Remark 5.2. If $X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, then $X \in \mathfrak{sp}(n, \mathbb{F})$ if and only if $D = -A^T$ and B, C symmetric.

Example 5.5. Suppose that H_n is a generalized Heisenberg group. Let X be an upper triangular matrix with 0 on the diagonals. It is clear that $e^{tX} \in H_n$ for all $t \in \mathbb{R}$. Conversely, if $e^{tX} \in H_n$ for all $t \in \mathbb{R}$ then X is an upper triangular matrix with 0 on the diagonals. Hence the Lie algebra of H_n is

$$\mathfrak{h}_n = \mathrm{Lie}(H_n) = \left\{ \begin{bmatrix} 0 & & & * \\ 0 & 0 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 0 \end{bmatrix} \right\}$$

Now we are ready to discuss some properties of Lie algebras.

Proposition 5.6. *Let G be a matrix Lie groups with Lie algebra $\mathfrak{g}$. Let $X \in \mathfrak{g}$ and $A \in G$. Then $AXA^{-1} \in \mathfrak{g}$. That is the conjugation by an element of G preserves $\mathfrak{g}$.*

Proof. One-line proof: $e^{t(AXA^{-1})} = Ae^{tX}A^{-1} \in G$ for all $t \in \mathbb{R}$. ☐

Theorem 5.7. $\mathfrak{g}$ *is a* real vector space. *That is,*

(1) *if $X \in \mathfrak{g}$ then $tX \in \mathfrak{g}$ for all $t \in \mathbb{R}$*
(2) *if $X, Y \in \mathfrak{g}$ then $X + Y \in \mathfrak{g}$.*

Additionally, we also have

(3) *if $X, Y \in \mathfrak{g}$ then $XY - YX \in \mathfrak{g}$.*

Proof. Let $t \in \mathbb{R}$. Since $e^{s(tX)} = e^{(st)X} \in G$ for all $s \in \mathbb{R}$, we have $tX \in \mathfrak{g}$. For (2), we divide into two cases. If $XY = YX$, then $e^{t(X+Y)} = e^{tX}e^{tY} \in G$ for all $t \in \mathbb{R}$. Hence $X + Y \in \mathfrak{g}$. If X and Y do not commute, then note that $e^{tX/m}e^{tY/m} \in G$ hence $(e^{tX/m}e^{tY/m})^m \in G$ also, for all nonzero m. Any matrix exponential is invertible, so $e^{t(X+Y)} \in \mathrm{GL}(n, \mathbb{C})$. Apply the Lie product formula to get

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

and since G is closed in $\mathrm{GL}(n, \mathbb{C})$, it follows that either the limit is not invertible or in G. Hence the limit is in G.

As for (3), recall that

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

Hence

$$\left.\frac{d}{dt}\right|_{t=0} e^{tX}Ye^{-tX} = XYe^{-0x} + e^{0x}Y(-X) = XY - YX.$$

If we let $Z(t) := e^{tX}Ye^{-tX}$, then $Z(t)$ is the conjugation of an element of $\mathfrak{g}$ (Y) by an element of G (namely, e^{tX}). Hence $Z(t) \in \mathfrak{g}$ for all $t \in \mathbb{R}$, and

$$\left.\frac{d}{dt}\right|_{t=0} Z(t) = XY - YX \in \mathfrak{g},$$

since $\mathfrak{g}$ is a subspace of $\mathfrak{gl}(n, \mathbb{C})$ and any subspace of a finite-dimensional $\mathbb{C}$-vector space is closed. ☐

Definition 5.8. Given $X, Y \in \mathfrak{gl}(n, \mathbb{C})$, their *Lie bracket* or *commutator* $[X, Y]$ is defined to be $[X, Y] = XY - YX$.

Remark 5.3. We showed that the Lie algebra $\mathfrak{g}$ of G is a real vector space that is closed under the Lie bracket, i.e., $X, Y \in \mathfrak{g} \Rightarrow [X, Y] \in \mathfrak{g}$.

Remark 5.4. Even if $G \in \mathrm{GL}(n, \mathbb{C})$ as opposed to $\mathrm{GL}(n, \mathbb{R})$, the Lie algebra $\mathfrak{g}$ of G is only a *real* vector space.

Example 5.9. Recall that $\mathrm{U}(n) \subset \mathrm{GL}(n, \mathbb{C})$ while $\mathrm{U}(n) \not\subset \mathrm{GL}(n, \mathbb{R})$. However, $\mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) : X^* = -X\}$, so if $X \in \mathfrak{u}(n)$ then $(iX)^* = -iX^* = (-i)(-X) = iX \neq -iX$. Therefore $iX \notin \mathfrak{u}(n)$ so $\mathfrak{u}(n)$ is not subspace of the n-dimensional $\mathbb{C}$-vector space.

Definition 5.10. The Lie algebra $\mathfrak{g}$ of a matrix Lie group G is a *complex Lie algebra* if $X \in \mathfrak{g} \Rightarrow iX \in \mathfrak{g}$ (i.e. $\mathfrak{g}$ is a *complex* subspace of $\mathfrak{gl}(n, \mathbb{C})$). In this case, we say G is a *complex Lie group*.

Example 5.11. $\mathrm{GL}(n, \mathbb{C}), \mathrm{SL}(n, \mathbb{C}), \mathrm{SO}(n, \mathbb{C}), \mathrm{O}(n, \mathbb{C}), \mathrm{Sp}(n, \mathbb{C})$ are complex Lie groups. On the other hand, $\mathrm{GL}(n, \mathbb{R}), \mathrm{SL}(n, \mathbb{R}), \mathrm{Sp}(n, \mathbb{R}), \mathrm{O}(n), \mathrm{SO}(n), \mathrm{O}(p, q), \mathrm{U}(n), \mathrm{SU}(n), \mathrm{Sp}(n)$ are not complex Lie groups.

Remark 5.5. Homomorphisms of matrix Lie groups induce homomorphisms of their Lie algebras.

Theorem 5.12. *Let G, H be matrix Lie groups with Lie algebras $\mathfrak{g}, \mathfrak{h}$ respectively, and let $\Phi : G \to H$ be a matrix Lie group homomorphism. Then there exists a* unique real linear map $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ *such that*

$$\Phi(e^X) = e^{\Phi_*(X)} \text{ for all } X \in \mathfrak{g}.$$

Φ_* *also satisfies*

1. *(1)* $\Phi_*(AXA^{-1}) = \Phi(A)\Phi_*(X)\Phi(A)^{-1}$ *for all* $A \in G, X \in \mathfrak{g}$
2. *(2)* $\Phi_*([X, Y]) = [\Phi_* X, \Phi_* Y]$ *(Lie algebra homomorphism)*
3. *(3)* $\Phi_*(X) = \frac{d}{dt}\big|_{t=0} \Phi(e^{tX})$ *for all* $X \in \mathfrak{g}$.
4. *(4) If* $\Psi : H \to K$ *is another Lie group homomorphism, then* $\Psi \circ \Phi : G \to K$ *and* $(\Psi \circ \Phi)_* = \Psi_* \circ \Phi_*$.

Proof. Since Φ is a group homomorphism and continuous, the map $t \in \mathbb{R} \overset{\Psi}{\mapsto} \Phi(e^{tX}) \in H$ is a 1-parameter subgroup. Use the fact that Φ is a group homomorphism to get $\Phi(e^{tX})\Phi(e^{sX}) = \Phi(e^{tX}e^{sX}) = \Phi(e^{(t+s)X})$. Similarly, $\Psi(s)\Psi(t) = \Psi(s + t)$. So by last week's theorem, there exists a unique matrix $W \in \mathfrak{gl}(n, \mathbb{C})$ such that $\Phi(e^{tX}) = e^{tW}$ for all $t \in \mathbb{R}$.

$\Phi(e^{tX}) \in H$ so $e^{tW} \in H$ for all $t \in \mathbb{R}$, which impiles that $W \in \mathfrak{h}$. So define $\Phi_*(X) = W$. So we have (3). Indeed,

$$\Phi_*(X) = W = \frac{d}{dt}\bigg|_{t=0} e^{tW} = \frac{d}{dt}\bigg|_{t=0} \Phi(e^{tX}).$$

$\Phi(e^{tX}) = e^{t\Phi_*(X)}$ for all t, so $\Phi(e^{(ts)\Phi_*(X)}) = \Phi(e^{(ts)X}) = \Phi(e^{t(sX)}) = e^{t\Phi_*(sX)}$. Take the derivative at $t = 0$ to get

$$\Phi_*(sX) = s\Phi_* X \text{ for all } s \in \mathbb{R}. \tag{$*$}$$

From $\Phi(e^{tX}) = e^{t\Phi_*(X)}$ at $t = 1$, we have $\Phi(e^X) = e^{\Phi_*(X)}$. Also,

$$\begin{aligned}
e^{t(\Phi_*(X+Y))} &= e^{\Phi_*(tX+tY)} \text{ (by } (*)) \\
&= \Phi(e^{tX+tY}) = \Phi\left(\lim_{m\to\infty}\left(e^{tX/m}e^{tY/m}\right)^m\right) \\
&= \lim_{m\to\infty}\Phi\left(\left(e^{tX/m}e^{tY/m}\right)^m\right) \text{ (}\Phi\text{ is a continuous homomorphism)} \\
&= \lim_{m\to\infty}\left(\Phi(e^{tX/m})\Phi(e^{tY/m})\right)^m = \lim_{m\to\infty}\left(e^{\frac{t}{m}\Phi_*X}e^{\frac{t}{m}\Phi_*Y}\right)^m \\
&= e^{t(\Phi_*X+\Phi_*Y)} \text{ (by the Lie product formula).}
\end{aligned}$$

For (1), note that

$$\begin{aligned}
e^{t\Phi_*(AXA^{-1})} = e^{\Phi_*(tAXA^{-1})} = \Phi(e^{tAXA^{-1}}) &= \Phi(Ae^{tX}A^{-1}) \\
&= \Phi(A)\Phi(e^{tX})\Phi(A)^{-1}.
\end{aligned}$$

Take d/dt at $t = 0$ to get $\Phi_*(AXA^{-1}) = \Phi(A)\Phi_*X\Phi(A)^{-1}$, which proves (1).

As for (2), recall that

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

So

$$\begin{aligned}
\Phi_*[X, Y] &= \Phi_*\left(\left.\frac{d}{dt}\right|_{t=0} e^{tX}Ye^{-tX}\right) \\
&= \left.\frac{d}{dt}\right|_{t=0} \Phi_*(e^{tX}Ye^{-tX}) \text{ (differentiation commutes with linear maps)} \\
&= \left.\frac{d}{dt}\right|_{t=0} \Phi_*(e^{tX}Ye^{-tX}) \text{ (by (1))} \\
&= \left.\frac{d}{dt}\right|_{t=0} e^{t\Phi_*X}\Phi_*Ye^{-t\Phi_*X} \\
&= [\Phi_*X, \Phi_*Y],
\end{aligned}$$

thereby proving (2).

For (4), note that

$$\begin{aligned}
e^{t(\Psi\circ\Phi)_*X} = (\Psi\circ\Phi)(e^{tX}) &= \Psi(\Phi(e^{tX})) \\
&= \Psi(e^{t\Phi_*X}) = e^{t\Psi_*\Phi_*X}.
\end{aligned}$$

Take d/dt at $t = 0$ to get $(\Psi \circ \Phi)_*X = \Psi_*(\Phi_*X)$ for all X, as required. □

Remark 5.6. We have shown that $\Phi : G \to H$ matrix Lie group homomorphism induces $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ Lie group homomorphism.

Question. Suppose that $\mathfrak{g}, \mathfrak{h}$ is a Lie group homomorphism (i.e., $\mathbb{R}$-linear, $\lambda[X, Y] = [\lambda(X), \lambda(Y)]$ if $\lambda : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism). Then does there exist a Lie group homomorphism $\Phi : G \to H$ such that $\Phi_* = \lambda$?

The answer is actually *no*. We will prove that a *sufficient condition* is that G is simply connected.

6. January 21

Recall that if $\Phi : G \to H$ is a Lie group homomorphism, then it induces a Lie algebra homomorphism $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ such that the diagram

$$\begin{array}{ccc} G & \xrightarrow{\Phi} & H \\ {\scriptstyle \exp}\uparrow & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[\Phi_*]{} & \mathfrak{h} \end{array}$$

commutes.

6.1. *The Adjoint mapping*

Definition 6.1. Let G be a matrix Lie group and $\mathfrak{g}$ its Lie algebra. Fix $A \in G$. If $x \in \mathfrak{g}$, then $AXA^{-1} \in \mathfrak{g}$. Define

$$\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$$

by $\mathrm{Ad}_A(X) = AXA^{-1}$. Then Ad_A is R-linear and Ad_A is invertible, because $\mathrm{Ad}_{A^{-1}} = (\mathrm{Ad}_A)^{-1}$. Therefore $\mathrm{Ad}_A \in \mathrm{GL}(\mathfrak{g})$, the set of invertible linear operators on $\mathfrak{g}$.

Proposition 6.2. *The map from G to $\mathrm{GL}(\mathfrak{g})$ defined by $A \mapsto \mathrm{Ad}_A$ is a matrix Lie group homomorphism.*

Proof. The map is clearly continuous (linear maps over finite-dimensional spaces). It is a homomorphism since

$$\mathrm{Ad}_{AB}(X) = (AB)X(AB)^{-1} = A(BXB^{-1})A^{-1} = (\mathrm{Ad}_A \circ \mathrm{Ad}_B)(X),$$

i.e., $\mathrm{Ad}_{AB} = \mathrm{Ad}_A \circ \mathrm{Ad}_B$. □

By last class, there exists an induced map $(\mathrm{Ad})_* : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$, where $\mathfrak{gl}(\mathfrak{g})$ denotes the Lie algebra of $\mathrm{GL}(\mathfrak{g})$, the space of all linear operators on $\mathfrak{g}$.

Proposition 6.3. *For all A, $\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra homomorphism.*

Proof. This is a straightforward computation:

$$\begin{aligned} \mathrm{Ad}_A[X, Y] &= A(XY - YX)A^{-1} \\ &= (AXA^{-1})(AYA^{-1}) - (AYA^{-1})(AXA^{-1}) \\ &= [\mathrm{Ad}_A X, \mathrm{Ad}_A Y]. \end{aligned}$$

□

So Ad induces a map Ad_* so that the following diagram commutes:

$$\begin{array}{ccc} G & \xrightarrow{\mathrm{Ad}} & \mathrm{GL}(\mathfrak{g}) \\ {\scriptstyle \mathrm{Ad}_*}\uparrow & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[(\mathrm{Ad})_*]{} & \mathfrak{gl}(\mathfrak{g}) \end{array}$$

Now we are ready to define the little ad:

Definition 6.4. We define the little "ad" as $\mathrm{ad} := (\mathrm{Ad})_* : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. That is, ad is the map such that $\exp(\mathrm{ad}_X) = \mathrm{Ad}(\exp X)$, or $e^{\mathrm{ad}_X} = \mathrm{Ad}(e^X)$.

Proposition 6.5. $\mathrm{ad}_X(Y) = [X, Y]$.

Proof. Note that $\mathrm{ad}_X = \frac{d}{dt}\big|_{t=0} \mathrm{Ad}(e^{tX})$, by definition of induced linear map on Lie algebras. Therefore,

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

by what we did last class. □

Corollary 6.6. $(e^{\mathrm{ad}_X})Y = \displaystyle\sum_{m=0}^{\infty} \frac{(\mathrm{ad}_X)^m}{m!} Y.$

Proof. Note that

$$\sum_{m=0}^{\infty} \frac{(\mathrm{ad}_X)^m}{m!} Y = Y + [X, Y] + \frac{1}{2}[X, [X, Y]] + \cdots = \mathrm{Ad}(e^X)Y = e^X Y e^{-X}.$$ □

6.2. *Detour to exponential mapping*

Remark 6.1. Recall that $\exp : \mathfrak{g} \to G$ is defined to be $\exp(X) = e^X$, and it is continuous in X. It exp injective or surjective? Unfortunately, the answer is *no* in general (i.e., neither injective nor surjective). The following example illustrates this point.

Example 6.7. We already know that $e^X \in G_0$, the connected component of identity. Therefore $\mathfrak{g} \overset{\exp}{\to} G$ cannot be surjective if G is not connected. For example, note that $\mathfrak{so}(n) = \mathfrak{o}(n)$, and the image of $\exp : \mathfrak{o}(n) \to \mathrm{O}(n)$ lies in $\mathrm{SO}(n) \subsetneq \mathrm{O}(n)$.

Example 6.8. However, even if G *is* connected, exp is in general still *not surjective*. Let $G := \mathrm{GL}(2, \mathbb{C})$, which is clearly connected. Let

$$A = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \in \mathrm{SL}(2, \mathbb{C}).$$

We claim that there cannot exist any $X \in \mathfrak{sl}(2, \mathbb{C}) = \{X \in \mathfrak{gl}(2, \mathbb{C}), \operatorname{tr} X = 0\}$ such that $e^X = A$. For this, consider the following cases:

<u>Case 1.</u> $\lambda_1 = \lambda_2 = 0$. Every 2×2 matrix in triangularizable over $\mathbb{C}$, i.e., there exists P such that $X = P \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix} P^{-1}$. Therefore, $e^X = P e^{\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}} P^{-1} = P \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} P^{-1}$. But since $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ has both eigenvalues 1, it cannot be A.

<u>Case 2.</u> $\lambda_1 + \lambda_2 = 0$ and $\lambda_1 \neq 0$, i.e., $(\lambda_1, \lambda_2) = (\lambda, -\lambda)$.

Then X has distinct eigenvalues, so X is diagonalizable. Hence there must exist P such that

$$P^{-1}XP = \begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix}$$
$$X = P\begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix}P^{-1}$$
$$e^X = P\begin{bmatrix} e^\lambda & 0 \\ 0 & e^{-\lambda} \end{bmatrix}P^{-1}.$$

Therefore e^X is diagonalizable. But A is not diagonalizable, hence a contradiction.

Example 6.9. We remark also that $\exp : \mathfrak{g} \to G$ is generally not injective either. Consider the following exponential map from $\mathfrak{so}(2)$ to $\mathrm{SO}(2)$: $\begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix} \overset{\exp}{\to} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. This is clearly not injective!

We will see however that $\exp : \mathfrak{g} \to G$ is a homeomorphism from an open neighbourhood U of 0 in $\mathfrak{g}$ to an open neighbourhood $V = \exp(U)$ of I in G.

Theorem 6.10. *For* $0 < \varepsilon < \log 2$ *(we need this to guarantee that* $\log A$ *is defined on* V_ε *and* $\log(e^X) = X$*), let* $U_\varepsilon := \{X \in \mathfrak{gl}(n, \mathbb{C}) : \|X\| < \varepsilon\}$ *and* $V_\varepsilon := \exp(U_\varepsilon)$*. Let* $G \subseteq \mathrm{GL}(n, \mathbb{C})$ *be a matrix Lie group with Lie algebra* $\mathfrak{g}$*. Then there exists* $\varepsilon_0 \in (0, \log 2)$ *such that for all* $A \in V_\varepsilon$ *with* $A \in G$ *if and only if* $\log A \in \mathfrak{g}$*. Hence,* $\exp : U_{\varepsilon_0} \to V_{\varepsilon_0}$ *is a homeomorphism.*

Proof. $(\Rightarrow)$ First, we need the following claim:

Claim. Let $B_m \in G$ for all $m \in \mathbb{N}$ such that B_m converges to I, and let $Y_m := \log B_m$, which is well-defined for m sufficiently large. up pose that $Y_m \neq 0$ for all m, and $\frac{Y_m}{\|Y_m\|} \to Y \in \mathfrak{gl}(n, \mathbb{C})$. Then $Y \in \mathfrak{g}$.

Proof of Claim. To show that $Y \in \mathfrak{g}$, we need to show that $e^{tY} \in G$ for all $t \in \mathbb{R}$. Let $t \in \mathbb{R}$. Then $\frac{t}{\|Y_m\|}Y_m \to tY$ as $m \to \infty$, but $B_m \to I$. Therefore $Y_m \to 0$ and $\|Y_m\| \to 0$. Therefore there exists a sequence $k_m \in \mathbb{Z}$ (depends on t) such that $k_m\|Y_m\| \to T$ as $m \to \infty$. Then

$$(B_m)^{k_m} = (e^{Y_m})^{k_m} = e^{k_m Y_m} = e^{k_m\|Y_m\|\frac{Y_m}{\|Y_m\|}} \to e^{tY},$$

so indeed $(B_m)^{k_m} \in G$ for all m. Since G is closed in $\mathrm{GL}(n, \mathbb{C})$, it follows that $e^{tY} \in G$. So $Y \in \mathfrak{g}$ as required. □

To finish off the proof, we start by observing that $\mathfrak{gl}(n, \mathbb{C}) \cong \mathbb{C}^{n^2} \cong \mathbb{R}^{2h^2}$ with usual topology, and that $\mathfrak{g}$ is a real subspace of $\mathbb{R}^{2n^2}$. Decompose $\mathfrak{gl}(n, \mathbb{C}) = \mathfrak{g} \oplus \mathfrak{g}^\perp$, where $\mathfrak{g}^\perp$ is the orthogonal complement with respect to the usual inner product. Define a map $F : \mathfrak{g} \oplus \mathfrak{g}^\perp = \mathfrak{gl}(n, \mathbb{C}) \to \mathrm{GL}(n, \mathbb{C})$ as $F(X, Y) = e^X e^Y$. Then $F : \mathbb{R}^{2n^2} \to \mathbb{R}^{2n^2}$ is a smooth map. Since

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

we have

$$(DF)|_0 : \mathbb{R}^{2n^2} \to \mathbb{R}^{2n^2}$$

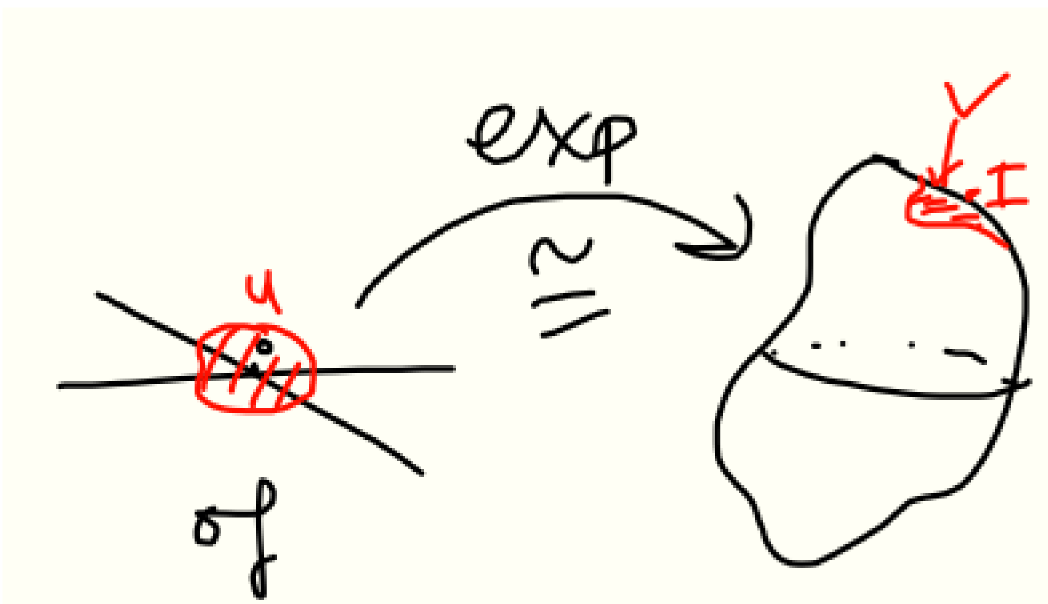
is the identity. In particular, $(DF)|_0$ is invertible, so by the inverse function theorem, there exists a neighbourhood $0 \in U$ such that $V = F(U)$ is a neighbourhood of I and $F : U \to V$ is a homeomorphism (in fact a *diffeo*morphism).

Now we need to show that there exists $\varepsilon_0 \in (0, \log 2)$ such that $A \in V_\varepsilon \cap G \Rightarrow \log A \in \mathfrak{g}$. Suppose not. That is, for all $m \in \mathbb{N}$ and for $\varepsilon = m^{-1}$, there exists $A_m \in V_{\frac{1}{m}} \cap G$ such that $\log A_m \notin \mathfrak{g}$. Using the local inverse for F, if m is sufficiently large then $A_m \approx I$ (A_m is close to I), so $A_m = e^{X_m} e^{Y_m}$ where $X_m \in \mathfrak{g}$ and $Y_m \in \mathfrak{g}^\perp$ such that $X_m \to 0$ and $Y_m \to 0$ (since $A_m \to I$). We must have $Y_m \neq 0$ since otherwise $A_m = e^{X_m} \Rightarrow X_m = \log A_m \in \mathfrak{g}$, which is a contradiction.

Let $B_m = e^{-X_m} A_m = e^{Y_m}$ so that $B_m \in G$ for all m and $B_m \to I$ as $m \to \infty$. Since the unit sphere in $\mathfrak{g}^\perp$ is compact, there must exist a subsequence of Y_m's (call it Y_m again) such that $Y_m / \|Y_m\| \to Y$ with $\|Y\| = 1$. But then by the claim we just proved, we have $Y \in \mathfrak{g}$: since $Y_m/\|Y_m\| \in \mathfrak{g}^\perp$ for all m we have $Y = \lim Y_m/\|Y_m\| \in \mathfrak{g}^\perp$. But this implies that $Y = 0$, and this is a contradiction.

($\Leftarrow$) This is obvious since $e^{\log A} = A$ for all $A \in V_{\varepsilon_0}$. □

Corollary 6.11. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then there exists a neighbourhood U of 0 in $\mathfrak{g}$ and a neighbourhood V of I in G such that $\exp : U \to V$ is a homeomorphism.*



Corollary 6.12. *Let G be* connected. *Then every $A \in G$ can be written in the form $A = e^{X_1} e^{X_2} \dots e^{X_k}$ for $X_1, \dots, X_k \in \mathfrak{g}$, with k depending on A.*

Remark 6.2. Informally speaking, every *connected* Lie group is generated by a neighbourhood of I. Note that we cannot take $k = 1$ in general.

Proof. Let $A \in G$. Since G is connected, there exists a continuous path $\alpha : [0,1] \to G$ with $\alpha(0) = I$ and $\alpha(1) = A$. Let VV be a neighbourhood of I in G on which $\exp : \exp^{-1}(V) \to V$ is a homeomorphism. Since α is continuous, $f(x) = (\alpha(s))^{-1}\alpha(t)$ is continuous in s. Note that $f(t) = I$.. Also, for $f : (t-\varepsilon, t+\varepsilon) \to G$, there exists $\delta_t > 0$ such that $s \in (t-\delta_k, t+\delta_k)$ means $F(s) = (\alpha(s))^{-1}\alpha(t) \in V$. Since $[0,1]$ is compact, we can find $0 = t_0 < t_1 < t_2 < \cdots < t_{k-1} < t_k = 1$ such that $\alpha(t_{j-1})^{-1}\alpha(t_j) \in V$ for all $j = 1, 2, \ldots, k$. Note that

$$\begin{aligned} A &= \underbrace{A(t_0)^{-1}A(t_1)}_{\in V}\underbrace{A(t_1)^{-1}A(t_2)}_{\in V}\cdots\underbrace{A(t_{k-1})^{-1}A(t_k)}_{\in V}, \\ &= e^{X_1}e^{X_2}\cdots e^{X_k}, \end{aligned}$$

where $A(t_0)^{-1} = I, A(t_k) = A(1) = A$ and $X_1, \ldots, X_k$ are some matrices in $\mathfrak{g}$. □

Corollary 6.13. *Let G, H be matrix Lie groups where G is connected. Let $\Phi_j : G \to H$ be matrix Lie group homomorphisms $j = 1, 2$. If $(\Phi_1)_* = (\Phi_2)_* : \mathfrak{g} \to \mathfrak{h}$ then $\Phi_1 = \Phi_2$.*

Remark 6.3. This corollary is *false* if G is not connected.

Proof. Let $A \in G$, and write $A = e^{X_1}e^{X_2}\cdots e^{X_k}$, where $X_i \in \mathfrak{g}$.

$$\begin{aligned} \Phi_1(A) &= \Phi_1(e^{X_1}e^{X_2}\cdots e^{X_k}) \\ &= \Phi_1(e^{X_1})\Phi_1(e^{X_2})\cdots\Phi_1(e^{X_k}) \\ &= e^{(\Phi_1)_*X_1}\cdots e^{(\Phi_1)_*X_k} \\ &= e^{(\Phi_2)_*X_1}\cdots e^{(\Phi_2)_*X_k} \\ &= \Phi_2(e^{X_1})\Phi_2(e^{X_2})\cdots\Phi_2(e^{X_k}) \\ &= \Phi_2(e^{X_1}e^{X_2}\cdots e^{X_k}) = \Phi_2(A). \end{aligned}$$

□

6.3. *Abstract Lie algebras*

Definition 6.14. Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. A *finite-dimensional* $\mathbb{F}$*-Lie algebra* $\mathfrak{g}$ is a finite-dimensional $\mathbb{F}$-vector space together with a map $[\_,\_] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that:

(1) $[\_,\_]$ is bilinear over $\mathbb{F}$
(2) $[X, Y] = -[Y, X]$ for all $X, Y \in \mathfrak{g}$ (i.e. skew-symmetric)
(3) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X] + Y] = 0$ for all $X, Y, Z \in \mathfrak{g}$ (Jacobi identity).

Remark 6.4. As defined, a Lie algebra need not be a subspace of $\mathfrak{gl}(n, \mathbb{F})$, and $[X, Y]$ need not be $XY - YX$ (since the matrix multiplication may not make sense). However it is indeed the case that $[X, Y] = XY - YX$ in $\mathfrak{gl}(n, \mathbb{C})$ satisfies (3). Therefore $\mathfrak{gl}(n, \mathbb{F})$ is a $\mathbb{F}$-Lie algebra with commutator as Lie bracket. Therefore, the Lie algebra $\mathfrak{g}$ of a matrix Lie group G is a $\mathbb{R}$-Lie algebra in this abstract sense.

Also, if $X \in \mathfrak{g}$ then $iX \in \mathfrak{g}$ (complex matrix Lie group) then $\mathfrak{g}$ is a $\mathbb{C}$-Lie algebra in this sense.

Definition 6.15. Let $\mathfrak{g}$ be a real (resp. complex) Lie algebra.

(1) A *subalgebra of* $\mathfrak{g}$ is a real (resp. complex) subspace $\mathfrak{h}$ that is closed under the Lie bracket, i.e., $\underbrace{[H_1, H_2]}_{\text{Lie bracket in } \mathfrak{g}} \in \mathfrak{h}$ for all $H_1, H_2 \in \mathfrak{h}$.

    If $\mathfrak{g}$ is a complex Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ is a *real* subspace closed under the bracket, then $\mathfrak{h}$ is a *real subalgebra of* $\mathfrak{g}$.
(2) A $\mathbb{F}$-linear map $\lambda : \mathfrak{g} \to \mathfrak{h}$ between two $\mathbb{F}$-lie algebra is a *Lie algebra homomorphism* if $\lambda[X, Y]_\mathfrak{g} = [\lambda X, \lambda Y]_\mathfrak{h}$ for all $X, Y \in \mathfrak{g}$.

Example 6.16. $\mathfrak{u}(n)$ is a real subalgebra of the complex Lie algebra $\mathfrak{gl}(n, \mathbb{C})$.

Definition 6.17. It is easy to see that a bijective Lie algebra homomorphism has inverse which is also a Lie algebra homomorphism. Therefore, a bijective Lie algebra homomorphism is a *Lie algebra isomorphism*.

Theorem 6.18 (Ado's theorem)**.** *Every finite-dimensional* $\mathbb{F}$*-Lie algebra is isomorphic (as an abstract* $\mathbb{F}$*-Lie algebra) to a subalgebra of* $\mathfrak{gl}(n, \mathbb{F})$ *for some* $\mathbb{F}$ *with the Lie bracket being the usual commutator.*

Remark 6.5. It is not true if the dimension is infinite, even for $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$. The proof of this theorem is beyond the scope of this course.

7. January 26

Definition 7.1. Let $\mathfrak{g}$ be a Lie algebra, and let $X \in \mathfrak{g}$. We defined $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ by $\operatorname{ad}_X(Y) = [X, Y]$. Last week, from $\operatorname{Ad} : G \to \operatorname{GL}(\mathfrak{g})$, we induced a Lie algebra homomorphism $\operatorname{ad} := (\operatorname{Ad})_* : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$.

Claim. $\operatorname{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathfrak{g}$.

Proof. ad is linear since $[\_, \_]$ is bilinear. We need to check that $\operatorname{ad}_{[X,Y]} = [\operatorname{ad}_X, \operatorname{ad}_Y] \overset{?}{=} \operatorname{ad}_X \operatorname{ad}_Y - \operatorname{ad}_Y \operatorname{ad}_X$. Note that $\operatorname{ad}_{[X,Y]} Z = [[X, Y], Z]$ and $(\operatorname{ad}_X \operatorname{ad}_Y - \operatorname{ad}_Y \operatorname{ad}_X)Z = [X, [Y, Z]] - [Y, [X, Z]]$. By the Jacobi identity, it follows $[[X, Y], Z] = [X, [Y, Z]] - [Y, [X, Z]]$. □

Proposition 7.2. *For all* $X \in \mathfrak{g}$*, the map* $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ *is a* derivation *of* $\mathfrak{g}$*. That is,* $\operatorname{ad}_X[Y, Z] = [\operatorname{ad}_X Y, Z] + [Y, \operatorname{ad}_X Z]$.

Proof. $[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$ by the Jacobi identity. □

7.1. *Complexification of a real Lie algebra*

Definition 7.3. Let V be a finite-dimensional *real* vector space. Then the *complexification* $V_{\mathbb{C}}$ is the finite-dimensional $\mathbb{C}$-vector space such that

$$V_{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C} = \{v_i + iv_2; v_1, v_2 \in V\},$$

where $i(v + iw) := iv - w$.

Remark 7.1. $V_{\mathbb{C}}$ is a $\mathbb{C}$-vector space with $\dim_{\mathbb{C}} V_{\mathbb{C}} = \dim_{\mathbb{R}} V$. For instance, note that $(\mathbb{R}^n)_{\mathbb{C}} = \mathbb{C}^n$.

Remark 7.2. If W is a complex vector space, then by restrictions on scalars, we can consider the underlying real vector space $W_{\mathbb{R}}$. Note that we have the identity $\dim_{\mathbb{R}}(W_{\mathbb{R}}) = 2 \dim_{\mathbb{C}} W$ and $(\mathbb{C}^n)_{\mathbb{R}} = \mathbb{R}^{2n}$. As *real vector spaces*, we have the isomorphism $(V_{\mathbb{C}})_{\mathbb{R}} \cong V \oplus V$.

Remark 7.3. If $\{v_1, \ldots, v_n\}$ is a basis for V as a $\mathbb{R}$-vector space and $\{v_1, \ldots, v_n\}$ a basis for $V_{\mathbb{C}}$ as a $\mathbb{C}$-vector space, then $\{v_1, \ldots, v_n, iv_1, \ldots, iv_n\}$ isa basis for $(V_{\mathbb{C}})_{\mathbb{R}} \cong V \oplus V$ as a real vector space.

Proposition 7.4. *Let* $\mathfrak{g}$ *be a real Lie algebra, and let* $\mathfrak{g}_{\mathbb{C}}$ *be its complexification as a vector space. Then the Lie bracket* $[\_, \_]$ *on* $\mathfrak{g}$ *has a* unique *extension to* $\mathfrak{g}_{\mathbb{C}}$ *which makes* $\mathfrak{g}_{\mathbb{C}}$ *into a complex Lie algebra.*

Proof. Let $X_1, X_2, Y_1, Y_2 \in \mathfrak{g}$. We have, by bilinearity, that $[X_1 + iX_2, Y_1 + iY_2] = [X_1, Y_1] - [X_2, Y_2] + i[X_2, Y_1] + i[X_1, Y_2]$. One can show this makes $\mathfrak{g}_{\mathbb{C}}$ into a $\mathbb{C}$-Lie algebra (straightforward but tedious computations). $\square$

Example 7.5. $\mathfrak{gl}(n, \mathbb{R})_{\mathbb{C}} = \mathfrak{gl}(n, \mathbb{C})$; similarly, $\mathfrak{gl}(n, \mathbb{R})_{\mathbb{C}} = \mathfrak{sl}(n, \mathbb{C})$. The complexification of $\mathfrak{sp}(n, \mathbb{R})$ is $\mathfrak{sp}(n, \mathbb{R})_{\mathbb{C}} = \mathfrak{sp}(n, \mathbb{C})$.

Example 7.6 (Complexification of $\mathfrak{u}(n)$). Recall that $\mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) : X^* = -X\}$. Let $Z \in \mathfrak{gl}(n, \mathbb{C})$. Write

$$Z = \frac{Z - Z^*}{2} + i\frac{Z + Z^*}{2i} =: X + iY.$$

Note that both X and Y and skew-adjoint, since

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

and one can similarly see that $Y^* = -Y$. Hence $\mathfrak{u}(n)_{\mathbb{C}} = \mathfrak{gl}(n, \mathbb{C})$. One can show that both have dimension n^2.

Therefore, $\mathfrak{gl}(n, \mathbb{R})$ and $\mathfrak{u}(n)$ both have the same complexification. But they are *not isomorphic as real Lie algebras* Similar fact is true for $\mathfrak{su}(n)$ and $\mathfrak{sl}(n, \mathbb{R})$: they have the same complexification, but they are not isomorphic as real Lie algebras either.

8. January 26: Campbell-Baker-Hausdorff formula and its consequences

Recall that if $\Phi : G \to H$ is a homomorphism of matrix Lie groups, it induce a Lie algebra homomorphism $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ such that the following diagram commutes.

$$\begin{array}{ccc} G & \xrightarrow{\Phi} & H \\ \uparrow{\scriptstyle \exp} & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[\Phi_*]{} & \mathfrak{h} \end{array}$$

Question. Let $\mathfrak{g}, \mathfrak{h}$ be two Lie algebras of matrix Lie groups G and H. Let $\lambda : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. Then does there exist $\Phi : G \to H$ such that $\Phi_* = \lambda$?

So we want

$$\begin{array}{ccc} G & \xrightarrow{\Phi} & H \\ \uparrow{\scriptstyle \exp} & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[\lambda]{} & \mathfrak{h} \end{array}$$

to commute: that is, $\Phi(e^X) = e^{\lambda(X)}$. But here is the problem. Not every $A \in G$ is in the image of exp – recall that exp is *not always surjective*. Even though it is, the problem is far from over, since X need not be unique – recall that exp is *not always injective either*. We will soon see though that if G is *simply connected*, then problems can be overcome. The CBH formula is used to show that the map we can define is in fact a *homomorphism* from G to H. The CBH formula says that the group multiplication (at least near I) is *completely determined* by the Lie bracket in $\mathfrak{g}$.

The idea goes as follows. Let $\mathfrak{g}$ be the Lie algebra of G. Let also that $X, Y \in \mathfrak{g}$ sufficiently close to 0, thereby making $e^X, e^Y, e^X e^Y$ sufficiently close to I, so that $\log(e^X e^Y)$ is defined.

If $[X.Y] = 0$, then $e^X e^Y = e^{X+Y}$, so $\log(e^X e^Y) \overset{!}{=} X + Y$. But in general, the $\overset{!}{=}$ part is not necessarily true. The CBH formula expresses $\log(e^X e^Y)$ in terms of X and Y and the bracket of these two. We need the following analytic function $g : \{|z-1| < 1\} \to \mathbb{C}$ defined by

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

for some $a_m \in \mathbb{C}$. Note that since $g(z)$ can be written as the power series (as shown above) as long as $|z-1| < 1$, $g(z)$ is indeed analytic on $\{|z-1| < 1\}$.

Let V be a k-dimensional $\mathbb{C}$-vector space (if we choose a basis, $V \cong \mathbb{C}^k$ and we can take the usual Euclidean norm on $\mathrm{M}_k(\mathbb{C}) \cong L(V, V)$, the space of linear transformations from V to itself. Hence, for any $A \in L(V, V)$, with $\|A - I\| < 1$, we can define

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

This converges absolutely and uniformly on compact subsets to a continuous function of A.

Theorem 8.1 (Campbell-Baker-Hausdorff formula (integral form))**.** *Let $X, Y \in \mathfrak{gl}(n, \mathbb{C})$, with $\|X\|, \|Y\|$ sufficiently small. Then*

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

Before discussing the proof of this theorem, we will discuss some corollaries of CBH.

Corollary 8.2. *Let G be a matrix Lie group and $\mathfrak{g}$ its Lie algebra. Then, for $X, Y \in \mathfrak{g}$ with their norm sufficiently small, we have $\log(e^X e^Y) \in \mathfrak{g}$ and $\lambda(\log(e^X e^Y)) = \log(e^{\lambda(X)} e^{\lambda(Y)})$.*

Remark 8.1. The above corollary will be used to construct a homomorphism Φ of Lie groups from a homomorphism of Lie algebras and to relate Lie subgroups to Lie subalgebras.

Proof of Corollary. If $X, Y \in \mathfrak{g}$, then ad_X and ad_Y leave $\mathfrak{g}$ invariant. So the RHS of CBH is indeed in $\mathfrak{g}$. Hence the LHS of CBH is in $\mathfrak{g}$ as well. If λ is a Lie algebra homomorphism, then $\lambda[Y, X] = [\lambda X, \lambda Y]$, so $\lambda(\mathrm{ad}_Y X) = \mathrm{ad}_{\lambda(Y)} \lambda(X)$. So by induction, we have $\lambda((\mathrm{ad}_Y)^m X) = (\mathrm{ad}_{\lambda(Y)})^m \lambda(X)$, hence

$$\begin{aligned}
\lambda(e^{t\, \mathrm{ad}_Y}(X)) &= \sum_{m \geq 0} \frac{t^m}{m!} \lambda((\mathrm{ad}_Y)^m X) \ (\because\ \lambda \text{ continuous}) \\
&= \sum_{m \geq 0} \frac{t^m}{m!} (\mathrm{ad}_{\lambda(Y)})^m \lambda(X) = e^{t\, \mathrm{ad}_{\lambda(Y)}} \lambda(X).
\end{aligned}$$

We can repeat this computation to get $\lambda(e^{\mathrm{ad}_X} e^{t\, \mathrm{ad}_Y}) = e^{\mathrm{ad}_{\lambda(X)}} e^{t\, \mathrm{ad}_{\lambda(Y)}}$. We let X and Y small enough so that we can apply CBH, and also to $\lambda(X), \lambda(Y)$:

$$\begin{aligned}
\lambda(\log(e^X e^Y)) &= \lambda(X) + \int_0^1 \sum_{m \geq 0} a_m \lambda((e^{\mathrm{ad}\, X} e^{t\, \mathrm{ad}_Y} - I)^m (Y)) \, dt \ (\because \text{CBH and } \lambda \text{ linear}) \\
&= \lambda(X) + \sum_0^1 \sum_{m \geq 0} a_m (e^{\mathrm{ad}_{\lambda(X)}} e^{t\, \mathrm{ad}_{\lambda(Y)}} - I)^m (\lambda(Y)) \, dt \\
&= \log(e^{\lambda(X)} e^{\lambda(Y)}),
\end{aligned}$$

with the last equality following from the CBH for $\lambda(X)$ and $\lambda(Y)$. So $\log(e^X e^Y) \in \mathfrak{g}$ and $\lambda(\log(e^X e^Y)) = \log(e^{\lambda(X)} e^{\lambda(Y)})$. □

Remark 8.2 (Idea of proof of CBH). we ail show that two different expressions (one being $Z(X) = e^X e^{tY}$) satisfy the same ordinary differential equation (ODE) with the same initial conditions. By existence and uniqueness, we will get the equality. More in detail:

First step: we need to understand the linearization, or "differential", of the exponential map $\exp : \mathfrak{gl}(n, \mathbb{C}) \to \mathrm{GL}(n, \mathbb{C})$ at points other than $0 \in \mathfrak{g}$. We saw that

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

Recall that if $F : \mathbb{R}^k \to \mathbb{R}^k$ is smooth (or just differentiable), then the directional derivative of F at X in the direction of Y $(DF)_X : \mathbb{R}^k \to \mathbb{R}^k$ is the linear map defined as

$$(DF)_X(Y) = \left.\frac{d}{dt}\right|_{t=0} F(X + tY).$$

Since

$$(D\exp)_0(Y) = \left.\frac{d}{dt}\right|_{t=0} \exp(0 + tY) = Y,$$

it follows that $(D\exp)_0(Y) = I$, regardless of the dimension of the given space. So we want

$$\left.\frac{d}{dt}\right|_{t=0} \exp(X + tY) \text{ for } X \neq 0.$$

If X and Y commute, then $\exp(X + tY) = e^X e^{tY}$. Then $\left.\frac{d}{dt}\right|_{t=0} \exp(X + tY) = e^X Y$. Hence $(D\exp)_X = e^X$ if $[X, Y] = 0$. Next time, we will show that

$$\left.\frac{d}{dt}\right|_{t=0} e^{X+tY} = e^X \left(\frac{I - e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}\right) Y.$$

We will try to make sense of this.

9. January 28

We start by recalling the CBH:

Theorem 9.1 (Campbell-Baker-Hausdorff formula)**.** *For all $X, Y \in \mathfrak{gl}(n, \mathbb{C})$ sufficiently close to 0, we have*

$$\log(e^X e^Y) = X + \int_0^1 g(e^{\mathrm{ad}_X} e^{\mathrm{ad}_Y})\, dt,$$

where $g(z) = \frac{z \log z}{z-1} = \sum a_m (z-1)^m$.

We need some preliminary results before the full proof of CBH:

Theorem 9.2. *For any $X, Y \in \mathfrak{gl}(n, \mathbb{C})$ we have*

$$\left.\frac{d}{dt}\right|_{t=0} e^{X+tY} = e^X \left(\frac{I - e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}\right) Y. \tag{2}$$

Lemma 9.3. *For all $X \in \mathfrak{gl}(n, \mathbb{C})$,*

$$\frac{1 - e^{-\operatorname{ad}_X}}{\operatorname{ad}_X} = \lim_{m\to\infty} \frac{1}{m} \sum_{k=0}^{m-1} \left(\exp\left(-\frac{\operatorname{ad}_X}{m} \right) \right)^k .$$

Proof. Observe that $\exp(\frac{-\operatorname{ad}_X}{m})^k = \exp(-\frac{k}{m}\operatorname{ad}_X)$; hence (and also by absolute convergence (the reason $\stackrel{!}{=}$ is justified)),

$$\begin{aligned} \frac{1}{m} \sum_{k=0}^{m-1} \exp\left(-\frac{k}{m} \operatorname{ad}_X \right) &\stackrel{!}{=} \sum_{l=0}^{\infty} \frac{1}{m} \sum_{k=0}^{m-1} \frac{1}{l!} \left(-\frac{k}{m} \operatorname{ad}_X \right)^l \\ &= \sum_{l=0}^{\infty} \left[\frac{1}{m} \sum_{k=0}^{m-1} \left(\frac{k}{m} \right)^l \right] \frac{(-1)^l}{l!} (\operatorname{ad}_X)^l. \end{aligned}$$

Note, in fact, that $\frac{1}{m}\sum_{k=0}^{m-1} \left(\frac{k}{m}\right)^l$ is a Riemann sum for $x \mapsto x^l$ on $[0, 1]$ – i is a lower approximation to $\int_0^1 x^l = \frac{1}{l+1}$. Therefore

$$\sum_{l=0}^{\infty} \left[\frac{1}{m} \sum_{k=0}^{m-1} \left(\frac{k}{m} \right)^l \right] \frac{(-1)^l}{l!} (\operatorname{ad}_X)^l,$$

so the series converges absolutely. The terms (in l) are bounded in norm, since

$$\left(\frac{1}{l+1} \right) \frac{1}{l!} \| \operatorname{ad}_X \|^l \leq \frac{1}{l!} \| \operatorname{ad}_X \|^l,$$

hence

$$\sum_l \left(\frac{1}{l+1} \right) \frac{1}{l!} \| \operatorname{ad}_X \|^l \leq e^{\| \operatorname{ad}_X \|} < \infty.$$

Therefore we can apply the dominated convergence theorem to interchange limits $m, l \to \infty$. □

Proof of Theorem 9.2. Consider the complex function

$$\frac{1 - e^{-z}}{z} = \frac{1 - (1 - z + \frac{z^2}{2} - \frac{z^3}{6} + \cdots)}{z} = \sum_{m=0}^{\infty} \frac{(-1)^m z^m}{(m+1)!},$$

and this power series has infinite radius of convergence. Therefore it is an entire function, hence we can substitute into this function for any $A \in L(V, V)$ where V is a finite-dimensional vector space. We want $V = \mathfrak{gl}(n, \mathbb{C})$ and $A = \operatorname{ad}_X$.

More generally, if $X(t) \in \mathfrak{gl}(n, \mathbb{C})$ is smooth in t, then

$$\frac{d}{dt} e^{X(t)} = e^{X(t)} \left(\frac{I - e^{-\operatorname{ad}_{X(t)}}}{\operatorname{ad}_{X(t)}} \right) \frac{dX}{dt} \tag{3}$$

To see why (3) and (2) are equivalent, note that (2) follows from (3) by letting $X(t) = X + tY$ with $t = 0$; for the other direction, you can use the chain rule.

Write $F(X, Y) = \frac{d}{dt}\Big|_{t=0} e^{X+tY} = (D\exp)_X Y$. So $F(X, Y)$ is linear in Y. Since $\exp : \mathfrak{gl}(n, \mathbb{C}) \to \mathfrak{gl}(n, \mathbb{C})$ is a C^1-map (you will prove this in Assignment #3), we know that

$F(X,Y)$ is continuous in X and continuous in Y (and also linear in Y). We have

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

Take $\frac{d}{dt}\big|_{t=0}$ on both sides:

$$\begin{aligned}
F(X,Y) &= \sum_{k=0}^{m-1} \left[\exp\left(\frac{X}{m}\right)\right]^{m-k-1} \left[\frac{d}{dt}\bigg|_{t=0} \exp\left(\frac{X}{m} + t\frac{Y}{m}\right)\right] \left[\exp\left(\frac{X}{m}\right)\right]^k \\
&= \exp\left(\frac{m-1}{m}X\right)\left(\sum_{k=0}^{m-1}\left[\exp\left(\frac{X}{m}\right)\right]\right)^{-k} F\left(\frac{X}{m},\frac{Y}{m}\right)\left[\exp\left(\frac{X}{m}\right)\right]^k \\
&= \exp\left(\frac{m-1}{m}X\right)\frac{1}{m}\sum_{k=0}^{m-1}\mathrm{Ad}\left(\exp\left(-\frac{k}{m}x\right)\right)F\left(\frac{X}{m},Y\right) \\
&= \exp\left(\frac{m-1}{m}X\right)\frac{1}{m}\sum_{k=0}^{m-1}\exp\left(\mathrm{ad}\left(\frac{k}{m}X\right)\right)F\left(\frac{X}{m},Y\right).
\end{aligned}$$

So we have

$$\begin{aligned}
F(X,Y) &= \exp\left(\frac{m-1}{m}X\right)\frac{1}{m}\sum_{k=0}^{m-1}\left[\exp\left(-\frac{\mathrm{ad}_X}{m}\right)\right]^k F\left(\frac{X}{m},Y\right) \quad \text{(for all } m \geq 0) \\
&= \lim_{m\to\infty}\left(\underbrace{e^{\frac{m-1}{m}X}}_{(*)}\underbrace{\left(\frac{1}{m}\sum_{k=0}^{m-1}\left[\exp\left(-\frac{\mathrm{ad}_X}{m}\right)\right]^k\right)}_{(**)}\underbrace{F\left(\frac{X}{m},Y\right)}_{(***)}\right).
\end{aligned}$$

Note that $(*) \to e^X$ and e^X is continuous; by Lemma 9.3, $(**) = \frac{I-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}$. Note that $(***)$ is $F(0,Y)$ since F is continuous in Y, and $F(0,tY) = \frac{d}{dt}\big|_{t=0} e^{tY} = Y$. The claim follows. $\square$

Proof of CBH. Let $Z(t) := \log(e^X e^{tY})$. If X,Y both sufficiently close to 0, then $Z(t)$ is defined for all $t \in [0,1]$. In fact, $Z(t)$ is *smooth* in t. We want to compute $Z(1) = \log(e^X e^Y)$. Our strategy is to take advantage of the fundamental theorem of calculus, i.e.,

$$Z(1) = Z(0) + \int_0^1 \frac{dZ}{dt}\,dt. \tag{4}$$

Therefore, the proof of CBH reduces to proving that $\frac{dZ}{dt} = g(e^{\mathrm{ad}_X}e^{t\,\mathrm{ad}_Y})Y$.

By Theorem 9.2, we have

$$e^{-Z(t)}\frac{d}{dt}e^{Z(t)} = \left(\frac{I-e^{-\mathrm{ad}_{Z(t)}}}{\mathrm{ad}_{Z(t)}}\right)\frac{dZ}{dt},$$

provided that X and Y are small. $Z(t)$ is small, so $\frac{I-e^{-\mathrm{ad}_{Z(t)}}}{\mathrm{ad}_{Z(t)}}$ is invertible. Recall that

$$\frac{I-e^{-\mathrm{ad}_{Z(t)}}}{\mathrm{ad}_{Z(t)}} = \sum_{l=0}^{\infty}\frac{(-\mathrm{ad}_{Z(t)})^l}{(l+1)!},$$

so $\frac{I-e^{-\operatorname{ad}_{Z(t)}}}{\operatorname{ad}_{Z(t)}}$ can be written as a sum of I and "small" matrices. Thus

$$\frac{dZ}{dt} = \left(\frac{I - e^{-\operatorname{ad}_{Z(t)}}}{\operatorname{ad}_{Z(t)}}\right)^{-1} Y.$$

$\operatorname{Ad} : G \to \operatorname{GL}(\mathfrak{g})$ is a homomorphism, so

$$e^{\operatorname{ad}_X} e^{t\operatorname{ad}_Y} = e^{\operatorname{ad}_{Z(t)}} = \operatorname{Ad}_{e^{Z(t)}} = \operatorname{Ad}_{e^X} \circ \operatorname{Ad}_{e^{tY}}$$

take log on both sides to get $\operatorname{ad}_{Z(t)} = \log(e^{\operatorname{ad}_X} e^{t\operatorname{ad}_Y})$. So

$$\frac{dZ}{dt} = \left(\frac{I - (e^{\operatorname{ad}_X} e^{t\operatorname{ad}_Y})^{-1}}{\log(e^{\operatorname{ad}_X} e^{t\operatorname{ad}_Y})}\right)^{-1} Y = g(e^{\operatorname{ad}_X} e^{t\operatorname{ad}_Y})Y, \tag{5}$$

where $g(z) = \left(\frac{1-z^{-1}}{\log z}\right)^{-1}$. The formula now follows upon letting $t = 1$ and replace the $\frac{dZ}{dt}$ in (4) with the RHS of (5). □

Recall from last time, that we proved a corollary of CBH:

Corollary 9.4. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. If $\lambda : \mathfrak{g} \to \mathfrak{gl}(n, \mathbb{C})$ is a Lie algebra homomorphism, then for X, Y sufficiently small we have* $\log(e^X e^Y) \in \mathfrak{g}$ *and* $\lambda(\log(e^X e^Y)) = \log(e^{\lambda(X)} e^{\lambda(Y)})$.

So what does this corollary mean? The corollary says that near 0, the exponential map $\exp : U \in \mathfrak{g} \to \exp(U) = V \in G$ is a homomorphism. Near $I \in G$, we can write elements of G as e^X for some unique $X \in \mathfrak{g}$. The corollary says that if we define $\Phi : G \to \operatorname{GL}(n, \mathbb{C})$ by $\Phi(e^X) = e^{\lambda(X)}$ (defined on a neighbourhood V of I) then Φ is a "local" homomorphism. Write $\lambda(\log(e^X e^Y)) = \log(e^{\lambda(X)} e^{\lambda(Y)}) = \log(\Phi(e^X)\Phi(e^Y))$. Take the exponential map on both sides:

$$e^{\lambda(\log(e^X e^Y))} = \Phi(e^{\log(e^X e^Y)}) = \Phi(e^X e^Y)$$
$$e^{\lambda(\log(e^X e^Y))} = \Phi(e^X)\Phi(e^Y).$$

So Φ is a homomorphism whenever it is well-defined. That is, if $AB \in V$ for every $A, B \in V$, then we have $\Phi(AB) = \Phi(A)\Phi(B)$.

Theorem 9.5 (Campbell-Baker-Hausdorff formula (series form, up to the third order))**.**

$$\log(e^X e^Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + (\dagger\dagger),$$

with $(\dagger\dagger)$ being the higher order terms (i.e., bracket of X, Y of brackets of brackets $(\dots)$).

Proof. Let $g(z) := \frac{z \log z}{z-1}$. Write

$$\begin{aligned} g(z) &= \frac{(1+(z-1))[(z-1) - \frac{(z-1)^2}{2} + \frac{(z-1)^3}{3} - \cdots]}{z-1} \\ &= (1+(z-1))\left(1 - \frac{z-1}{2} + \frac{(z-1)^2}{3} - \cdots\right) \\ &= 1 + \frac{1}{2}(z-1) - \frac{1}{6}(z-1)^2 + \cdots = 1 + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m(m+1)}(z-1)^m. \end{aligned}$$

So

$$e^{\mathrm{ad}_X} e^{t\,\mathrm{ad}_Y} - I = \left(I + \mathrm{ad}_X + \frac{(\mathrm{ad}_X)^2}{2} + \cdots\right)\left(I + t\,\mathrm{ad}_Y + \frac{t^2}{2}(\mathrm{ad}_Y)^2 + \cdots\right) - I$$
$$= \mathrm{ad}_X + t\,\mathrm{ad}_Y + \frac{(\mathrm{ad}_X)^2}{2} + \frac{t^2}{2}(\mathrm{ad}_Y)^2 + t\,\mathrm{ad}_X\,\mathrm{ad}_Y + (\dagger),$$

where $(\dagger)$ denotes the higher-order terms in $\mathrm{ad}_X, \mathrm{ad}_Y$. Note that there is no zeroth order term, so $(e^{\mathrm{ad}_X} e^{t\,\mathrm{ad}_Y} - I)^m$ only has terms of degree $\geq m$ in $\mathrm{ad}_X, \mathrm{ad}_Y$. So up to degree 2 in $\mathrm{ad}_X, \mathrm{ad}_Y$, we have

$$g(e^{\mathrm{ad}_X} e^{t\,\mathrm{ad}_Y}) = I + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m(m+1)}(e^{\mathrm{ad}_X} e^{t\,\mathrm{ad}_Y} - I)^m$$
$$= I + \underbrace{\frac{1}{2}\left(\mathrm{ad}_X + t\,\mathrm{ad}_Y + \frac{(\mathrm{ad}_X)^2}{2} + \frac{t^2}{2}(\mathrm{ad}_Y)^2 + t\,\mathrm{ad}_X\,\mathrm{ad}_Y + \cdots\right)}_{m=1}$$
$$\underbrace{- \frac{1}{6}\left((\mathrm{ad}_X)^2 + t^2(\mathrm{ad}_Y)^2 + t\,\mathrm{ad}_X\,\mathrm{ad}_Y + t\,\mathrm{ad}_Y\,\mathrm{ad}_X + \cdots\right)}_{m=2} + \cdots .$$

Now apply to Y, with using the fact that $\mathrm{ad}_Y(Y) = [Y, Y] = 0$. So

$$g(e^{\mathrm{ad}_X} e^{t\,\mathrm{ad}_Y})Y = Y + \frac{1}{2}[X, Y] + \frac{1}{4}[X, [X, Y]] - \frac{1}{6}[X, [X, Y]] + \frac{t}{6}[Y, [X, Y]] + \cdots .$$

Thus

$$\log(e^X e^Y) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + (\dagger\dagger),$$

with $(\dagger\dagger)$ being the higher order terms (i.e., bracket of X, Y of brackets of brackets $(\dots)$). □

Remark 9.1. If $[X, Y] = 0$, then we already know that $\log(e^X e^Y) = X + Y$. Hence, the non-triviality of the Lie bracket is the infinitesimal measure of the non-commutativity of the Lie group. That is, if G is abelian then $e^X e^Y = e^Y e^X$ for all $X, Y \in \mathfrak{g}$ (Hint: replace X with tX and Y with sY: take the derivative $\frac{\partial f}{\partial t \partial s}$ at $(s, t) = (0, 0)$ and see what happens.) and $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$ (clear!).

Conversely, if $[X, Y] = 0$ for all X, Y then G is abelian near I.

Remark 9.2. The bracket is "infinitesimal" measure of (non-)commutativity. One can show that the Jacobi identity is a consequence of the fact that the Lie group G is *associative*. One can define a "weaker" notion of "Lie groups", i.e., a group with associativity dropped (but still want multiplication and inversion to be continuous). For such objects, some of this theory holds, but for instance the Jacobi identity *fails*.

Theorem 10.1. *Let G, H be matrix Lie groups with lie algebras $\mathfrak{g}, \mathfrak{h}$ respectively. Let $\lambda : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. If G is* simply connected, *then there exists a* unique *Lie group homomorphism $\Phi : G \to H$ such that $\Phi_* = \lambda$, i.e.,*

$$\begin{array}{ccc} G & \xrightarrow{\Phi} & H \\ {\scriptstyle \exp}\uparrow & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[\lambda]{} & \mathfrak{h} \end{array}$$

commutes.

Before proving this, we will first prove a corollary

Corollary 10.2. *Suppose G and H are both simply connected. If $\mathfrak{g} \cong \mathfrak{h}$ then $G \cong H$ also.*

Proof. Tthere exists $\lambda : \mathfrak{g} \cong \mathfrak{h}$ so $\lambda^{-1} : \mathfrak{h} \cong \mathfrak{g}$ is also an isomorphism. By Theorem 10.1, there exist Lie group homomorphisms $\Phi : G \to H$ and $\Psi : H \to G$ such that $\Phi_* = \lambda, \Psi_* = \lambda^{-1}$. Hence $(\Psi \circ \Phi)_* = \Psi_* \circ \Phi_* = \lambda^{-1}\lambda = \mathrm{id}_{\mathfrak{g}}$. So by Assignment #3, we have $\Psi \circ \Phi = \mathrm{id}$ and $\Phi \circ \Psi = \mathrm{id}$. □

Proof of Theorem 10.1. The proof will be broken into multiple steps:

Step 1: We have shown that there exists U open neighbourhood of $0 \in \mathfrak{g}$ and V open neighbourhood of $I \in G$ such that $\exp : U \to V = \exp(U)$ is a homeomorphism. Without loss of generality, take V small enough so that if $A, B \in V$ then $\log A, \log B$ are small enough to apply the CBH. On V, define Φ by $\Phi : V \to H$ where $\Phi(A) = e^{\lambda(\log A)}$.

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & H \\ {\scriptstyle \exp}\uparrow\downarrow{\scriptstyle \log} & & \uparrow{\scriptstyle \exp} \\ U & \xrightarrow[\lambda]{} & \mathfrak{h} \end{array}$$

Step 2: Since G is simply connected, it is connected. Let $A \in G$ be arbitrary. Then there must exist a continuous path $\alpha : [0, 1] \to G$ with $\alpha(0) = I$ and $\alpha(1) = A$. Just as we did last week, there must exist a partition $0 = t_0 < t_1 < \cdots < t_m = 1$ such that for all s and t satisfying $t_i \leq s \leq t \leq t_{i+1}$ we have $\alpha(t)\alpha(s)^{-1} \in V$.

Write

$$A = \alpha(1) = \underbrace{(\alpha(t_m)\alpha(t_{m-1})^{-1})}_{\in V}\underbrace{(\alpha(t_{m-1})\alpha(t_{m-2})^{-1})}_{\in V}\cdots\underbrace{(\alpha(t_2)\alpha(t_1)^{-1})}_{\in V}\alpha(t_1).$$

We want Φ to be a homomorphism so define

$$\Phi(A) = \Phi(\alpha(t_m)\alpha(t_{m-1})^{-1})\Phi(\alpha(t_{m-1})\alpha(t_{m-2})^{-1})\cdots\Phi(\alpha(t_2)\alpha(t_1)^{-1})\Phi(\alpha(t_1)),$$

using Φ defined on V by Step 1.

Step 3: We need to show that the definition of $\Phi(A)$ from Step 2 is independent of the partition.

We will show that $\Phi(A)$ is unchanged if we refine our partition. Let $s \in (t_i, t_{i+1})$ be an extra partition point. Then

$$\alpha(t_{i+1})\alpha(t_i)^{-1} = (\alpha(t_{i+1})\alpha(s)^{-1})(\alpha(s)\alpha(t_i)^{-1}).$$

So by a corollary of CBH (did it last time) on V, Φ as defined in Step 1 is a local homomorphism. Hence

$$\Phi(\alpha(t_{i+1})\alpha(t_i)^{-1}) = \Phi(\alpha(t_{i+1})\alpha(s)^{-1})\Phi(\alpha(s)\alpha(t_i)^{-1}).$$

So the definition of $\Phi(A)$ from step 2 is unchanged if we refine the partition. Therefore the definition of $\Phi(A)$ is independent of the choice of admissible partition (by considering their common refinement). Thus $\Phi(A)$ depends only on the path of α chosen. In the next step, we will show that in fact $\Phi(A)$ is independent of the path chosen.

Step 4: We now show that the definition of $\Phi(A)$ is independent of path (uses simple connectedness of G).

We said that G is simply connected if and only if every closed loop is continuously deformable to a constant loop. This notion is equivalent to the following notion (courtesy of algebraic topology): any two paths $\alpha, \widetilde{\alpha}$ joining the same two endpoints are homotopic as paths with fixed end points. This means that if $\alpha, \widetilde{\alpha} : [0,1] \to G$ such that $\alpha(0) = \widetilde{\alpha}(0)$ and $\alpha(1) = \widetilde{\alpha}(1)$ then there exists a continuous map $\beta : [0,1] \times [0,1] \to G$ such that

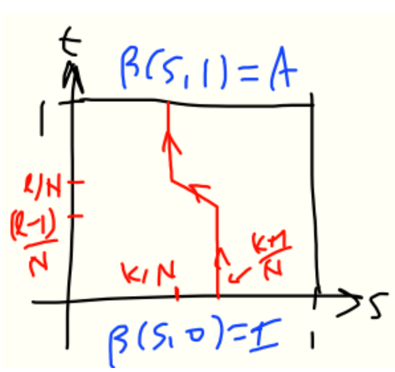
- $\beta(0,t) = \alpha(t)$ for all t
- $\beta(1,t) = \widetilde{\alpha}(t)$ for all t
- $\beta(s,0) = \alpha(0) = \widetilde{\alpha}(0)$ for all s
- $\beta(s,1) = \alpha(1) = \widetilde{\alpha}(1)$ fora ll s.

Let $\alpha, \widetilde{\alpha}$ be two paths from I to A. Need to show that the definitions of $\Phi(A)$ using each of these paths agrees. Let β be the one defined as above. Since β is continuous and $[0,1]\times[0,1]$ is compact, there exists $N > 0$ such that if $(s,t), (s',t') \in [0,1] \times [0,1]$ with $|s - s'| < 2N^{-1}$ and $|t - t'| < 2N^{-1}$ then $\beta(s,t)\beta(s',t')^{-1} \in V$. We will deform α to $\widetilde{\alpha}$ a little bit of time and show value of $\Phi(A)$ is unchanged with each step.

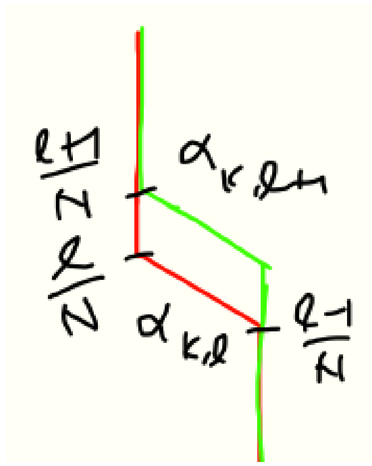
Define $\alpha_{k,l} : [0,1] \to G$ a continuous path from I to A for $k = 0, 1, \dots, N-1$ and $l = 0, 1, 2, \dots, N-1$ as follows:

$$\alpha_{k,l}(t) = \begin{cases} \beta\left(\frac{k+1}{N}, t\right) & t \in \left[0, \frac{l-1}{N}\right] \\ \beta\left(\frac{k}{N}, t\right) & t \in \left[\frac{l}{N}, 1\right]. \end{cases}$$

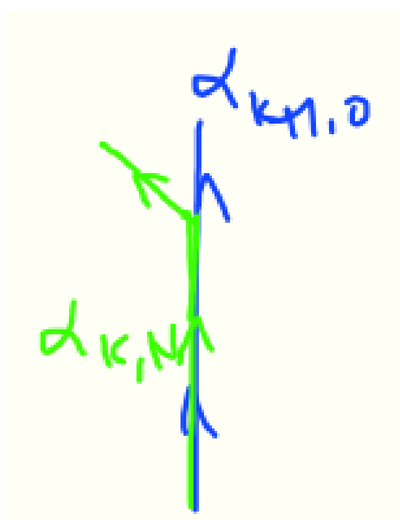
As for $[\frac{l-1}{N}, \frac{l}{N}]$, we interpolate between them with the straight line.



So more or less we deform the following way: $\alpha = \alpha_{0,0} \to \alpha_{0,1} \to \alpha_{0,2} \to \cdots \to \alpha_{0,N} \to \alpha_{1,0} \to \cdots \to \alpha_{N-1,N} \to \alpha_{N,0} = \widetilde{\alpha}$. Notice that $\alpha_{k,l}$ and $\alpha_{k,l+1}$ are identical except in the interval $[\frac{l-1}{N}, \frac{l}{N}]$.

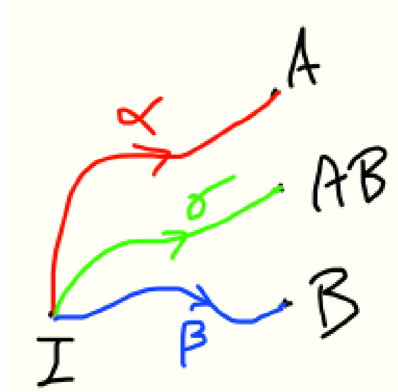


We have shown that the definition of $\Phi(A)$ using a given bath is independent of admissible partition. For both $\alpha_{k,l}$ and $\alpha_{k,l+1}$, choose the partition points to be $0, N^{-1}, 2N^{-1}, \ldots, (l-1)N^{-1}, (l+1)N^{-1}, \ldots, 1$ (note that we omitted l/N). So distance between any two consecutive partition points is $2N^{-1}$. So this is a valid partition since $\beta(s,t)\beta(s',t')^{-1} \in V$ if $s = s', |t - t'| < 2N^{-1}$. Note also that the values of $\Phi(A)$ obtained using paths $\alpha_{k,l}, \alpha_{k,l+1}$ using this partition for each path are *identical* because $\alpha_{k,l}(t_0) = \alpha_{k,l+1}(t_0)$ for any of these partition points t_0. So the definition of $\Phi(A)$ using $\alpha_{k,l}$ or $\alpha_{k,l+1}$ is the same. As for the case of $\alpha_{k,N} \to \alpha_{k+1,0}$, it's the same idea: note $\alpha_{k,N}(1) = \alpha_{k+1,0}(1) = A$.



Hence, we can define $\Phi(A)$ independent of path.

Step 5: We still need to show Φ is a homomorphism and that the diagram commutes. Let α be a path from I to A and β a path from I to B. We want to show that $\Phi(AB) \stackrel{?}{=} \Phi(A)\Phi(B)$.



For this, define

$$\gamma(t) = \begin{cases} \beta(2t) & 0 \le t \le \frac{1}{2} \\ \alpha(2t-1)B & \frac{1}{2} \le t \le 1. \end{cases}$$

Note that $\gamma(1/2) = \beta(1) = B$ and $\gamma(1/2) = \alpha(0)B = IB = B$, so γ is indeed a path from I to AB.

Let $0 = t_0 < t_1 < t_2 < \cdots < t_m = 1$ be an admissible partition for α. Also, let $0 = s_0 < s_1 < \cdots < s_n = 1$ be an admissible partition for β. Then

$$0 = \frac{t_0}{2} = \frac{s_0}{2} < \frac{s_1}{2} < \frac{s_2}{2} < \cdots < \frac{s_N}{2} = \frac{1}{2} < \frac{1}{2} + \frac{t_1}{2} < \frac{1}{2} + \frac{t_2}{2} < \cdots < \frac{1}{2} + \frac{t_m}{2} = 1$$

is an admissible partition for γ. Notice that

$$\begin{aligned} \gamma\left(\frac{1}{2} + \frac{t_i}{2}\right)\gamma\left(\frac{1}{2} + \frac{t_{i+1}}{2}\right)^{-1} &= (\alpha(t_i)B)(\alpha(t_{i-1})B)^{-1} \\ &= \alpha(t_i)BB^{-1}\alpha(t_{i-1})^{-1} = \alpha(t_i)\alpha(t_{i-1})^{-1} \in V, \end{aligned}$$

and similarly we have $\gamma(\frac{s_i}{2})\gamma(\frac{s_{i+1}}{2})^{-1} = \beta(s_i)\beta(s_{i-1})^{-1}$. So since

$$\begin{aligned} A &= \alpha(t_m)\alpha(t_{m-1})^{-1}\alpha(t_{m-1})\alpha(t_{m-2})^{-1}\cdots\alpha(t_2)\alpha(t_1)^{-1}\alpha(t_1) \\ B &= \beta(s_n)\beta(s_{n-1})^{-1}\beta(s_{n-1})\beta(s_{n-2})^{-1}\cdots\beta(s_2)\beta(s_1)^{-1}\beta(s_1), \end{aligned}$$

we have

$$AB = \gamma(r_{n+m})\gamma(r_{n+m-1})^{-1}\cdots\gamma(r_2)\gamma(r_1)^{-1}\gamma(r_1).$$

So $\Phi(AB) = \Phi(A)\Phi(B)$, as desired.

Step 6: Finally, near I, we have $\Phi = \exp \circ \lambda \circ \log$, so

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

for t sufficiently small. Hence

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

This proves that $\lambda = \Phi_*$. So the diagram

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & H \\ {\scriptstyle \exp}\uparrow\downarrow{\scriptstyle \log} & & \uparrow{\scriptstyle \exp} \\ U & \xrightarrow[\lambda]{} & \mathfrak{h} \end{array}$$

does commute, which is what we wanted. □

11. February 4

Let $G \in \mathrm{GL}(n, \mathbb{C})$ be matrix Lie group. Let $H \leq G$ be a matrix Lie subgroup (i.e., it is closed in G). Then the Lie algebra of H, $\mathfrak{h} := \{X \in \mathfrak{gl}(n, \mathbb{C}) : e^{tX} \in H \leq G \text{ for all } t \in \mathbb{R}\} \leq \mathfrak{g}$. One can verify that $\mathfrak{h}$ is closed under bracket, so $\mathfrak{h}$ is a Lie algebra of $\mathfrak{g}$.

What about converse? If G is a matrix Lie group with $\mathrm{Lie}(G) = \mathfrak{g}$ and $\mathfrak{h}$ a subalgebra of $\mathfrak{g}$, does there exist a matrix Lie subgroup H of G such that $\mathrm{Lie}(H) = \mathfrak{h}$? Answer is that, in general, *no* because the notion of matrix Lie subgroup is too restrictive. Consider the following example.

Example 11.1. Let $G = \mathrm{GL}(2, \mathbb{C})$ and $\mathfrak{g} = \mathfrak{gl}(2, \mathbb{C})$. Define

$$\mathfrak{h} := \left\{ \begin{bmatrix} it & 0 \\ 0 & it\alpha \end{bmatrix} : t \in \mathbb{R} \right\}$$

for some fixed irrational α. Then $\mathfrak{h}$ is clearly a Lie subalgebra. So does there exist a matrix Lie group $H \subset \mathrm{GL}(2, \mathbb{C})$ such that $\mathrm{Lie}(H) = \mathfrak{h}$? If so, then H contains

$$\exp\left(\begin{bmatrix} it & 0 \\ 0 & it\alpha \end{bmatrix} \right) = \begin{bmatrix} e^{it} & 0 \\ 0 & e^{it\alpha} \end{bmatrix}$$

for all $t \in \mathbb{R}$. Therefore

$$H' := \left\{ \begin{bmatrix} e^{it} & 0 \\ 0 & e^{it\alpha} \end{bmatrix} : t \in \mathbb{R} \right\} \subset H.$$

We want H to be closed, so it must contain the closure of H'. But the Lie algebra of H' is two-dimensional while $\mathfrak{h}$ is one-dimensional. Contradiction! Note that H' is isomorphic (as a *group*) to a matrix Lie group, but it is not isomorphic as a *matrix Lie subgroup* of $\mathrm{GL}(2, \mathbb{C})$.

While the answer to our previous question is no in general, one can nonetheless prove a weaker converse, as we shall see later.

Definition 11.2. Let H be any subgroup of $\mathrm{GL}(n, \mathbb{C})$, which is *not necessarily closed.* Then its *Lie algebra* $\mathfrak{h}$ is defined to be

$$\mathrm{Lie}(H) = \mathfrak{h} = \{X \in \mathfrak{gl}(n, \mathbb{C}) : e^{tX} \in H \text{ for all } t \in \mathbb{R}\}.$$

Definition 11.3. Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. A subgroup H of G (again, not necessarily closed) is called an *analytic subgroup* (or a *connected Lie subgroup*) if:

- $\mathfrak{h} := \mathrm{Lie}(H)$ is a vector subspace of $\mathfrak{g}$.
- Every $A \in H$ can be written in the form $A = \prod_{i=1}^{m} e^{X_i}$ for some m and $X_1, \ldots, X_m \in \mathfrak{h}$.

Proposition 11.4. *Let H be an analytic subgroup of G. Then H is path-connected.*

Proof. Let $A \in H$. Then there exist $X_1, \ldots, X_m \in \mathfrak{h}$ such that $A = e^{X_1}e^{X_2}\cdots e^{X_m}$. Let $\alpha(t) = Ae^{-tX_m} = e^{X_1}e^{X_2}\cdots e^{(1-t)X_m}$, which is a continuous path such that $\alpha(0) = A$ and $\alpha(1) = e^{X_1}\cdots e^{X_{m-1}}$ (and then iterate). $\square$

Proposition 11.5. *Let H be an analytic subgroup of G. Then $\mathfrak{h} = \mathrm{Lie}(H)$ is a subalgebra of $\mathfrak{g}$.*

Proof. We need to show that $[X, Y] \in \mathfrak{h}$ if $X, Y \in \mathfrak{h}$. Let $A \in H$ and $Y \in \mathfrak{h}$. Then $e^{tAYA^{-1}} = Ae^{tY}A^{-1}$, so $e^{tAYA^{-1}} \in H$ since $e^{tY} \in H$, for all t. Therefore $AYA^{-1} \in \mathfrak{h}$. And if $X, Y \in \mathfrak{h}$, then $e^{tX}Ye^{-tX} \in \mathfrak{h}$. Take $\frac{d}{dt}\big|_{t=0}$, and use the fact that $\mathfrak{h}$ is a subspace of $\mathfrak{g}$. The proof is complete upon noticing that

$$\frac{d}{dt}\bigg|_{t=0} e^{tX}Ye^{-tX} = [X, Y] \in \mathfrak{h}. \qquad \square$$

Theorem 11.6 (Main theorem for today's lecture)**.** *Let G be a matrix Lie group and let $\mathfrak{g} := \mathrm{Lie}(G)$. If $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then there exists a* unique *analytic subgroup H of G with $\mathrm{Lie}(H) = \mathfrak{h}$. In fact, $H = \{e^{X_1}e^{X_2}\cdots e^{X_m} : X_i \in \mathfrak{h}\}$.*

Before we prove the theorem, notice that without loss of generality, we can take $G = \mathrm{GL}(n, \mathbb{C})$ because an analytic subgroup of $\mathrm{GL}(n, \mathbb{C})$ such that $\mathrm{Lie}(H) = \mathfrak{h} \subseteq \mathfrak{g}$ is an analytic subgroup of G. We need two technical lemmas for the proof of Theorem 11.6.

Definition 11.7. Let $\mathcal{B}$ be a basis of $\mathfrak{h}$. An element $R \in \mathfrak{h}$ is called *rational* with respect to $\mathcal{B}$ if its coordinates with respect to this basis are rational.

Lemma 11.8. *For all $\delta > 0$ and for all $A \in H$, there exist rational elements $R_1, \ldots, R_m \in \mathfrak{h}$ such that $A = e^{R_1}e^{R_2}\cdots e^{R_m}e^X$ with $X \in \mathfrak{h}$ and $\|X\| < \delta$.*

Proof. Let $\varepsilon > 0$ be sufficiently small so that CBH (for $\mathrm{GL}(n, \mathbb{C})$) holds for all X, Y such that $\|X\|, \|Y\| < \varepsilon$:

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

$C(X, Y)$ is continuous at X and Y. Since $C(X, Y)$ is continuous, choose $\varepsilon' > 0$ (without loss of generality assume $\varepsilon' < \varepsilon$) such that $\|X\|, \|Y\| < \varepsilon$. Then $\|C(X, Y)\| < \varepsilon$. Since $e^X = (e^{\frac{X}{k}})^k$, any element of H can be written in the form $A = e^{X_1}e^{X_2}\cdots e^{X_m}$ with $\|X_i\| < \varepsilon'$ for all i. Since $\mathfrak{h}$ is a subalgebra of $\mathfrak{gl}(n, \mathbb{C})$ (assumption), by CBH we have $C(X_1, X_2) \in \mathfrak{h}$.

Choose a rational element $R_1 \in \mathfrak{h}$ close to $C(X_1, X_2)$ such that $\|R_1\| < \varepsilon$ (possible since $\|X_1\|, \|X_2\| < \varepsilon' \Rightarrow C(X_1, X_2) < \varepsilon$). Then

$$e^{X_1}e^{X_2} = e^{C(X_1,X_2)} = e^{R_1}e^{-R_1}e^{C(X_1,X_2)} = e^{R_1}e^{C(-R,C(X_1,X_2))}.$$

Write $\widetilde{X_2} = C(-R_1, C(X_1, X_2))$. Notice that $C(-X, X) = 0$ for all X by choosing R_1 sufficiently close to $C(X_1, X_2)$. Thus we can ensure that $\|\widetilde{X_2}\| < \varepsilon'$. So $A = e^{X_1}e^{X_2}\cdots e^{X_m} = e^{R_1}e^{\widetilde{X_2}}e^{X_3}\cdots e^{X_m}$, with $\|\widetilde{X_2}\|, \|X_j\| < \varepsilon'$ for all $j = 3, \ldots, m$. Iterate this process to get $A = e^{R_1}e^{R_2}\cdots e^{R_{m-1}e^{\widetilde{X_m}}}$ with $R_1, \ldots, R_{m-1}$ rational elements and $\|\widetilde{X_m}\| < \varepsilon'$. $\square$

Recall that $\mathfrak{gl}(n,\mathbb{C}) = \mathfrak{h} \oplus \mathfrak{h}^\perp$. We proved earlier that there exist a neighbourhood U of 0 in $\mathfrak{h}$, a neighbourhood W of 0 in $\mathfrak{h}^\perp$, and neighbourhood V of I in $\mathrm{GL}(n,\mathbb{C})$ such that every $A \in V$ can be written uniquely as $A = e^Y e^X$ for some $Y \in W \subseteq \mathfrak{h}^\perp$ and $X \in U \subseteq \mathfrak{h}$ where X and Y depend continuously on A.

Lemma 11.9. *The set of $\{Y \in W : e^Y \in H\}$ is at most countable.*

Proof. Fix $\delta > 0$ such that for all $X, Y \in \mathfrak{h}$ with $\|X\|, \|Y\| < \delta$, $C(X,Y)$ is defined and contained in U (possible since $C(0,0) = 0$ and C is continuous in X and Y). We claim that for each finite set $\{R_1, \ldots, R_m\}$ of rational element in $\mathfrak{h}$,, there is at most one $X \in \mathfrak{h}$ with $\|X\| < \delta$ such that $e^{R_1} e^{R_2} \cdots e^{R_m} e^X \in \exp(W)$.

Proof of the above claim. Suppose that there are more than one. If $e^{R_1} e^{R_2} \cdots e^{R_m} e^{X_1} = e^{Y_1}$ and $e^{R_1} e^{R_2} \cdots e^{R_m} e^{X_2} = e^{Y_2}$ with $X_i \in \mathfrak{h}, \|X_i\| < \delta$ and $Y_i \in W \in \mathfrak{h}^\perp$, then $e^{Y_2} e^{-X_2} = e^{Y_1} e^{-X_1}$. Hence $e^{Y_2} = e^{Y_1} e^{-X_1} e^{X_2} = e^{Y_1} e^{C(-X_1, X_2)}$ with $C(-X_1, X_2) \in U \subseteq \mathfrak{h}$. But each element of U has a unique representative $e^Y e^X$ with $Y \in W, X \in U$. Hence, by uniqueness $Y_2 = Y_1$, and $C(-X_1, X_2) = 0$, and $e^{X_1} = e^{X_2}$. So $X_1 = X_2$ since exp is injective in U. ☐

By Lemma 11.8, for every $A \in H$, there exist rational $R_1, \ldots, R_m \in \mathfrak{h}$ such that $A = e^{R_1} \cdots e^{R_m} e^X$ with $X \in \mathfrak{h}$ and $\|X\| < \delta$. But there exist countably many rational elements in $\mathfrak{h}$, so countably many $e^{R_1} \cdots e^{R_m}$, each of which (by the claim above) products at most one element $e^Y = e^{R_1} \cdots e^{R_m} e^X$ for some $Y \in W$. The lemma follows. ☐

Proof of Theorem 11.6. Recall that we defined $H = \{e^{X_1} e^{X_2} \cdots e^{X_m} : X_i \in \mathfrak{h}\}$. This is clearly a subgroup of $\mathrm{GL}(n,\mathbb{C})$. We want to show that H is an analytic subgroup of $\mathrm{GL}(n,\mathbb{C})$ and $\mathrm{Lie}(H) = \mathfrak{h}$, since uniqueness is clear from the properties of analytic subgroups. In order for H to be an analytic subgroup, we need to prove two things: (a) $\mathrm{Lie}(H)$ is a subgroup of $\mathfrak{gl}(n,\mathbb{C})$; and (b) every element is of the form $e^{X_1} e^{X_2} \cdots e^{X_m}$ with $X_i \in \mathrm{Lie}(H)$. If we show that $\mathrm{Lie}(H) = \mathfrak{h}$ then we are done, since $\mathfrak{h}$ is a subalgebra (hence a subspace) of $\mathfrak{gl}(n,\mathbb{C})$ by construction. Let $\widetilde{\mathfrak{h}} = \mathrm{Lie}(H)$. Need to show that $\widetilde{\mathfrak{h}} = \mathfrak{h}$. If $X \in \mathfrak{h}$, then $e^{tX} \in H$ for all $t \in \mathbb{R}$ by the definition of H, since $\mathfrak{h}$ is a subspace. So $X \in \widetilde{\mathfrak{h}} = \mathrm{Lie}(H)$. Hence $\mathfrak{h} \subseteq \widetilde{\mathfrak{h}}$. Let $z \in \widetilde{\mathfrak{h}} = \mathrm{Lie}(H)$, for all sufficiently small t. Write $e^{tZ} = e^{Y(t)} e^{X(t)}$ with $Y(t) \in W \subseteq \mathfrak{h}^\perp$ and $X(t) \in U \subseteq \mathfrak{h}$, with $X(t)$ and $Y(t)$ continuous in t. But $e^{tZ}, e^{X(t)} \in H$. Hence $e^{Y(t)} = e^{tZ} e^{-X(t)} \in H$, for all sufficiently small t. But by Lemma 11.9, we must have $Y(t)$ is constant. If not, it takes on uncountably many values (just apply the intermediate value theorem to one of its components). Hence $Y(t)$ must be constant. But $Y(0) = 0$ so $e^{tZ} = e^{X(t)}$ for all sufficiently small t. Hence $tZ = X(t) \in \mathfrak{h}$ for all sufficiently small t. Thus $z \in \mathfrak{h}$ so $\widetilde{\mathfrak{h}} \subseteq \mathfrak{h}$, as required. ☐

Remark 11.1 (Quick summary of what we did today). There exists a one-to-one correspondence between analytic subgroups of $\mathrm{GL}(n,\mathbb{C})$ and subalgebras of $\mathfrak{gl}(n,\mathbb{C})$. The one-to-one correspondence map is given by $H \mapsto \mathfrak{h} = \mathrm{Lie}(H)$ where H is an analytic subgroup of G. Conversely, if $\mathfrak{h}$ is a subalgebra of $\mathfrak{gl}(n,\mathbb{C})$ then this subalgebra corresponds to $H = \{e^{X_1} e^{X_2} \cdots e^{X_m} : X_i \in \mathfrak{h}\}$ (an analytic subgroup of $\mathrm{GL}(n,\mathbb{C})$).

12. February 9: Representation theory of Lie groups and Lie algebras

Definition 12.1. Let G be a matrix Lie group. A *finite-dimensional complex representation of G* is a matrix Lie group homomorphism

$$\Pi : G \to \mathrm{GL}(V),$$

for some finite-dimensional complex vector space V, i.e., $V \cong \mathbb{C}^n$ and $\mathrm{GL}(V) \cong \mathrm{GL}(n, \mathbb{C})$ for some $n = \dim(V) \geq 1$.

A *finite-dimensional real representation of G* is a matrix Lie group homomorphism $\Pi : G \to \mathrm{GL}(V)$ where V is a finite-dimensional real vector space V such that $\dim V \geq 1$.

Remark 12.1. If $g \in G$, then $\Pi(g) \in \mathrm{GL}(V)$. We shall abuse notation and write $g \cdot v := \Pi(g)(v)$ for all $v \in V$ when the representation Π of G is understood. We say that G *acts on* V by the representation Π. W sometime also say that V is a representation of G. Note that since $\Pi(gh) = \Pi(g)\Pi(h)$, we have $g \cdot (h \cdot v) = (gh) \cdot v$.

Definition 12.2. Let $\mathfrak{g}$ be a real or complex abstract Lie algebra. A *finite-dimensional complex representation of* $\mathfrak{g}$ is a Lie algebra homomorphism

$$\pi : \mathfrak{g} \to \mathfrak{gl}(V)$$

for some finite-dimensional complex vector space $\dim(V) \geq 1$. Recall that $\mathfrak{gl}(V)$ is a Lie algebra whose Lie bracket is the usual commutator.

If $\mathfrak{g}$ is a real algebra then a *real representation of* $\mathfrak{g}$ is a Lie algebra homomorphism $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ where V is a finite-dimensional real vector space.

Remark 12.2. Throughout this course, every representation we consider is *finite-dimensional* unless stated otherwise.

Definition 12.3. A representation (real or complex) of a Lie group or a Lie algebra is called *faithful* if the homomorphism is *injective*.

Definition 12.4. Let Π be a real or complex representation of G, acting on V. Then subspace W of V is called *G-invariant* (or *invariant under G*) if $\Pi(g)w \in W$ for all $w \in W$, i.e., $\Pi(g)W \subseteq W$.

If $W \neq \{0\}, V$ then the subspace W is called *non-trivial*. A representation is called *irreducible* if it has no non-trivial invariant subspace.

The same type of definitions are applicable for representations of Lie algebras.

Remark 12.3. Any one-dimensional representation is necessarily irreducible.

Definition 12.5. Let $\Pi : G \to \mathrm{GL}(V)$ and $\Sigma : G \to \mathrm{GL}(W)$ be two representations of a matrix Lie group G (both real or both complex). A linear map $T : V \to W$ is said to be a *morphism of representations*, or an *intertwining map* if $(T \circ \Pi(g))(v) = (\Sigma(g) \circ T)(v)$ for all $g \in G$ and $v \in V$. That is, the following diagram commutes:

$$\begin{array}{ccc} V & \xrightarrow{\Pi(g)} & V \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ W & \xrightarrow[\Sigma(g)]{} & W \end{array}$$

An intertwining map that is an isomorphism of vector spaces is called an *isomorphism of representations* or an *equivalence of representations*. We consider the isomorphic G-representations to be "the same".

The same definition is applicable for Lie algebra representations. (i.e., $(T \circ \pi(X))(v) = (\sigma(X) \circ T)(v)$ for all $X \in \mathfrak{g}$ and $v \in V$.

Question. Given G (or $\mathfrak{g}$), what are all the irreducible representations up to equivalence, real or complex? We will answer this question in the next few lectures.

Proposition 12.6. *Let* Π *be a representation of* G *acting on* V*. Then there exists a* unique *representation of* $\mathfrak{g} = \mathrm{Lie}(G)$ *acting on the same vector space* V *such that:*

- $\Pi(e^X) = e^{\pi(X)} = e^{\Pi_*(X)}$ *for all* $X \in \mathfrak{g}$
- $\pi(X) = \frac{d}{dt}\big|_{t=0} \Pi(e^{tX})$
- $\pi(gXg^{-1}) = \Pi(g)\pi(X)(\Pi(g))^{-1}$ *for all* $g \in G, X \in \mathfrak{g}$.

Proof. This is immediate from our earlier results, namely $\pi \equiv \Pi_* : \mathfrak{g} \to \mathfrak{gl}(V)$, where $\mathfrak{g} = \mathrm{Lie}(G)$ and $\mathfrak{gl}(V) = \mathrm{Lie}(\mathrm{GL}(V))$. □

Proposition 12.7. *Let* G *be a* connected *matrix Lie group with the Lie algebra* $\mathfrak{g}$.

(1) Let Π *be a representation of* G *and let* $\pi = (\Pi)_*$ *be the associated representation of* $\mathfrak{g}$*. Then* Π *is irreducible if and only if* π *is irreducible.*

(2) Let Π_1, Π_2 *be two representations of* G *with associated Lie algebra representations* π_1, π_2*, respectively. Then* $\Pi_1 \cong \Pi_2$ *if and only if* $\pi_1 \cong \pi_2$.

Proof. ((1), $\Rightarrow$) Suppose that Π is irreducible. Let W be a subspace of V invariant under $\pi(X)$ for all $X \in \mathfrak{g}$. Since G is connected, any $g \in G$ is of the form $G = e^{X_1}e^{X_2}\cdots e^{X_m}$ for some $X_1, \ldots, X_m \in \mathfrak{g}$. If W is invariant under $\pi(X_i)$, then W is invariant under $e^{\pi(X_i)} = \Pi(e^{X_i})$ (to see why this follows, recall that $W \leq V$ is topologically closed). Hence $\Pi(g) = \Pi(e^{X_1})\cdots\Pi(e^{X_m})$ leaves W invariant for all $g \in G$. Since Π is irreducible, $W = \{0\}$ or V. This proves that π is irreducible.

((1), $\Leftarrow$) Conversely, suppose π is an irreducible representation of $\mathfrak{g}$ acting on V. Let W be a subspace of V invariant under $\Pi(g)$ for all $g \in G$. Hence $\Pi(e^{tX})W \subseteq W$ for all $t \in \mathbb{R}$ and $X \in \mathfrak{g}$. Take

$$\frac{d}{dt}\bigg|_{t=0} \Pi(e^{tX})W \subseteq W$$

for all $X \in \mathfrak{g}$. So W is invariant under all $\pi(X)$. hence $W = \{0\}$ or V. So π is irreducible.

((2), $\Leftrightarrow$) You will prove this in Assignment #4. □

Proposition 12.8. *Let* $\mathfrak{g}$ *be a real Lie algebra and let* $\mathfrak{g}_{\mathbb{C}}$ *be its complexification. Every* complex representation π *of* $\mathfrak{g}$ *has a* unique *extension to a complex representation of* $\mathfrak{g}_{\mathbb{C}}$*, also denoted* π*, given by*

$$\pi(X + iY) = \pi(X) + i\pi(Y)$$

for all $X, Y \in \mathfrak{g}$*. Moreover,* π *is irreducible as a representation of* $\mathfrak{g}$ *if and only if* π *is irreducible as a representation of* $\mathfrak{g}_{\mathbb{C}}$.

Remark 12.4. This proposition does not make sense if the word "complex" is replaced with "real".

Proof. If such an extension is to exist, we must have, by complex linearity, that $\pi(X+iY) = \pi(X) + i\pi(Y)$ for all $X, Y \in \mathfrak{g}$.

Define it by the above, then we need to show that it is a complex Lie algebra homomorphism. One can show that (computational steps omitted) $\pi([X_1 + iY_1, X_2 + iY_2]) = [\pi(X_1 + iY_1), \pi(X_2 + iY_2)]$.

What is more interesting is the second part of the proposition.

($\Rightarrow$) Let $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ be a complex representation of $\mathfrak{g}$. Suppose that π is irreducible as a representation of $\mathfrak{g}$. Let $W \subseteq V$ be an irreducible subspace for $\mathfrak{g}_{\mathbb{C}}$, i.e., $\pi(X+iY)W \subseteq W$ for all $X, Y \in \mathfrak{g}$. If $Y = 0$, then we have $\pi(X)W \subseteq W$ for all $X \in \mathfrak{g}$. So W is invariant subspace for $\mathfrak{g}$, hence $W = \{0\}$ or V. Thus π is also irreducible as a representation of $\mathfrak{g}_{\mathbb{C}}$.

($\Leftarrow$) Conversely, suppose that $W \subseteq V$ is an invariant subspace for $\mathfrak{g}$. Then $\pi(X)W \subseteq W$ for all $X \in \mathfrak{g}$, and W is a complex subspace, hence $i\pi(Y)W \subseteq W$ for all $Y \in \mathfrak{g}$. Thus $\pi(X+iY)W \subseteq W$ for all $X, Y \in \mathfrak{g}$. Thus W is an invariant subspace for $\mathfrak{g}_{\mathbb{C}}$. But since $W = \{0\}$ or V, it follows that π is irreducible as a representation of $\mathfrak{g}$. $\square$

Example 12.9 (The standard representation). By definition, $G \subseteq \mathrm{GL}(n, \mathbb{C})$ for some n. The inclusion map $\iota : G \hookrightarrow \mathrm{GL}(n, \mathbb{C})$ is a complex finite-dimensional representation of G.

If $G \subseteq \mathrm{GL}(n, \mathbb{R})$ for some n, then the inclusion map $\iota : G \hookrightarrow \mathrm{GL}(n, \mathbb{R})$ is a real finite-dimensional representation of G.

Similarly, if $\mathfrak{g}$ is a Lie algebra of matrices (such as the Lie algebra of a matrix Lie group) (i.e., $\mathfrak{g}$ is a subalgebra of $\mathfrak{gl}(n, \mathbb{C})$ or $\mathfrak{gl}(n, \mathbb{R})$), then the inclusion map is a finite-dimensional (complex or real) representation of $\mathfrak{g}$.

Remark 12.5. If an abstract Lie group is not a matrix Lie group, then there is *no* "standard representation".

Example 12.10 (The trivial representation). Consider $\Pi : G \to \mathrm{GL}(1, \mathbb{C})$ given by $\Pi(g) = 1$ for all $g \in G$. This is irreducible since this representation is one-dimensional. Another non-trivial representation is $\pi : \mathfrak{g} \to \mathfrak{gl}(1, \mathbb{C})$ given by $\pi(X) = 0$ for all $X \in \mathfrak{g}$. This is irreducible, for the same reason (one-dimensional).

Example 12.11 (The Adjoint representation). We have already seen a non-trivial representation of a matrix Lie group and its Lie algebra, namely the big "Ad". (Let $V = \mathfrak{g}$.) Recall that $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$, and $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is defined as $\mathrm{Ad}_g(X) = gXg^{-1}$ is a matrix Lie group homomorphism. Since $\mathrm{Ad}_{gh} = \mathrm{Ad}_g \circ \mathrm{Ad}_h$, it follows that Ad is a representation of G on $\mathfrak{g}$.

On the other hand, $(\mathrm{Ad})_* = \mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ (again, $V = \mathfrak{g}$ here) defined as $\mathrm{ad}_X(Y) = [X, Y]$ is a representation, since $\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y]$.

Claim. Let $G = \mathrm{SO}(3)$ and $\mathfrak{g} = \mathfrak{so}(3) \cong \mathbb{R}^3$ (isomorphic as vector spaces). Then the standard representation of $\mathrm{SO}(3)$ on $\mathbb{R}^3$ and the adjoint representation of $\mathrm{SO}(3)$ on $\mathfrak{so}(3)$ are equivalent/isomorphic.

Proof. You will prove this in Assignment #4. $\square$

Definition 12.12. We define V_m to be a complex vector space of homogeneous polynomials of degree $m \geq 0$ in two complex variables. That is,

$$V_m := \{a_0 z_1^m + a_2 z_1^{m-1} z_2 + \cdots + a_{m-1} z_1 z_2^{m-1} + a_m z_2^m : a_i \in \mathbb{C}\}.$$

Thus the basis of V_m is $\{z_1^k z_2^{m-k} : k = 0, 1, \ldots, m\}$. Thus $\dim V_m = m + 1$.

13. February 11

We will show that, up to equivalence, the following representations are all the finite-dimensional complex irreducible representations of SU(2).

13.1. *A complex irreducible representation of* SU(2) *of dim* $m+1$ *(*$m \geq 0$*)*

Remark 13.1. Recall that SU(2) acts on $\mathbb{C}^2$ in the standard way: for any $A \in \mathrm{SU}(2)$ and $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \mathbb{C}^2$, we have

$$A = \begin{bmatrix} A_{11}z_1 + A_{12}z_2 \\ A_{21}z_1 + A_{22}z_2 \end{bmatrix} \in \mathbb{C}^2.$$

Define $\Pi_m : \mathrm{SU}(2) \to \mathrm{GL}(V_m)$ to be

$$(\Pi_m(A)f)(z) := f(A^{-1}z).$$

If $A, B \in \mathrm{SU}(w)$, then

$$(\Pi_m(A)\Pi_m(B)f)(z) = (\Pi_m(B)f)(A^{-1}z) = f(B^{-1}A^{-1}z) = f((AB)^{-1}z).$$

Hence $\Pi_m(A)\Pi_m(B)f = \Pi_m(AB)f$ for all $f \in V_m$. So Π_m is indeed a representation of SU(2).

If $f = a_0z_1^m + \cdots + a_mz_2^m = \sum_{k=0}^{m} a_kz_1^{m-k}z_2^k$, then

$$(\Pi_m(A)f)(z_1,z_2) = \sum_{k=0}^{m} a_k((A^{-1})_{11}z_1 + (A^{-1})_{12}z_2)^{m-k}((A^{-1})_{21}z_1 + (A^{-1})_{22}z_2)^k,$$

hence $\Pi_m(A)f(z_1,z_2) \in V_m$ for all $A \in \mathrm{SU}(2)$. Hence, Π_m induces a representation π_m of $\mathfrak{su}(2)$ on the same space V_m

$$\pi_m(X) = \left.\frac{d}{dt}\right|_{t=0} \Pi_m(e^{tX})$$
$$(\pi_m(X)f)(z) = \left.\frac{d}{dt}\right|_{t=0} f(e^{-tX}z).$$

Let $z(t) = e^{-tX}z$. Apply the chain rule:

$$\begin{aligned}(\pi(X)f)(z) &= \left.\frac{d}{dt}\right|_{t=0} f(z(t)) = \frac{\partial f}{\partial z_1}(z)\left.\frac{dz_1}{dt}\right|_{t=0} + \frac{\partial f}{\partial z_2}(z)\left.\frac{dz_2}{dt}\right|_{t=0} \\ &= (-X_{11}z_1 - X_{12}z_2)\frac{\partial f}{\partial z_1}(z) + (-X_{21}z_1 - X_{22}z_2)\frac{\partial f}{\partial z_2}(z).\end{aligned}$$

Since $\mathfrak{sl}(2,\mathbb{C}) = \mathfrak{su}(2)_{\mathbb{C}}$, this π_m is a representation of $\mathfrak{sl}(2,\mathbb{C})$ given by the same formula.

Consider $\pi(X)$ for some specific $\mathfrak{sl}(2,\mathbb{C})$. For instance, consider

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

Recall that H, X, Y form a basis of $\mathfrak{sl}(2,\mathbb{C})$. So we have

$$\begin{aligned}
\pi_m(H)f &= -z_1\frac{\partial f}{\partial z_1} + z_2\frac{\partial f}{\partial z_2} = \left(-z_1\frac{\partial}{\partial z_1} + z_2\frac{\partial}{\partial z_2}\right)f\\
\pi_m(X)f &= -z_2\frac{\partial f}{\partial z_1} = \left(-z_2\frac{\partial}{\partial z_1}\right)f\\
\pi_m(Y)f &= -z_1\frac{\partial f}{\partial z_2} = \left(-z_1\frac{\partial}{\partial z_2}\right)f.
\end{aligned}$$

Observe that

$$\begin{aligned}
\pi_m(H)(z_1^k z_2^{m-k}) &= \left(-z_1\frac{\partial}{\partial z_1} + z_2\frac{\partial}{\partial z_2}\right)(z_1^k z_2^{m-k})\\
&= -kz_1^k z_2^{m-k} + (m-k)z_1^k z_2^{m-k}\\
&= (m-2k)z_1^k z_2^{m-k}.
\end{aligned}$$

hence, $\pi_m(H)$ is *diagonalizable* as an operator on V_m. Meanwhile, as for the remaining two:

$$\begin{aligned}
\pi_m(X)(z_1^k z_2^{m-k}) &= -kz_1^{k-1}z_2^{m-k+1}\\
\pi_m(Y)(z_1^k z_2^{m-k}) &= -(m-k)z_1^{k+1}z_2^{m-k-1}.
\end{aligned}$$

Claim. V_m is an *irreducible* $\mathbb{C}$-representation of $\mathfrak{sl}(2,\mathbb{C})$.

Proof. Let $W \neq \{0\}$ be an invariant subspace. We need to show that $W = V_m$. We claim that there exists $w \neq 0 \in W$. That is we need to find $w = a_0 z_1^m + a_1 z_1^{m-1}z_2 + \cdots + a_m z_2^m$ where at least one of a_k is non-zero. Let k_0 be the smallest integer with $a_{k_0} \neq 0$. So we can write $w = a_{k_0} z_1^{m-k_0} z_2^{k_0} + \cdots + a_m z_2^m$.

Now consider $\pi_m(X)^{m-k_0}w = a_{k_0}(-1)^{m-k_0}(m-k_0)!z_2^m \neq 0$. Therefore $z_2^m \in W$. Note that $\pi_m(X)^{m-k_0}(w) \in W$ since W is invariant under π_m. Then $\pi_m(Y)^k(z_2^m) = (*)z_1^k z_2^{m-k}$ where $(*)$ is some non-zero stuff. This means $z_1^k z_2^{m-k} \in W$ for all $0 \leq k \leq m$ meaning $W = V_m$. □

13.2. *Complex representations of* $\mathfrak{sl}(2,\mathbb{C})$

We can just use the same basis for $\mathfrak{sl}(2,\mathbb{C})$. Note that $[H, X] = 2X$. Similarly we have $[H, Y] = -2Y$. Finally, note $[X, Y] = H$.

Let V be a complex vector space. Let $A, B \in \mathfrak{gl}(V)$. If $[A, B] = 2B, [A, C] = -2C, [B, C] = A$ then the map $\pi : \mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(V)$ where $\pi(H) = A, \pi(X) = B, \pi(Y) = C$ is a complex representation of $\mathfrak{sl}(2,\mathbb{C})$.

Remark 13.2. Consider $\pi(H) \in \mathfrak{gl}(V)$. We know that there exists at least one eigenvector, i.e., there exists non-zero $u \in V$ and $\alpha \in \mathbb{C}$ such that $\pi(H)u = \alpha u$.

Lemma 13.1. *Let* $\pi : \mathfrak{gl}(2, C) \to \mathfrak{gl}(V)$ *be a complex,* not necessarily irreducible *representation of* $\mathfrak{sl}(2,\mathbb{C})$*. Then* $\pi(H)\pi(X)u = (\alpha+2)\pi(X)u$ *and* $\pi(H)\pi(Y)u = (\alpha-2)\pi(Y)u$*. Hence, either* $\pi(X)u = 0$ *or* $4\pi(X)u$ *is an eigenvector of* $\pi(H)$ *with eigen value* $\alpha + 2$*. Similarly, either* $\pi(Y) = u$ *or* $\pi(Y)$ *is an eigenvector of* $\pi(H)$ *with eigenvalue* $\alpha - 2$*.*

Proof. $\pi(H)\pi(X)u - \pi(X)\pi(H)u = [\pi(H), \pi(X)]u = \pi([H, X])u = \pi(2X)u = 2\pi(X)u$. Therefore $\pi(H)\pi(X)u - \alpha\pi(X)u = 2\pi(X)u$. One can prove the other claim in a similar manner. □

Theorem 13.2 (Main theorem)**.** *The following are true:*

(1) *For every integer $m \geq 0$, there exists an irreducible complex representation of $\mathfrak{sl}(2, \mathcal{C})$ with dimension $m+1$.*

(2) *Any two irreducible complex representations of $\mathfrak{sl}(2, \mathbb{C})$ with same dimension are equivalent.*

(3) *If π is an irreducible complex representation of $\mathfrak{sl}(2, \mathbb{C})$ of dimension $m+1$, then $\pi \cong \pi_m$.*

Proof. First, we want to diagonalize $\pi(H)$, since we don't know yet that $\pi(H)$ is always diagonalizable. By Lemma 13.1, there exists a non-zero u so that $\pi(H)u = \alpha u$ for some $\alpha \in \mathbb{C}$, and $\pi(X)^n u = 0$ or $\pi(X)^n u$ an eigenvector of $\pi(H)$ with eigenvalue $\alpha + 2n$ for all $n \in \mathbb{N}$.

Recall that eigenvectors with distinct eigenvalues are linearly independent; and since V is finite-dimensional, there exists a non-negative integer N such that $\pi(X)^N u \neq 0$ but $\pi(X)^{NH} u = 0$. Let $u_0 = \pi(X)^N u \neq 0$. Then $\pi(X)u_0 = 0$ and $\pi(H)(u_0) = (\alpha + 2N)u_0 = \lambda u_0$, where $\pi = \alpha + 2N$.

Define $u_k = \pi(Y)^k u_0$. By Lemma 13.1, either $u_k = 0$ or u_k is an eigenvector of $\pi(H)$ with eigenvalue $\lambda - 2k$. So there exists $m \geq 0$ such that $u_k = \pi(Y)^k u_0 \neq 0$ for all $k \leq m$ and $u_{m+1} = \pi(Y)^{m+1} u_0 = 0$. We need the following claim to proceed further:

Claim. $\pi(X)u_0 = 0$, and $\pi(X)u_k = (k\lambda - k(k-1))u_{k-1}$ for $k > 0$.

Proof of the claim. We prove by induction on k. If $k = 1$, then $\pi(X)u_1 = \pi(X)\pi(Y)u_0 = \pi(H)u_0 = \lambda u_0$ since $\pi(X)u_0 = 0$.

Now assume that the claim holds for k (the induction hypothesis). Observe that

$$\begin{aligned}
\pi(X)u_{k+1} = \pi(X)\pi(Y)u_k &= \pi(Y)\pi(X)u_k + \pi(H)u_k \\
&= \pi(Y)[k\lambda - k(k-1)]u_{k-1} + (\lambda - 2k)u_k \\
&= (k\lambda - k^2 + k + \lambda - 2k)u_k \\
&= ((k+1)\lambda - (k+1)k)u_k.
\end{aligned}$$

□

$u_{m+1} = 0$ so $\pi(X)u_{m+1} = 0$. By the claim above we have $\pi(X)u_{m+1} = ((m+1)\lambda - (m+1)m)u_m = 0$, or $(m+1)(\lambda - m)u_m = 0$. Therefore $\lambda = m$, which is a non-negative integer.

So far, we have that there exist a non-negative integer m such that $u_0, \dots, u_m$ are non-zero vectors such that $\pi(H)u_k = (m-2k)u_k, \pi(Y)u_k = u_{k+1}, \pi(Y)u_m = 0, \pi(X)u_k = (km - k(k-1))u_{k-1}$, and $\pi(X)u_0 = 0$. Call the collection of these facts $(*)$.

Let $W = \mathrm{span}_{\mathbb{C}}\{u_0, \dots, u_m\}$. By construction, W is an invariant subspace for π. If $\pi : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{gl}(V)$ is an irreducible complex representation of $\mathfrak{sl}(2, \mathbb{C})$ then $V = \mathrm{span}\{u_0, \dots, u_m\}$.

Claim. $(*)$ is also sufficient. That is, if $u_0, \dots, u_m$ is a basis for an $(m+1)$-dimensional complex vector space V and if we define $\pi(H), \pi(X), \pi(Y)$ acting on V by $(*)$, and extend by linearly to an action of $\mathfrak{sl}(2, \mathbb{C})$, then this *is* a Lie algebra representation.

Proof of the claim. Let $k > 0$. Then $[\pi(H), \pi(X)]u_k = \pi(H)\pi(X)u_k - \pi(X)\pi(H)u_k = (m - 2(k-1))\pi(X)u_k - (m-2k)\pi(X)u_k$. Hence $2\pi(X)u_k = \pi(2X)u_k = \pi([H, X])u_k$. Also $\pi([H, X])u_0 = \pi(2X)u_0 = 0$. So $\pi([H, X]) = [\pi(H), \pi(X)]$. Also, $\pi([H, X])u_0 = \pi(2X)u_0 = 0$. Thus $\pi([H, X]) = [\pi(H), \pi(X)]$. Similarly, $[\pi(H), \pi(Y)] = \pi(-2Y) = \pi([H, Y])$. And finally (for $0 < k < m$),

$$\begin{aligned}[\pi(X), \pi(Y)]u_k &= \pi(X)\pi(Y)u_k - \pi(Y)\pi(X)u_k = \pi(X)u_{k+1} - \pi(Y)[km - k(k-1)]u_{k-1} \\ &= [m(k+1) - (k+1)k]u_k - [km - k(k-1)]u_{k-1} \\ &= (m-2k)u_k = \pi(H)u_k.\end{aligned}$$

Same holds for $k = 0, m$ so $[\pi(X), \pi(Y)]u_k = \pi(H)u_k$. Therefore $[\pi(X), \pi(Y)] = \pi([X, Y])$ so π is an $\mathfrak{sl}(2, \mathbb{C})$-Lie algebra representation. □

Now let $\pi : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{gl}(V), \widetilde{\pi} : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{gl}(\widetilde{V})$ be two finite-dimensional irreducible complex representations of $\mathfrak{sl}(2, \mathbb{C})$ of same dimension $m + 1$. We have seen that there exist a basis $\{u_0, \ldots, u_m\}$ of V and a basis $\{\widetilde{u_0}, \ldots, \widetilde{u_m}\}$ of $\widetilde{V}$ such that $(*)$ holds. Then the linear map $T : V \to \widetilde{V}$ defined by $T(u_k) = \widetilde{u_k}$ is an equivalence of representations. So T is a bijective morphism because, for any $z \in \mathfrak{sl}(2, \mathbb{C})$, the following diagram

$$\begin{CD} V @>{\pi(z)}>> V \\ @V{T}VV @VV{T}V \\ \widetilde{V} @>>{\widetilde{\pi}(z)}> \widetilde{V} \end{CD}$$

commutes, i.e., $\widetilde{\pi}(z)T(u_k) = \widetilde{\pi}(z)\widetilde{u_k} = T(\pi(z)u_k)$. This proves (2) so the proof is complete. □

Remark 13.3 (Summary of what we have done). Let π be a finite-dimensional complex representation of $\mathfrak{sl}(2, \mathbb{C})$ acting on V, where π is not necessarily irreducible. Then:

(1) every eigenvalue of $\pi(H)$ is an *integer*.
(2) If $v \in V$ is non-zero such that $\pi(X)v = 0$ and $\pi(H) = \lambda v$, then λ is a non-negative integer m and $\{v, \pi(Y)v, \ldots, \pi(Y)^m v\}$ is an irreducible invariant subspace of V of dimension $m + 1$.

Remark 13.4 (What's coming up). First, given a finite-dimensional representation of a matrix Lie group or a Lie algebra, we want to use them to construct new representations. Let V, W be finite-dimensional vector space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. There are three basic constructions: direct sum $V \oplus W$, dual space $V^* := L(V, \mathbb{F})$, and tensor product $V \otimes W$.

If Π, Σ are representations of G with $\Pi : G \to \mathrm{GL}(V)$ and $\Sigma : G \to \mathrm{GL}(W)$, we want to define

$$\begin{gathered}\Pi \oplus \Sigma : G \to \mathrm{GL}(V \oplus W) \\ \Pi^* : G \to \mathrm{GL}(V^*) \\ \Pi \otimes \Sigma : G \to \mathrm{GL}(V \otimes W),\end{gathered}$$

and the same can be defined for Lie algebra representations. Note that $\Pi \oplus \Sigma$ is *never irreducible* since $V \oplus \{0\}$ and $\{0\} \oplus W$ are non-trivial invariant subspaces. Now suppose that $\Pi : G \to \mathrm{GL}(V)$ is a finite-dimensional complex representation of G. So the question: is Π equivalent (isomorphic) to a direct sum of irreducible representation? The answer is, unfortunately, *not always*. But this holds only for "nice" groups or Lie algebras. This prompts us to introduce a new definition.

Definition 13.3. A representation Π is *completely reducible* if it is isomorphic to a direct sum of irreducible representations.

Remark 13.5. Not all representations are completely reducible. Also, note that any irreducible representation is completely reducible.

Remark 13.6. Another question: if Π, Σ are *irreducible* representations, then is $\Pi \otimes \Sigma$ irreducible also? The answer is again unfortunately no in general. If G has the property that all finite-dimensional representations are completely reducible, then $\Pi \otimes \Sigma \cong \bigoplus_{k=0}^{N} \mathcal{P}_k$ where each $\mathcal{P}_k$ is an irreducible representation.

14. February 23: Constructing new representations from existing ones

14.1. *First way: direct sum*

Definition 14.1. Let G be a matrix Lie group. Let $\Pi_1, \ldots, \Pi_m$ be representations of G acting on $V_1, V_2, \ldots, V_m$ respectively, all over some field. Then the direct sum $\Pi_1 \oplus \cdots \oplus \Pi_m$ of G acting on $V_1 \oplus V_2 \oplus \cdots \oplus V_m$ is given by (for any $g \in G$)

$$(\Pi_1 \oplus \Pi_2 \oplus \cdots \oplus \Pi_m)(g)(v_1 \oplus v_2 \oplus \cdots \oplus v_m) = (\Pi_1(g)v_1) \oplus \cdots \oplus (\Pi_m(g)v_m),$$

for each $v_i \in V_i$. We shall abuse notation by writing

$$g \cdot (v_1 \oplus v_2 \oplus \cdots \oplus v_m) = (gv_1) \oplus \cdots \oplus (gv_m),$$

and

$$\begin{aligned} g(h \cdot (v_1 \oplus v_2)) &= g((hv_1) \oplus (hv_2)) = (g(hv_1)) \oplus (g(hv_2)) \\ &= ((gh)v_1) \oplus ((gh)v_2) \ (\because \Pi_1, \Pi_2 \text{ are representations}) \\ &= (gh) \cdot (v_1 \oplus v_2). \end{aligned}$$

So $\Pi_1 \oplus \cdots \oplus \Pi_m : G \to \mathrm{GL}(V_1 \oplus V_2 \oplus \cdots \oplus V_m)$ is a representation of G.

Similarly, let $\pi_1, \ldots, \pi_m$ be representations of $\mathfrak{g}$ (of a Lie algebra) acting on $V_1, \ldots, V_m$ respectively. Then

$$(\pi_1 \oplus \cdots \oplus \pi_m)(X)(v_1 \oplus \cdots \oplus v_m) = (\pi_1(X)v_1) \oplus \cdots \oplus (\pi_m(X)v_m).$$

Thus $\pi_1 \oplus \cdots \oplus \pi_m : \mathfrak{g} \to \mathfrak{gl}(V_1 \oplus \cdots \oplus V_m)$ is a Lie algebra representation.

Definition 14.2. Let G be a matrix Lie group. We say G has the *complete reducibility property* if every finite-dimensional *complex* representation of G is isomorphic to a direct sum of irreducible representations. Similarly, we define CRP for a Lie algebra $\mathfrak{g}$.

Remark 14.1. Not every G or $\mathfrak{g}$ has the CRP. We will see later on soon.

14.2. *Second way: tensor product of representations*

Definition 14.3. Let G be a matrix Lie group, and Π_1, Π_2 representations of G acting on V_1, V_2 respectively. Then the *tensor product* $\Pi_1 \otimes \Pi_2$ of Π_1 and Π_2 is given by

$$(\Pi_1 \otimes \Pi_2)(g)(v_1 \otimes v_2) := (\Pi_1(g)v_1) \otimes (\Pi_2(g)v_2)$$

where $v_i \in V_i$ (extend this definition by linearity). Then again, by abuse of notation, we write $g(v_1 \otimes v_2) = (gv_1) \otimes (gv_2)$.

Proposition 14.4. $\Pi_1 \otimes \Pi_2$ *is a representation.*

Proof. First we need to show that it is well-defined. Let $\lambda \in \mathbb{F}$. Then

$$(\lambda v_1) \otimes v_2 = v_1 \otimes (\lambda v_2) = \lambda(v_1 \otimes v_2).$$

Therefore

$$\begin{aligned} g((\lambda v_1) \otimes v_2) &= (g(\lambda v_1)) \otimes (gv_2) \\ &= (\lambda(gv_1)) \otimes (gv_2)(\because g \text{ is linear}) \\ &= \lambda((gv_1) \otimes (gv_2)), \end{aligned}$$

and similarly,

$$g(v_1 \otimes (\lambda v_2)) = \lambda((gv_1) \otimes (gv_2)),$$

so $\lambda g(v_1 \otimes v_2) = g(\lambda(v_1 \otimes v_2))$. Similarly, since $(v_1 + u_1) \otimes v_2 = v_1 \otimes v_2 + u_1 \otimes v_2$. Apply g on both sides to get $(gv_1 + gu_1) \otimes gv_2 = gv_1 \otimes gv_2 + gu_1 \otimes gv_2$. Hence the representation is well-defined.

It is easy to check that $\Pi_2 \otimes \Pi_2 : G \to \mathrm{GL}(V_1 \otimes V_2)$ is a homomorphism, so indeed it is a representation. □

Let $\pi_k : \mathfrak{g} \to \mathfrak{gl}(V_k)$ for $k = 1, 2$ be representations of $\mathfrak{g}$. We want to define $\pi_1 \otimes \pi_2 : \mathfrak{g} \to \mathrm{GL}(V_1 \otimes V_2)$. Let $X \in \mathfrak{g}$. Define (again with abuse of notation)

$$\begin{aligned} (\pi_1 \otimes \pi_2)(X)(v_1 \otimes v_2) &:= (\pi_1(X)v_1) \otimes v_2 + v_1 \otimes (\pi_2(X)v_2) \\ X(v_1 \otimes v_2) &:= (Xv_1) \otimes v_2 + v_1 \otimes (Xv_2), \end{aligned}$$

and extend by linearity.

Recall that a representation of Lie algebras is a homomorphisms of Lie algebras. So in particular it is a linear map. Note that if we define $(\pi_1 \otimes \pi_2)(X)(v_1 \otimes v_2) = \pi_1(X)v_1 \otimes \pi_2(X)v_2$ is *no longer linear*.

If $\pi_k = (\Pi_k)_*$ i.e., induced from $\Pi_k : G \to \mathrm{GL}(V_k)$, then

$$\pi_k(X)v = \frac{d}{dt}\bigg|_{t=0} \Pi_k(e^{tX})v.$$

Since $e^{\pi_k(X)} = \Pi_k(e^X)$, we have, for $\Pi_1 \otimes \Pi_2 : G \to \mathrm{GL}(V_1 \otimes V_2)$,

$$\begin{aligned} (\Pi_1 \otimes \Pi_2)_*(X)(v_1 \otimes v_2) &= \frac{d}{dt}\bigg|_{t=0} (\Pi_1 \otimes \Pi_2)(e^{tX})(v_1 \otimes v_2) \\ &= \frac{d}{dt}\bigg|_{t=0} (\Pi_1(e^{tX})v_1) \otimes (\Pi_2(e^{tX})v_2) \\ &= \pi_1(X)v_1 \otimes Iv_2 + Iv_1 \otimes \pi_2(X)v_2 \quad \text{(by the product rule)}. \end{aligned}$$

Thus $(\Pi_1 \otimes \Pi_2)_*(X) = \pi(X) \otimes Iv_2 + Iv_1 \otimes \pi_2(X)$. One can also check directly that $\pi_1 \otimes \pi_2$ as defined earlier is a Lie algebra homomorphism.

Remark 14.2. If V_1, V_2 are irreducible representations of G or $\mathfrak{g}$, then $V_1 \otimes V_2$ *need not be irreducible in general.*

14.3. *Third way: dual representation*

Definition 14.5. Let $T : V \to W$ be a linear map of vector spaces. Then $T^t : W^* \to V^*$ is the *dual linear map* if for all $\alpha \in W^*$ and $T^t\alpha \in V^*$, then $(T^t\alpha)(v) = \alpha(T(v))$ for all $v \in V$.

Suppose γ is a basis of W and β is a basis of V. Then $A = [T]_{\gamma,\beta}$ is a $\dim W \times \dim V$ matrix. Then if β^* is a dual basis of V^* and γ^* is a dual basis of W^*, tun $B = [T^t]_{\beta^*,\gamma^*}$ is a $\dim V \times \dim W$ matrix, and $B = A^t$.

Definition 14.6. Let $\Pi : G \to \mathrm{GL}(V)$ be a representation of G. Define the *dual representation* $\Pi^* : G \to \mathrm{GL}(V^*)$ by $\Pi^*(g) = (\Pi(g^{-1})^t : V^* \to V^*$.

Remark 14.3. We see that Π^* is a homomorphism, since, for any $g, h \in G$,

$$\begin{aligned}\Pi^*(g)\Pi^*(h) &= (\Pi(g^{-1}))^t(\Pi(h^{-1}))^t = (\Pi(h^{-1})\Pi(g^{-1}))^t \\ &= (\Pi(h^{-1}g^{-1}))^t = (\Pi((gh)^{-1}))^t = \Pi^*(gh).\end{aligned}$$

Proposition 14.7. *Let $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ be a Lie algebra representation. Then the dual representation $\pi^* : \mathfrak{g} \to \mathfrak{gl}(V^*)$ is given by*

$$\pi^*(X) := -(\pi(X))^t.$$

Proof. Since the Lie bracket is a bracket in $\mathfrak{gl}(V^*)$, the bracket is the usual commutator.

$$\begin{aligned}[\pi^*(X), \pi^*(Y)] &= \pi^*X\pi^*Y - \pi^*Y\pi^*X \\ &= (-\pi X)^t(-\pi Y)^t - (-\pi Y)^t(-\pi X)^t = ((\pi Y)(\pi X))^t - ((\pi X)(\pi Y))^t \\ &= [\pi Y, \pi X]^t = (\pi([Y, X]))^t \quad (\because \pi \text{ is a representation}) \\ &= (\pi(-[X, Y]))^t = \pi^*([X, Y]).\end{aligned}$$

□

Recall that an intertwining map is a linear map $T : V \to W$, where V and W are representations of a matrix Lie group G or a Lie algebra $\mathfrak{g}$ satisfying $T(g \cdot v) = g \cdot T(v)$ for all $g \in G$ (or $T(X \cdot v) = X \cdot T(v)$ for all $X \in \mathfrak{g}$).

14.4. *Schur's lemma*

Theorem 14.8 (Schur's lemma)**.** *The following are true:*

1. *Let V and W be* irreducible *representations of a matrix Lie group or a Lie algebra, and let $T : V \to W$ be an intertwining map. Then either $T \equiv 0$ or T is an isomorphism. Therefore, any morphism between two* irreducible representations *is either the zero map or an isomorphism.*
2. *Let V be an* irreducible complex *representation of a matrix Lie group or Lie algebra. Then if $T : V \to V$ is an intertwining map of V with itself then $T = \lambda I$ for some $\lambda \in \mathbb{C}$.*
3. *Let V, W be* irreducible complex *representation of a matrix Lie group or Lie algebra. Let $T_1, T_2 : V \to W$ be two* non-zero *intertwining maps. Then $T_1 = \lambda T_2$ for some $\lambda \in \mathbb{C}^*$.*

Proof. We will only do the Lie group case, since the Lie algebra case only requires modifying notation.

(Proof of (1)) Suppose that $g \cdot T(v) = T(g \cdot v)$ for all $v \in V, g \in G$. Let $v \in \ker(T)$. Then $T(g \cdot v) = g \cdot T(v) = g \cdot 0 = 0$. Thus $g \cdot v \in \ker(T)$. Thus $\ker(T)$ is an invariant subspace of V. Let $w \in \mathrm{im}(T)$. So $w = T(v)$ for some $v \in V$, whence $g \cdot w = g \cdot T(v) = T(g \cdot v) \in \mathrm{im}(T)$.

It follows that $\text{im}(T)$ is an invariant subspace of W. But then since V and W are both irreducible, $\ker(T)$ is either trivial or the entire V; similarly, $\text{im}(T)$ is either trivial or the entire W. Therefore either $T \equiv 0$ (if $\ker(T) = V, \text{im}(T) = 0$) or T is an isomorphism (if $\ker(T) = 0, \text{im}(T) = W$).

(Proof of (2)) Let $T(g \cdot v) = g \cdot T(v)$. There exists at least one eigenvector $v_0 \neq 0$ and eigenvalue $\lambda \in \mathbb{C}$ for T, i.e., $T(v_0) = \lambda v_0$.

Let E_λ be the eigenspace with eigenvalue λ. If $v \in E_\lambda$, then $T(g \cdot v) = g \cdot T(v) = g \cdot (\lambda v) = \lambda \cdot (g \cdot v)$, by the linearity of the group action. Hence E_λ is a non-zero invariant subspace of V, so $E_\lambda = V$. Thus $T(x) = \lambda x$ for all $x \in V$.

(Proof of (3)) By (1), both T_1, T_2 are isomorphisms, Hence $T_1 \circ T_2^{-1} : W \to W$ is an intertwining map. By (2), we have $T_1 \circ T_2^{-1} = \lambda I$ for some $\lambda \in \mathbb{C}^*$. □

Corollary 14.9. *Let* Π *be an* irreducible complex *representation of* G. *Let* $Z(G)$ *be the centre of* G, *i.e.,* $Z(G) = \{g \in G : gh = hg \, \forall h \in G\}$. *Then if* $g \in Z(G)$, *we have* $\Pi(g) = \lambda I$ *for some* $\lambda \in \mathbb{C}$. *That is, every element in* $Z(G)$ *acts by scalar multiplication.*

Similarly, if π *is an* irreducible complex *representation of* $\mathfrak{g}$, *let* $z(\mathfrak{g}) = \{Y \in \mathfrak{g} : [X, Y] = 0 \, \forall X \in \mathfrak{g}\}$, *or the centre of* $\mathfrak{g}$. *Then if* $X \in z(\mathfrak{g})$, *then* $\pi(X) = \lambda I$ *for some* $\lambda \in \mathbb{C}$.

Proof. Let $g \in Z(G)$. Then for any $h \in G$, we have $\Pi(g)\Pi(h) = \Pi(gh) = \Pi(hg) = \Pi(h)\Pi(g)$. So $\Pi(g)$ is an intertwining map of V to itself. Thus by (2) of Schur's lemma, indeed $\Pi(g) = \lambda I$. The Lie algebra case can be proved similarly. □

Corollary 14.10. *Any complex irreducible representation of an* abelian *group or an* abelian *Lie algebra is one-dimensional.*

Proof. Since $Z(G) = G$, by the preceding corollary, for any $g \in G$ we have $\Pi(g) = \lambda_g I$ for some $\lambda_G \in \mathbb{C}$. Hence, every subspace of V is invariant. But since V is irreducible, it cannot have any non-trivial subspace. Therefore $\dim(V) = 1$. □

14.5. *Relation between representations of matrix Lie groups and Lie algebras*

We have seen that if $F : G \to H$ is a homomorphism, it induces $F_* : \mathfrak{g} \to \mathfrak{h}$ a Lie algebra homomorphism such that $F(e^X) = e^{F_* X}$ for all $X \in \mathfrak{g}$. Also, if G is *simply connected*, then we have a converse: that is, given $\lambda : \mathfrak{g} \to \mathfrak{h}$ a Lie algebra homomorphism, there exists a *unique* $F : G \to H$ such that $F_* = \lambda$.

Any Lie group representation $\Pi : G \to \text{GL}(V)$ always induces a Lie algebra representation $\Pi_* = \pi : \mathfrak{g} \to \mathfrak{gl}(V)$. Also, we can go back provided G is simply connected. Therefore, we can conclude the following:

(1) If G is simply connected, then we have a *one-to-one correspondence* between representations of G and representations of $\mathfrak{g}$.
(2) Moreover, this one-to-one correspondence restricts to the *irreducible representations*.
(3) However, if G is *not* simply connected, then we *may not have* such one-to-one correspondence. We can always go from G-representations to $\mathfrak{g}$-representations, but we can't always "lift" a $\mathfrak{g}$-representation to a G-representation.

Example 14.11. $G = \text{SU}(2)$ is homeomorphic to S^3, so G is indeed simply connected. Therefore we have the bijective correspondence between irreducible complex representations of $\text{SU}(2)$ and of $\mathfrak{su}(2)$. We also have the bijective correspondence between the irreducible complex representations of $\mathfrak{su}(2)$ and those of $\mathfrak{su}(2)_\mathbb{C} = \mathfrak{sl}(2, \mathbb{C})$ (the complexification of

$\mathfrak{su}(2)$). Recall that we found all of the complex irreducible representations of $\mathfrak{sl}(2,\mathbb{C})$ with $\dim(V_m) = m+1$ unique up to isomorphism. Recall that $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ which gives an isomorphism $\mathfrak{g} \cong \mathfrak{h}$. Thus there exists a one-to-one correspondence between irreducible complex representations of $\mathfrak{g}$ and those of $\mathfrak{h}$. However, recall that SO(3) is *not* simply connected. We shall continue the discussion next class.

15. February 25

Recall the following bijective correspondences, which preserve irreducibility:

$$\text{complex rep's of } G \Leftrightarrow \text{complex rep's of } \mathfrak{g} \Leftrightarrow \text{complex rep's of } \mathfrak{g}_{\mathbb{C}}.$$

We considered an example where $G = \mathrm{SU}(2), \mathfrak{g} = \mathfrak{su}(2)$. Recall that $\mathfrak{g}_{\mathbb{C}} = \mathfrak{sl}(2,\mathbb{C})$. It is easy to see that if $\mathfrak{g} \cong \mathfrak{h}$ as a Lie algebra, then there is a one-to-one correspondence between the irreducible representations of $\mathfrak{g}$ and the irreducible representations of $\mathfrak{h}$. Thus we have a bijective correspondence between the complex irreducible representations of $\mathfrak{su}(2)$ and the complex irreducible representations of $\mathfrak{so}(3)$, since $\mathfrak{su}(2) \cong \mathfrak{so}(3)$. However SO(3) is *not* simply connected. So do there exist representations of $\mathfrak{so}(3)$ that *do not lift* to representations of SO(3)? Answer: Yes. We hope to prove this in today's lecture. For this, we need more details about the relation between SU(2) and SO(3).

Recall that $\mathfrak{su}(2)$ consists of skew-Hermitian 2×2 complex matrices, and that it is three-dimensional as a *real* vector space, with basis

$$E_1 = \frac{1}{2}\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad E_2 = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad E_3 = \frac{1}{2}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Recall also that $\mathfrak{so}(3)$ consists of skew-symmetric 3×3 real matrices, and that it is also a three-dimensional real vector space, with basis

$$F_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad F_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that $[E_i, E_j] = E_k$ and $[F_i, F_j] = F_k$ where (i,j,k) is a cyclic permutation of $(1,2,3)$. Therefore the map $T : \mathfrak{su}(2) \to \mathfrak{so}(3)$ defined by $E_i \mapsto F_i$ (and extend by linearity) is a Lie algebra isomorphism. Thus there is a one-to0one correspondence between the representations of $\mathfrak{su}(2)$ and the representations of $\mathfrak{so}(3)$ given by

$$\pi : \mathfrak{su}(2) \to \mathfrak{gl}(V) \to \pi \circ T^{-1} : \mathfrak{so}(3) \to \mathfrak{gl}(V)$$
$$\sigma \circ T \leftarrow \sigma.$$

We have determined, up to isomorphism, all the irreducible representations of $\mathfrak{su}(2)$. So let $\pi_m : \mathfrak{su}(2) \to \mathfrak{gl}(V_m)$ where $\dim V_m = m+1$ and $m \geq 0$. So by the correspondence, all the complex irreducible representations of $\mathfrak{so}(3)$ are of the form $\sigma_m := \pi_m \circ T^{-1} : \mathfrak{so}(3) \to \mathfrak{gl}(V_m)$. Also, since SU(2) is simply connected, there is a one-to-one correspondence between Π_m and $\pi_m := (\Pi_m)_*$. What of SO(3)? For each $m \geq 0$, does there exist a finite-dimensional irreducible complex representation $\Sigma_m : \mathrm{SO}(3) \to \mathrm{GL}(V_m)$ such that $(\Sigma_m)_* = \sigma_m$?

Lemma 15.1. *There exists a matrix Lie group homomorphism $P : \mathrm{SU}(2) \to \mathrm{SO}(3)$ that is two-to-one such that P is surjective with $\ker P = \{\pm I\}$ and such that $P_* : \mathfrak{su}(2) \to \mathfrak{so}(3)$ is the map T that sends E_i to F_i for $i = 1,2,3$.*

Proof. Recall that $\mathfrak{su}(2) \cong \mathfrak{so}(3) \cong \mathbb{R}^3$ as vector spaces *and* as Lie algebras. If $\{e_1, e_2, e_3\}$ is the standard basis of $\mathbb{R}^3$, then the map $E_i \leftrightarrow F_i \leftrightarrow e_i$ is an isomorphism.

Since $\mathrm{ad}_{E_i} : \mathfrak{su}(2) \to \mathfrak{su}(2)$ and $\mathfrak{su}(2) \cong \mathbb{R}^3$, we have $\mathrm{ad}_{E_i} \in \mathfrak{gl}(\mathbb{R}^3)$. We also have that $F_i e_j = e_k$ and $F_i e_k = -e_j$ where (i, j, k) is the cyclic permutation of $(1, 2, 3)$. Under this identification, we have

$$\mathrm{ad} : \mathfrak{su}(2) \to \mathfrak{so}(3) \subsetneq \mathfrak{gl}(\mathbb{R}^3),$$

and $\mathrm{ad}_{E_i} = F_i$. hence $T := \mathrm{ad} : \mathfrak{su}(2) \to \mathfrak{so}(3)$ is a Lie algebra isomorphism.

Now consider $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{GL}(\mathfrak{su}(2)) \cong \mathrm{GL}(3, \mathbb{R})$. by construction, Ad is a matrix Lie group homomorphism. We only need to show that:

(1) $\mathrm{im}(\mathrm{Ad}) = \mathrm{SO}(3) \subseteq \mathrm{GL}(3, \mathbb{R})$;
(2) Ad is surjective onto its image; and that
(3) $\ker(\mathrm{Ad}) = \{\pm I\}$.

Recall that if $\langle \_, \_ \rangle$ is the standard Hermitian inner product on 2×2 complex matrices, then

$$\langle A, B \rangle = \mathrm{tr}(A^* B) = \sum_{i,j} \overline{A_{ij}} B_{ij}.$$

When restricted to $\mathfrak{su}(2)$, then this inner product is, up to a factor of 2, the Euclidean inner product of $\mathbb{R}^3$. Therefore,

$$\left\langle \sum_i a_i E_i, \sum b_j E_j \right\rangle = 2 \sum_{i=1}^{3} a_i b_i.$$

Let $g \in \mathrm{SU}(2)$. Then $g^* = g^{-1}$. Let $v \in \mathbb{R}^3 \cong \mathfrak{su}(2)$ with $\mathrm{Ad}_g v = g v g^{-1}$. Thus it follows

$$\begin{aligned} \langle \mathrm{Ad}_g v, \mathrm{Ad}_g w \rangle &= \mathrm{tr}((g v g^{-1})^* g w g^{-1}) = \mathrm{tr}((g^{-1})^* v^* \underbrace{g^* g}_{I} w g^{-1}) \\ &= \mathrm{tr}(g v^* w g^{-1}) = \mathrm{tr}(v^* w g^{-1} g) = \mathrm{tr}(v^* w) = \langle v, w \rangle. \end{aligned}$$

Therefore $\mathrm{Ad}_g \in \mathrm{O}(3)$ and $\mathrm{Ad}_I = I \in \mathrm{SO}(3)$. Recall that Ad is continuous and SU(2) is connected, and that $\det(\mathrm{Ad}_g) = \pm 1$ for all $g \in \mathrm{SU}(2)$. So $\det(\mathrm{Ad}_g) = 1$ for all $g \in \mathrm{SU}(2)$. So indeed $\mathrm{im}(\mathrm{Ad}) \subseteq \mathrm{SO}(3)$, which proves (1).

Recall that, for any θ,

$$\exp \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\theta \\ 0 & \theta & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix},$$

i.e., the counterclockwise rotation by θ about e_1 axis. Since $e^{\theta F_1} = e^{\theta\, \mathrm{ad}_{E_1}} = e^{\mathrm{ad}_{\theta E_1}} = \mathrm{Ad}(e^{\theta E_1})$ and $e^{\theta E_1} \in \mathrm{SU}(2)$, it follows that $\mathrm{Ad}(e^{\theta E_1})$ is the counterclockwise rotation by θ about the e_1 axis. Similarly, we can show that $\mathrm{Ad}(e^{\theta E_j})$ for $j = 2, 3$ are counterclockwise rotations by θ about the e_j axis.

Let $R \in \mathrm{SO}(3)$. Then by the Cartan-Dieudonné theorem, R can be written as a product of rotations about e_1, e_2, e_3 axes. So we have

$$R = \prod_i \mathrm{Ad}(e^{\theta E_i}) = \mathrm{Ad}\left(\prod_i e^{\theta E_i} \right),$$

with the last equality following from the fact that Ad is a homomorphism. So $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is surjective, which proves (2).

For the last part, let $g \in \ker(\mathrm{Ad})$. Then $\mathrm{Ad}_g = I \in \mathrm{SO}(3) \subseteq \mathrm{GL}(\mathbb{R}^3) = \mathrm{GL}(\mathfrak{su}(2))$. Thus $\mathrm{Ad}_g E_i = E_i$ for all $i = 1, 2, 3$ SO $gE_ig^{-1} = E_i$. Hence $gE_i = E_ig$ fofr all i. So g is of the form

$$\begin{bmatrix} a & -\bar{b} \\ b & \bar{a} \end{bmatrix},$$

with $|a|^2 + |b|^2 = 1$. Since $g \in \mathrm{SU}(2)$, it follows that $|a|^2 = 1$ so $a = e^{i\theta}$. Now solve for g:

$$\begin{bmatrix} a & -\bar{b} \\ b & \bar{a} \end{bmatrix} E_1 = E_1 \begin{bmatrix} a & -\bar{b} \\ b & \bar{a} \end{bmatrix},$$

which gives

$$\begin{bmatrix} a & \bar{b} \\ b & -\bar{a} \end{bmatrix} = \begin{bmatrix} a & -\bar{b} \\ -b & -\bar{a} \end{bmatrix}.$$

Thus $b = -b$, or $b = 0$. Also,

$$\begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} E_2 = E_2 \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix},$$

so

$$\begin{bmatrix} 0 & e^{i\theta} \\ e^{-i\theta} & 0 \end{bmatrix} = \begin{bmatrix} 0 & e^{-i\theta} \\ e^{i\theta} & 0 \end{bmatrix},$$

hence $e^{i\theta} = e^{-i\theta}$. So $e^{i\theta} = \pm 1$, so $g = \pm I$, as required.

Now that we showed all (1), (2), and (3), we can say that $P = \mathrm{Ad}$ and $P_* = \mathrm{Ad}_* = \mathrm{ad} = T$, as we claimed initially. □

Theorem 15.2. *Let $\sigma_m = \pi_m \circ T^{-1}$ be the irreducible complex representations of $\mathfrak{so}(3)$ on V_m. Then:*

1. *If m is* even, *then there exists a representation of Σ_m of* SO(3) *on V_m such that $(\Sigma_m)_* = \sigma_m$.*
2. *If m is* odd, *then there* does not exist *such representation.*

Proof. We will start by proving (2) first. Suppose that there exists a representation Σ_m such that $(\Sigma_m)_* = \sigma_m$. Thus we must have

$$\Sigma_m(e^X) = e^{\sigma_m(X)},$$

for all $X \in \mathfrak{so}(3)$. Let $X = 2\pi F_1$. Then

$$e^{2\pi F_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(2\pi) & -\sin(2\pi) \\ 0 & \sin(2\pi) & \cos(2\pi) \end{bmatrix} = I.$$

So $I = \Sigma_m(e^{2\pi F_1}) = e^{2\pi\sigma_m(F_1)}$. So we have

$$\sigma_m(F_1) = (\pi_m \circ T^{-1})(F_1) = \pi_m(E_1) = \frac{i}{2}\pi_m(H),$$

where

$$E_1 = \frac{i}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \frac{i}{2}H$$

from the representation theory of $\mathfrak{su}(2)$ and $\mathfrak{sl}(2,\mathbb{C})$. So there exists a basis $u_0, u_1, \ldots, u_m$ of V_m such that $\pi_m(H)u_k = (m-2k)u_k$. So we have $\sigma_m(F_1)u_k = \frac{i}{2}(m-2k)u_k$. So in this basis, we have

$$\sigma_m(F_1) = \begin{bmatrix} \frac{i}{2}m & & & \\ & \frac{i}{2}(m-2 & & \\ & & \ddots & \\ & & & \frac{i}{2}(-m) \end{bmatrix}.$$

Since m is odd, $m-2k$ is also odd for all k. Hence

$$e^{2\pi\sigma_m(F_1)} = \begin{bmatrix} e^{\pi i m} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & e^{-m\pi i} \end{bmatrix} = -I, \qquad (\dagger)$$

which is a contradiction. Therefore such Σ_m cannot exist.

Now suppose that m is even. Consider the representation Π_m of SU(2). Then we have

$$e^{2\pi E_i} = \exp \begin{bmatrix} \pi i & 0 \\ 0 & -\pi i \end{bmatrix} = \begin{bmatrix} e^{\pi i} & 0 \\ 0 & e^{-\pi i} \end{bmatrix} = -I.$$

Therefore, we have

$$\Pi_m(-I) = \Pi_m(e^{2\pi E_1}) = e^{(2\pi)\pi_m(E_1)} = e^{2\pi\sigma_m(F_1)} = I,$$

by (†). Hence,

$$\Pi_m(-U) = \Pi_m((-I)U) = \Pi_m(-I)\Pi_m(U) = I\Pi_m(U) = \Pi_m(U),$$

with the second equality following from the fact that Π is a homomorphism. So $\Pi_m(U) = \Pi_m(-U)$ for all $U \in \mathrm{SU}(2)$. By the previous theorem. given $R \in \mathrm{SO}(3)$, there exists a *unique* pair $\{U, -U|\}$ in SU(2) such that $\mathrm{Ad}_U = \mathrm{Ad}_{-U} = R$. Now define $\Sigma_m : \mathrm{SO}(3) \to \mathrm{GL}(V_m)$ by $\Sigma_m(R) = \Pi_m(U)$. Note that this is well-defined since m is even. One can verify that Σ_m is a representation of SO(3) on V_m, and that $\Pi_m = \Sigma_m \circ \mathrm{Ad}$ by construction. So it follows that

$$(\Pi_m)_* = \pi_m = (\Sigma_m \circ \mathrm{Ad})_* = (\Sigma_m)_* \circ \mathrm{Ad}_* = \sigma_m \circ T,$$

from which it follows that $\sigma_m = \pi_m \circ T^{-1}$ as required. □

Remark 15.1. Hence, if G is not simply connected, then there can be complex representations of $\mathfrak{g}$ that do not a rise from representations of G.

Remark 15.2 (On complete reducibility). Recall the following definitions, which we covered last class. A finite-dimensional representation V of a group or an algebra is called *completely reducible* if it is isomorphic to a direct sum of irreducible representations. A group or algebra is said to have the compete reducibility property (CRP) if its every finite-dimensional representation is completely reducible.

Notice that if $\mathfrak{g}$ is the Lie algebra of a simply connected group G, then $\mathfrak{g}$ has CRP if and only if G has CRP. Also, we remark that the direct sum of irreducible representations is preserved by this one-to-one correspondence. However, it is possible for a group to have CRP even when it's not connected. Here is a fact: if G is *compact* (*not necessarily simply connected*), then it has CRP. The idea behind proving this is introducing an invariant inner product on V by "averaging" over G, but this needs "Haar measure", which is more or less a left-invariant volume form.

We will use the fact about compact groups in the following way. Recall that $(3, \mathbb{C}) = \mathfrak{su}(3)_{\mathbb{C}}$. Note that SU(3) is compact and simply connected. Thus, by the fact we just mentioned, SU(3) has CRP. Also, since SU(3) is connected, it follows that $\mathfrak{su}(3)$ and $\mathfrak{su}(3)_{\mathbb{C}} = \mathfrak{sl}(3, \mathbb{C})$ also have CRP. Generalizing this, we see that SU(n) is compact and simply connected, whence it follows that $\mathfrak{sl}(n, \mathbb{C})$ has CRP.

15.1. *Representation theory of* SU(3)

From now on, *all* representations are *finite-dimensional* and *complex.* We will re-visit the bijective correspondences between

irreducible rep's of SU(3) $\leftrightarrow$ irreducible rep's of $\mathfrak{su}(3)$ $\leftrightarrow$ irreducible rep's of $\mathfrak{sl}(3, \mathbb{C})$.

Recall for $\mathfrak{sl}(2, \mathbb{C})$, we showed that there exists *exactly one* (up to isomorphism) irreducible representation in any positive dimension $m \geq 0$, with $\dim(V_m) = m + 1$.

For $\mathfrak{sl}(3, \mathbb{C})$, we will "parametrize" the irreducible representations. We will see that if $m_1, m_2 \geq 0$ then we will get an irreducible representation of $\mathfrak{sl}(3, \mathbb{C})$ $V_{(m_1, m_2)} \cong V_{\widetilde{m_1}, \widetilde{m_2}} \Leftrightarrow m_i = \widetilde{m_i}$ $(i = 1, 2)$. This time, we will not get (in general) only one for each positive dimension.

We first consider $\mathfrak{sl}(2, \mathbb{C})$. Recall that $\mathfrak{sl}(2, \mathbb{C})$ consists of traceless 2×2 complex matrices. Recall also that

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

serves as a basis, and that $[H, X] = 2X, [H, Y] = -2Y, [X, Y] = H$. And we derived restriction on finite-dimensional representations of $\mathfrak{sl}(2, \mathbb{C})$ and in particular on irreducible representations.

As for $\mathfrak{sl}(3, \mathbb{C})$, note that $\mathfrak{sl}(3, \mathbb{C})$ is a complex vector space of dimension 8, with the basis

$$X_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad X_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$Y_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Y_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad Y_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

(Recall that H_1 and H_2 are basis for diagonal traceless matrices.)

16. March 2: Representations of $\mathfrak{sl}(3, \mathbb{C})$

Idea: given an irreducible representation (V, π) of $\mathfrak{sl}(3, \mathbb{C})$, we will try to simultaneously diagonalize $\pi(H_1), \pi(H_2)$ (as of now, it is not obvious that this is possible). Recall that for $\mathfrak{sl}(2, \mathbb{C})$, we found a basis of V for which $\pi(H)$ was diagonal. Note that $\pi(H_1), \pi(H_2)$ commute, since $[\pi(H_1), \pi(H_2)] = \pi([H_1, H_2]) = 0$, since H_1 and H_2 commute and π preserves the bracket.

Definition 16.1. Let $\pi : \mathfrak{sl}(3,\mathbb{C}) \to \mathfrak{gl}(V)$ be a representation. Then an ordered pair $\mu = (m_1, m_2) \in \mathbb{C}^2$ is called a *weight of* π if there exists some non-zero vector $v \in V$ such that

$$\begin{aligned} \pi(H_1)v &= m_1 v \\ \pi(H_2)v &= m_2 v. \end{aligned} \tag{$*$}$$

That is, a weight is a pair of simultaneous eigenvalues for $\pi(H_1), \pi(H_2)$. We call v in this case a *weight vector* for weight μ. If μ is a weight for π, then the space $W_\mu \subseteq V$ consisting of all weight vectors for weight μ plus zero vector is a subspace. The *multiplicity* of μ is $\dim(W_\mu)$.

Proposition 16.2. *Every representation of* $\mathfrak{sl}(3,\mathbb{C})$ *has at least one weight.*

Proof. $\pi(H)$ is an operator on V, so it has at least one eigenvalue m_1. So there exists non-zero $v \in V$, with $\mu(H_1)v = m_1 v$. Let E_{m_1} be the eigenspace of $\pi(H_1)$ with eigenvalue m_1. So $E_{m_1} \neq \{0\}$. Let $v \in E_{m_1}$. Since $[\pi(H_1), \pi(H_2)] = 0$, we have

$$\pi(H_1)\pi(H_2)v = \pi(H_2)\pi(H_1)v = m_1\pi(H_2)v.$$

So $\pi(H_2)$ maps E_{m_1} to E_{m_1}. Hence $\pi(H_2)|_{E_{m_1}}$ is an operator on E_{m_1}, so $\pi(H_2)|_{E_{m_1}}$ has an eigenvalue $m_2 \in \mathbb{C}$ and (m_1, m_2) is a weight for π. □

Proposition 16.3. *If* π *is a representation of* $\mathfrak{sl}(3,\mathbb{C})$ *and* $\mu = (m_1, m_2)$ *is a weight for* π *then* $m_1, M-2$ *are both* integers.

Proof. Let $\mathfrak{g}_k = \operatorname{span}\{X_k, Y_k, H_k\}$ with $k = 1, 2$. Note that each $\mathfrak{g}_k \cong \mathfrak{sl}(2,\mathbb{C})$. Restrict π to $\mathfrak{g}_k$ we get a representation of $\mathfrak{sl}(2,\mathbb{C})$. Suppose that $\pi(H_k)$ has eigenvalue m_k. We already proved that for every finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ the eigenvalues of $\pi(H)$ are integers. So $m_1, m_2 \in \mathbb{Z}$. □

Remark 16.1. The weights of π depend on the choice $\{H_1, H_2\}$ of basis vectors for the subspace of $\mathfrak{sl}(3,\mathbb{C})$ consisting of diagonal traceless matrices.

Definition 16.4. An ordered pair $\alpha = (a_1, a_2) \in \mathbb{C}^2$ is called a *root of* $\mathfrak{sl}(3,\mathbb{C})$ if

(1) $\alpha \neq (0,0)$
(2) there exists non-zero $z \in \mathfrak{sl}(3,\mathbb{C})$ such that $\operatorname{ad}_{H_1} z = [H_1, z] = a_1 z$ and $\operatorname{ad}_{H_2} z = [H_2, z] = a_2 z$.

Therefore, a root of $\mathfrak{sl}(3,\mathbb{C})$ is a non-zero weight for the adjoint representation of $\mathfrak{sl}(3,\mathbb{C})$. We call z a *root vector* for the root α.

Example 16.5. For $\mathfrak{sl}(3,\mathbb{C})$ we have the following six roots:

root	root vector z	root	root vector z
$(2,-1)$	X_1	$(-2,1)$	Y_1
$(-1,2)$	X_2	$(1,-2)$	Y_2
$(1,1)$	X_3	$(-1,-1)$	Y_3

Claim. There are no other roots.

Here's the strategy: suppose that $\alpha = (\alpha_1, \alpha_2)$ is a root. Then there exists $z \neq 0$ in $\mathfrak{sl}(3, \mathbb{C})$ such that $\operatorname{ad}_{H_k} z = a_k z$. We can write

$$z = t_1 H_1 + t_2 H_2 + \lambda_1 X_1 + \lambda_2 X_2 + \lambda_3 X_3 + \mu_1 Y_1 + \mu_2 Y_2 + \mu_3 Y_3,$$

and so

$$\begin{aligned} \operatorname{ad}_{H_1} z = a_1 z &= 0H_1 + 0H_2 + 2\lambda_1 X_1 - \lambda_2 X_2 + \lambda_3 X_3 - 2\mu_1 Y_1 + \mu_2 Y_2 - \mu_3 Y_3 \\ \operatorname{ad}_{H_2} z = a_2 z &= 0H_1 + 0H_2 - \lambda_1 X_1 + 2\lambda_2 X_2 + \lambda_3 X_3 + \lambda_1 Y_1 - 2\mu_2 Y_2 - \mu_3 Y_3. \end{aligned}$$

So we have

$$\begin{aligned} a_1 t_1 = 0, &\quad a_2 t_1 = 0, \\ a_1 t_2 = 0, &\quad a_2 t_2 = 0, \\ a_1 \lambda_1 = 2\lambda_1, &\quad a_2 \lambda_1 = -\lambda_1, \\ a_1 \lambda_2 = -\lambda_2, &\quad a_2 \lambda_2 = 2\lambda_2, \\ a_1 \lambda_3 = \lambda_3, &\quad a_2 \lambda_3 = \lambda_3, \\ a_1 \lambda_1 = -2\mu_1, &\, a_2 \lambda_1 = \lambda_1, \\ a_1 \mu_2 = \mu_2, &\quad a_2 \mu_2 = -2\mu_2, \\ a_1 \mu_3 = -\mu_3, &\quad a_2 \mu_3 = \mu_3. \end{aligned}$$

Since both a_1 and a_2 cannot be zero, we have $t_1, t_2 = 0$. Suppose that $\lambda_1 \neq 0$. Then $a_1 = 2$, and $\lambda_2, \lambda_3, \mu_1, \mu_2, \mu_3$ all zero and so forth.

More generally, we have the fact that weight vectors corresponding to distinct weight are *linearly independent*. So we can use this to argue that $X_1, X_2, X_3, Y_1, Y_2, Y_3$ are the only root vectors for $\mathfrak{sl}(2, \mathbb{C})$. So the roots are non-zero eigenvalues of ad_H. Since $\operatorname{ad}_H X = 2X$ and $\operatorname{ad}_H Y = -2Y$, we have

roots	(2)	(-2)
root vector	X	Y

Lemma 16.6. *Let $\alpha = (a_1, a_2)$ be a root of $\mathfrak{sl}(3, \mathbb{C})$ be a root of $\mathfrak{sl}(3, \mathbb{C})$, let $z_\alpha \neq 0$ be the corresponding root vector. Let π be a representations of $\mathfrak{sl}(3, \mathbb{C})$. Let $\pi = (m_1, m_2)$ be a weight for π. Let $v \neq 0$ be the corresponding weight vector. Then*

$$\begin{aligned} \pi(H_1)\pi(z_\alpha)v &= (m_1 + \alpha_1)\pi(z_\alpha)v \\ \pi(H_2)\pi(z_\alpha)v &= (m_2 + \alpha_2)\pi(z_\alpha)v. \end{aligned} \tag{$**$}$$

Hence, $\pi(z_\alpha)v$ is either zero or is a weight vector for π with weight $\mu + \alpha = (m_1 + a_1, m_2 + a_2)$.

Two-line proof. $[H_k, z_\alpha] = a_k z_\alpha$, so $\pi(H_k)\pi(z_\alpha)v = \pi(z_\alpha)\pi(H_k)v + [\pi(H_k), \pi(z_\alpha)]v$
$= m_k \pi(z_\alpha)v + a_k \pi(z_\alpha)v$. □

Recall that for $\mathfrak{sl}(2, \mathbb{C})$, $\pi(X)$ increases the eigenvalue of $\pi(H)$ by 2 (or gives zero). By finite-dimensionality, there exist only finitely many eigenvalues of $\pi(H)$, so there exists a non-zero $v \neq 0$ with $\pi(X)v = 0$. This v had the "highest eigenvalue" for $\pi(H)$. We want the $\mathfrak{sl}(3, \mathbb{C})$ analogue of the "highest" eigenvalue. But it's not the obvious thing.

Now let's go back to the case of $\mathfrak{sl}(3, \mathbb{C})$. If $(2, -1) = \alpha_1$ and $(-1, 2) = \alpha_2$, then $(1, 1) = \alpha_1 + \alpha_2, (-2, 1) = -\alpha_1, (1, -2) = -\alpha_2, (-1, -1) = -\alpha_1 - \alpha_2$. Then $(2, -1), (-1, 2), (1, 1)$ are called the *positive simple roots*. Note that all roots are linear combinations of α_1 and α_2 with *integer* coefficients that are all ≥ 0 or ≤ 0.

Definition 16.7. Let μ_1, μ_2 to be two weights of π, a representation of $\mathfrak{sl}(3,\mathbb{C})$. We say that μ_1 is *higher than* μ_2 ($\mu_1 \succeq \mu_2$), or equivalently μ_2 is *lower than* μ_1 ($\mu_2 \preceq \mu_1$) if $\mu_1 - \mu_2 = s\alpha_1 + t\alpha_2$ with $s, t \geq 0$.

Remark 16.2. Notice that $\preceq$ is a *partial order*. If $\mu_1 \succeq \mu_2$ and $\mu_2 \succeq \mu_3$ then $\mu_1 \succeq \mu_3$, since $\mu_1 - \mu_3 = (\mu_1 + \mu_2) + (\mu_2 - \mu_3)$. If $\mu_1 \succeq \mu_2$ and $\mu_2 \succeq \mu_1$ then $\mu_1 - \mu_2 = s\alpha_1 + t\alpha_2$ and $\mu_2 - \mu_1 = s'\alpha_1 + t'\alpha_2$ with $s, t, s', t' \geq 0$. So $s\alpha_1 + t\alpha_2 = -s'\alpha_1 - t'\alpha_2$, whence we have $s, t = 0$.

However, given two weights μ_1 and μ_2, neither one need to be higher than the other. For instance, if $\mu = \alpha_1 - \alpha_2 = (3, -3)$ then μ is neither higher nor lower than, say, $(0,0)$. Also note that the coefficients s, t *need not be* integers, even though μ_1 and μ_2 may have integer entries. For example, clearly $(1,0) \succeq (0,0)$ because $(1,0) = (1,0) - (0,0) = \frac{2}{3}(2,-1) + \frac{1}{3}(-1,2)$.

Definition 16.8. Let π be a representation of $\mathfrak{sl}(3,\mathbb{C})$. Then a weight μ_0 for π is called the *highest weight* if, for any weight μ of π_1 we have $\mu_0 \succeq \mu$. Clearly, if a highest weight exists, then it is unique.

Theorem 16.9 (The highest weight for $\mathfrak{sl}(3,\mathbb{C})$)**.** *The following are true for* $\mathfrak{sl}(3,\mathbb{C})$*:*

(1) every irreducible complex representation π of $\mathfrak{sl}(3,\mathbb{C})$ is the direct sum of its weight spaces (i.e., $\pi(H_1)$ and $\pi(H_2)$ are simultaneously diagonalizable in every irreducible representation).
(2) Every irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ has a unique highest weight μ_0 and two isomorphic irreducible representations have the same highest weight.
(3) Two irreducible representations of $\mathfrak{sl}(3,\mathbb{C})$ with the same highest weight are isomorphic.
(4) If π is an irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ then the highest weight is $\mu_0 = (m_1, m_2)$ where $m_1, m_2 \in \mathbb{Z}_{\geq 0}$.
(5) If $\mu_0 = (m_1, m_2)$ is an ordered pair of non-negative integers, then there exists an irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ with the highest weight μ_0.

Remark 16.3 (Quick summary of what's going on in the above theorem). Any finite-dimensional irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ are parametrized up to isomorphism by pairs (m_1, m_2) of non-negative integers. We *do not* say that there exists one in any given dimension. There may exist non or more than one in a given dimension.

We first need the following definitions:

Definition 16.10. An ordered pair (m_1, m_2) of non-negative integers is called a *dominant integral element.*

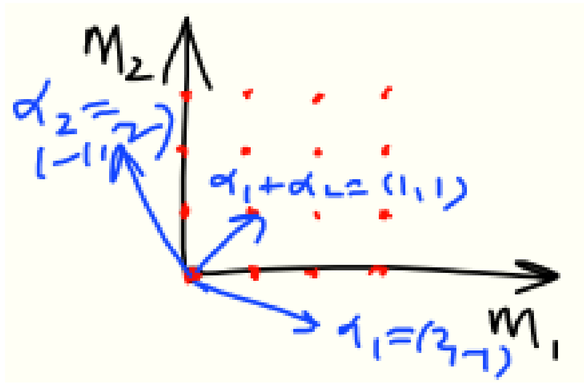
Definition 16.11. A representation (V, π) of $\mathfrak{sl}(3,\mathbb{C})$ is called a *highest weight cyclic representation with height* $\mu_0 = (m_1, m_2)$ if there exists a non-zero $v \in V$ such that:

(1) v is a weight vector with weight μ_0.
(2) $\pi(X_1)v = 0$ and $\pi(X_2)v = 0$ (which imply $\pi(X_3)v = 0$ also)
(3) the smallest invariant subspace of V contains v is all of V.

This vector v is called a *cyclic vector* for π.

So the theorem of the highest weight says that every irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ has a unique highest weight which is a dominant integral element; and that conversely, every dominant integral element arises as the highest weight of some irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$.

Note that $(1,0) = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$ and $(0,1) = \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2$ where $\alpha_1 = (2,-1)$ and $\alpha_2 = (-1,2)$. So every dominant integral element is higher than $(0,0)$. However, not every $(a,b) \in \mathbb{Z}\times\mathbb{Z}$ that is higher than $(0,0)$ is dominant integral. For example, α_1 is higher than $(0,0)$.



(Red dots = dominant integral elements)

Proof. (Proof of (1)) In every irreducible representation (V,π) of $\mathfrak{sl}(3,\mathbb{C})$ we have that $\pi(H_1)$ and $\pi(H_2)$ can be simultaneously diagonalizable (i.e., V is the direct sum of its weight spaces). Let W be the direct sum of weight spaces of π (equivalently, linear combination of simultaneous eigenvalues of $\pi(H_1), \pi(H_2)$). By the main lemma if z_α is a root vector of $\mathfrak{sl}(3,\mathbb{C})$ then $\pi(z_\alpha)(W) \subseteq W$ and $\pi(z_\alpha)(W_\mu) \subseteq W_{\mu+\alpha}$. Also, $\pi(H_k) : W \subseteq W$. So W is invariant under π and $W \neq \{0\}$. SInce V is irreducible we have $W = V$, as desired.

(Idea of the proof of (2) and (3)) We will show that being a finite-dimensional complex irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ with the highest weight is equivalent to being a highest weight cyclic representation. We will show that if (V,π) is a HWC with weight μ_0 then μ_0 is a highest weight. And then we shall show that HWC with weight μ_0 is equivalent to "irreducible with highest weight μ_0. After this, we shall get to (2) and (3) in the process on Wednesday. ∎

17. March 4

Before proving (2) and (3) of Theorem 16.9, we need to prove the following lemma:

Lemma 17.1. *Let $\mathfrak{g}$ be any Lie algebra and let $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite-dimensional representation of $\mathfrak{g}$. Let $z_1, \ldots, z_m$ be an ordered basis of $\mathfrak{g}$. Then any expression of the form*

$$\pi(z_{i_1}) \cdots \pi(z_{i_n})$$

can be expressed as a linear combination of terms of the form

$$\pi(z_m)^{k_m} \cdots \pi(z_2)^{k_2}\pi(z_1)^{k_1}$$

with $k_i \geq 0$ and $k_1 + \cdots + k_m \leq N$.

Proof. We prove by induction on N. For $N = 1$, the claim is immediate. Now assume that for N,

$$\pi(z_i)\pi(z_j) = \pi(z_j)\pi(z_i) + \sum_{k=1}^{m} c_{ij}^k \pi(z_k)$$

where $[z_i, z_j] = \sum_{k=1}^{m} c_{ij}^k z_k$ if we have

$$\pi(z_{i_1})\cdots\pi(z_{i_N}) = \pi(z_{i_1})[\text{linear combination of terms } \pi(z_m)^{k_m}\cdots\pi(z_1)^{k_1}],$$

with $k_i \geq 0$ and $k_1 + \cdots + k_m \leq N$. Now jump over to put i_1 in the right spot, and each time introduce new terms where each of them is a product of $\leq N$ factors. Apply induction hypothesis on each of those terms. □

Definition 17.2. Recall that a representation (V, π) of $\mathfrak{sl}(3, \mathbb{C})$ is called *highest weight cyclic (HWC) with weight* μ_0 if there exists a non-zero $v \in V$ such that

(1) v is a weight vector with weight μ_0
(2) $\pi(X_1)v = 0$ and $\pi(X_2)v = 0$ implies that $\pi(X_3)v = 0$ where $X_3 = [X_1, X_2]$
(3) the smallest invariant subspace of V that contains V is all of V.

And the vector v is called a *cyclic vector* for π.

Proposition 17.3. *Let (V, π) be HWC with height μ_0. Then*

(1) π has highest weight μ_0
(2) the height space W_{μ_0} corresponds to μ_0 is one-dimensional.

Proof. Let W be the subspace of V spanned by the elements of the form

$$\pi(Y_{i_1})\pi(Y_{i_2})\cdots\pi(Y_{i_m})v,$$

where each i_k is 1, 2, or 3 and $N \geq 0$. We will show that W is an invariant subspace.

Suppose that $X_1, X_2, X_3, H_1, H_2, Y_1, Y_2, Y_3$ is our ordered basis of $\mathfrak{sl}(3, \mathbb{C})$. Apply the lemma that we just proved to see that for any z n the basis, $\pi(z)\pi(Y_{i_1})\cdots\pi(Y_{i_N})$ is a linear combination of the terms $\pi(Y_3)^{k_8}\pi(Y_2)^{k_7}\pi(Y_1)^{k_6}\pi(H_2)^{k_5}\pi(H_1)^{k_4}\pi(X_3)^{k_3}\pi(X_2)^{k_2}\pi(X_1)^{k_1}$. But since $\pi(X_k)v = 0$ for $k = 1, 2, 3$, we can assume that $k_1 = k_2 = k_3 = 0$. Powers of $\pi(H_1)$ and $\pi(H_2)$ only introduce constants because v is a simultaneous eigenvector of $\pi(H_1)$ and $\pi(H_2)$. We are left with an element of W. Since π is linear, we get $\pi(Z)(W) \subseteq W$ for all $Z \in \mathfrak{sl}(3, \mathbb{C})$. So W is an invariant subspace containing v.

Hence by (3) of the definition of HWC, we have $W = V$. We know that Y_1, Y_2, Y_3 are root vectors of $\mathfrak{sl}(3, \mathbb{C})$ with roots $-\alpha_1, -\alpha_2, -\alpha_1 - \alpha_2$ respectively. Hence, every element of V is a linear combination of weight vectors with weight $\mu_0 - n_1\alpha_1 - n_2\alpha_2$ for n_1, n_2 non-negative integers.

$$\underbrace{\pi(Y_3)^{k_3}\,\underbrace{\pi(Y_2)^{k_2}\,\underbrace{\pi(Y_1)k_1v}_{\substack{\text{weight}\\ \mu_0-k_1\alpha_1}}}_{\mu_0-k_1\alpha_1-k_2\alpha_2}}_{\mu_0-k_1\alpha_1-k_2\alpha_2-k_3(\alpha_1+\alpha_2)}$$

Note that μ_0 is higher than $\mu_0 - n_1\alpha_1 - n_2\alpha_2$ since $\mu_0 - (\mu_0 - n_1\alpha_1 - n_2\alpha_2) = n_1\alpha_1 + n_2\alpha_2$ and $n_1, n_2 \geq 0$. Thus if (V, π) is a HWC representation with weight μ_0 then μ_0 is a highest weight for π.

Let $w \in V$ be a weight vector with weight μ_0. If $v, v_1, \ldots, v_r$ are weight vectors of corresponding distinct heights, then there exists a constant c such that $w = cv + v_1 + \cdots + v_r$. Since $cv - w$ has weight μ_0, all the vectors have distinct weights, so they are linearly independent. Thus $cv - w = 0$ so $w = cv$. Thus W_{μ_0} is one-dimensional as we wanted. □

The following proposition will prove (2) of Theorem 16.9.

Proposition 17.4. *Every irreducible representation of* $\mathfrak{sl}(3,\mathbb{C})$ *is a HWC representation with a unique highest weight* μ_0.

Proof. We already know by (1) of Theorem 16.9 that every irreducible representation is a direct sum of its weight spaces. Since our representations are finite-dimensional, there can be at most finitely many distinct weights. Hence there exists a weight μ_0 such that there is no weight $\mu \neq \mu_0$ that is higher than μ_0. Note that this is *not equivalent* to saying that μ_0 is a highest weight.

Since there is no weight higher than μ_0, if $v \neq 0$ is a weight vector with weight μ_0 then $\pi(X_1)v = 0$ and $\pi(X_2)v = 0$. (Otherwise, there would be weight vectors with weights $\mu_0 + \alpha_1$ and $\mu_0 + \alpha_2$ respectively, both of which are distinct from μ_0 and higher than μ_0.) Since π is irreducible, the smallest invariant subspace of V containing v is all of V. Thus the all the definitions of HWC are satisfied. □

Corollary 17.5. *Every irreducible representation of* $\mathfrak{sl}(3,\mathbb{C})$ *has a* unique *highest weight.*

Remark 17.1. This is (2) of Theorem 16.9.

Proposition 17.6. *Every HWC representation of* $\mathfrak{sl}(3,\mathbb{C})$ *is irreducible.*

Proof. Let (V,π) be a HWC representation of $\mathfrak{sl}(3,\mathbb{C})$ with weight μ_0. Here is where we use the fact that $\mathfrak{sl}(3,\mathbb{C})$ has CRP. So (V,π) is isomorphic to

$$V \cong \bigoplus_{i=1}^{N} V_i,$$

where each (V_i,π_i) is irreducible, and $\pi \cong \pi_1 \oplus \cdots \oplus \pi_N$. We have already shown in (1) of Theorem 16.9 that each V_i is a direct sum of its weight spaces. Since the weight μ_0 occurs in V, it must occur in some V_i. If not, then we can get a vector $v \in V_1 \oplus V_2$ such that v is a weight vector with weight μ_0 and $v = v_1 + \cdots + v_r$ such that weight vectors of either V_1 or V_2 with weight distinct from μ_0. This implies that $v = 0$, a contradiction.

So at least one V_i has weight μ_0. But we proved that fro every HWC representation with weight μ_0 the height space W_{μ_0} is one-dimensional. So V_i must contain v. Also, V_i is an invariant subspace of V containing v. By (3) of the definition of HWC, we have $V_1 = V$. Therefore $V = V_i$ is indeed irreducible. □

So we just proved that the irreducible representations of $\mathfrak{sl}(3,\mathbb{C})$ is the same as the HWC representation of $\mathfrak{sl}(3,\mathbb{C})$.

Proposition 17.7. *Two irreducible representations of* $\mathfrak{sl}(3,\mathbb{C})$ *with the same highest weight are isomorphic.*

Remark 17.2. This is (3) of the main theorem.

Proof. We have shown that irreducibility and HWC are equivalent. Let (V,π) and (W,σ) be two such representations with the *same* highest weight $\mu_0 = (m_1, m_2)$. Let $v \neq 0$ in V and $w \neq 0$ in W be the corresponding cyclic vectors. Consider the representation $(V \oplus W, \pi \oplus \sigma)$, and let U be the smallest invariant subspace of $V \oplus W$ containing (v,w). Then

$$(\pi \oplus \sigma)(X_k)(v,w) = (\pi(X_k)v, \sigma(X_k)w) = (0,0),$$

since v, w are cyclic vectors for π and σ respectively. Similarly,

$$(\pi \oplus \sigma)(H_k)(v, w) = (\pi(H_k)v, \sigma(H_k)w) = (m_k v, m_k w) = m_k(v, w),$$

with $k = 1, 2$. So $(v, w) \neq (0, 0)$ in $U \subseteq V \oplus W$, and $(\pi \oplus \sigma)(X_k)(v, w) = (0, 0)$ for $k = 1, 2$ and (v, w) is a weight vector for $\pi \oplus \sigma$ with weight μ_0. And U is the smallest invariant subspace containing (v, w). Hence $(U, (\pi \oplus \sigma)|_U)$ is a HWC representation of $\mathfrak{sl}(3, \mathbb{C})$ with height μ_0. Hence $(U, (\pi \oplus \sigma)|_U)$ is irreducible.

Now consider the canonical projection maps $P_1 : V \oplus W \to V$ and $P_2 : V \oplus W \to W$. It's easy to verify that P_1, P_2 are intertwining maps, so $P_1|_U$ and $P_2|_U$ are intertwining maps also. So it follows that $P_1|_U : U \to V$ and $P_2|_U : U \to W$ are intertwining maps between the irreducible representations. By Schur's lemma, each is either zero or an isomorphism. But then neither of them are zero because $P_1|_U(v, w) = v \neq 0$ and $P_2|_U(v, w) = w \neq 0$. Thus they are both isomorphisms. Hence $V \cong U$ as representations of $\mathfrak{sl}(3, \mathbb{C})$. Thus $V \cong W$ as desired. □

Proof of Theorem 16.9(4). Let π be an irreducible representation of $\mathfrak{sl}(3, \mathbb{C})$. THen the highest weight μ_0 of π is of the form $\mu_0 = (m_1, m_2)$ for $m_1, m_2 \in \mathbb{Z}$. We need to show that m_1 and m_2 are non-negative integers. Let $v \neq 0$ be a cyclic vector for π with $\pi(X_1)v = \pi(X_2)v = 0$. Hence if we restrict π to a representation of the subalgebra

$$\mathfrak{g}_k = \operatorname{span}\{H_k, X_k, Y_k\} \cong \mathfrak{sl}(2, \mathbb{C})$$

for $k = 1, 2$, then we get that $\pi(H_k)v = m_k v$ and $\pi(X_k)v = 0$ by our results for $\mathfrak{sl}(2, \mathbb{C})$ with $m_k \geq 0$. This completes the proof. □

Now all that remains is (5) of Theorem 16.9: given $\mu_0 = (m_1, m_2)$ with $m_1, M - 2$ non-negative integers, we need to show that there exists an irreducible representation of $\mathfrak{sl}(3, \mathbb{C})$ with highest weight μ_0.

Proof of Theorem 16.9(5). First, construct in the cases $\mu_0 = (0, 0), (1, 0), (0, 1)$. If $\mu_0 = (0, 0)$ then $V = \mathbb{C}$ and define $\pi_{(0,0)} : \mathfrak{sl}(3, \mathbb{C}) \to \mathfrak{gl}(1, \mathbb{C})$ by $\pi_{(0,0)}(z)v = 0$ for all $z \in \mathfrak{sl}(3, \mathbb{C})$ and $v \in \mathbb{C}$. This is the trivial representation, which is irreducible (evidently). Since $\pi(H_2)v = 0$ for all v, this representation has only one weight $(0, 0)$.

Now suppose $\mu_0 = (1, 0)$. Let $V = \mathbb{C}^3$. Consider the standard representation π. Then $\pi(z)v = zv$ (the regular matrix multiplication). Recall that you showed in Assignment #4 that this is irreducible. Let e_1, e_2, e_3 be the standard basis of $\mathbb{C}^3$. Then e_1 has weight $(1, 0)$; e_2 has weight $(-1, 1)$; and e_3 has weight $(0, -1)$. We claim that $(1, 0)$ is the highest weight since $(1, 0) - (-1, 1) = (2, -1) = \alpha_1$ and $(1, 0) - (0, -1) = (1, 1) = \alpha_1 + \alpha_2$. So the irreducible representation with highest weight $(1, 0)$ is the standard representation.

In the case of $(0, 1)$, take the same vector space $V = \mathbb{C}^3$. Let $\pi^* : \mathfrak{sl}(3, \mathbb{C}) \to \mathfrak{gl}(3, \mathbb{C})$ be the dual of the standard representation. Recall that $\pi^*(z) = -(\pi(z))^t$. Then we have $\pi^*(z)v = -z^t v$ (the usual matrix multiplication). By Assignment #4, we see that since π is irreducible, so is π^*. Since $\pi^*(H_k)v = -H_k^t v = -H_k v$, we see that e_1 is a weight vector of π^* with height $(-1, 0)$; e_2 has weight (for π^*) $(1, -1)$; and e_3 has weight $(0, 1)$. One can easily check that $(0, 1)$ is the highest of the three. Hence the irreducible representation of $\mathfrak{sl}(3, \mathbb{C})$ with the highest weight $(0, 1)$ is the dual of the standard representation.

We ran out of time, so we will continue this proof next Monday. We will start on constructing an irreducible representation of $\mathfrak{sl}(3, \mathbb{C})$ of highest weight (m_1, m_2) using the "fundamental representation" (i.e., the irreducible representations of $\mathfrak{sl}(3, \mathbb{C})$ of highest weight

$(1,0)$ and $(0,1)$). We will do $(1,1)$ in class and $(2,0)$ in Assignment #5. We will also start on Weyl group. □

Remark 17.3. The three cases for $(0,0),(1,0),(0,1)$ prove that the standard representation of $\mathfrak{sl}(3,\mathbb{C})$ *is not isomorphic* to its dual. On the other hand, for $\mathfrak{sl}(2,\mathbb{C})$ *is indeed isomorphic* to its dual: note that they have the same dimension, and for $\mathfrak{sl}(2,\mathbb{C})$ there exists exactly one in each dimension.

Remark 17.4. For $\mathfrak{sl}(3,\mathbb{C})$, it is not the case that there exists at least one irreducible representation in every dimension (unlike $\mathfrak{sl}(2,\mathbb{C})$). In particular, there cannot exist an irreducible representation of dimension 2 (there exists a formula for the dimension in terms of $\mu_0 = (m_1, m_2)$).

18. March 9

Proposition 18.1. *Let $\mu = (m_1, m_2)$ be a dominant integral element (i.e., m_i are non-negative integers). Then there exists an irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ with highest weight (m_1, m_2).*

Remark 18.1. Recall that:

highest weight	representation
$(0,0)$	trivial representation
$(1,0)$	standard representation ($\pi(Z)v = Zv$)
$(0,1)$	dual of the standard representation ($\pi^*(Z)v = \overline{Z}v, \overline{Z} = -Z^t$)
$(1,1)$	adjoint representation (we will do this in class today)

Proof. Let (V_1, π_1) be the $(1,0)$ representation with cyclic vector v. Then $\pi_1(X_k)v = 0, \pi_1(H_1)v = v, \pi_1(H_2)v = 0$ where $k = 1,2$. Let (V_2, π_2) be the $(0,1)$ representation with cyclic vector w. Then $\pi_2(X_k)2 = 0, \pi_2(H_1)w = 0, \pi_2(H_2)w = w$ where again $k = 1,2$.

Define $V = V_1^{\otimes m_1} \otimes V_2^{\otimes m_2} \cong \mathbb{C}^{3(m_1+m_2)}$, and let $\pi_{m_1,m_2} = \pi_1^{\otimes m_1} \otimes \pi_2^{\otimes m_2}$ be the tensor product representation of $\mathfrak{sl}(3,\mathbb{C})$ on V. This is a finite-dimensional representation of $\mathfrak{sl}(3,\mathbb{C})$. The action of π_{m_1,m_2} is therefore (where $Z \in \mathfrak{sl}(3,\mathbb{C})$):

$$\pi_{m_1,m_2}(Z) = \pi(Z) \otimes I \otimes \cdots \otimes I + \cdots + I \otimes I \otimes \cdots \otimes \pi_2(Z).$$

Consider $v_{m_1,m_2} := v^{\otimes m_1} \otimes w^{\otimes m_2}$. This is non-zero, and we have

$$\begin{aligned}
\pi_{m_1,m_2}(H_1)v_{m_1,m_2} &= m_1 v_{m_1,m_2} \\
\pi_{m_1,m_2}(H_2)v_{m_1,m_2} &= m_2 v_{m_1,m_2} \\
\pi_{m_1,m_2}(X_k)v_{m_1,m_2} &= 0.
\end{aligned}$$

So v_{m_1,m_2} is a weight vector for π_{m_1,m_2} with height (m_1, m_2). Thus by Theorem 16.9, if we let U be the smallest invariant subspace of V containing v_{m_1,m_2}, then $\pi_{m_1,m_2}|_U : \mathfrak{sl}(3,\mathbb{C}) \to \mathfrak{gl}(U)$ is an irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ with the highest weight (m_1, m_2) (since it is a HWC rep'n with weight (m_1, m_2). □

Example 18.2 ("Good for your soul", according to Spiro). Recall that the basis of $\mathfrak{sl}(3,\mathbb{C})$ consists of

$$X_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad X_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$Y_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Y_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad Y_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Then the $(1,0)$ representation of $\mathfrak{sl}(3,\mathbb{C})$ satisfies

$$\begin{aligned} Y_1(e_1) &= 0, \quad Y_2(e_1) = 0 \\ Y_1(e_2) &= 0, \quad Y_2(e_2) = e_3 \\ Y_1(e_3) &= 0, \quad Y_2(e_3) = 0, \end{aligned}$$

and the highest weight vector is e_1. As for the $(0,1)$ representation, we have

$$\overline{Y_1} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \overline{Y_2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\overline{H_1} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \overline{H_2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In this case, e_3 is a highest weight vector. Switch the notation for the dual representation by letting $f_1 := e_3, f_2 := -e_2, f_3 := e_1$. Then we have

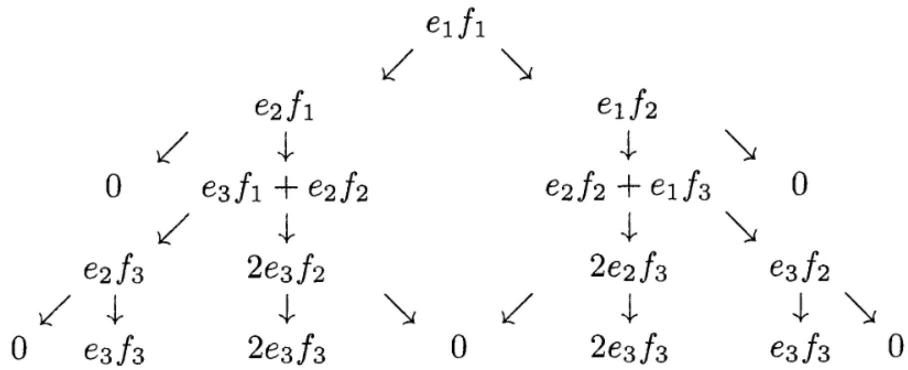
$$\begin{aligned} \overline{Y_1}(f_1) &= 0, \quad \overline{Y_2}(f_1) = f_2 \\ \overline{Y_1}(f_2) &= f_3, \quad \overline{Y_2}(f_2) = 0 \\ \overline{Y_1}(f_3) &= 0, \quad \overline{Y_2}(f_3) = 0. \end{aligned}$$

Thus the $(1,1)$ irreducible representation of $\mathfrak{sl}(3,\mathbb{C})$ is obtained by taking the smallest invariant subspace of $\mathbb{C}^3 \otimes \mathbb{C}^3 \cong V_1 \otimes V_2$ containing $e_1 \otimes f_1$. The subspace is determined by applying powers of $\pi(Y_1), \pi(Y_2)$ to $e_1 \otimes f_1$ where $\pi = \pi_{(1,0)} \otimes \pi_{(0,1)}$, i.e.,

$$\begin{aligned} \pi(Y_1) &= Y_1 \otimes I + I \otimes \overline{Y_1} \\ \pi(Y_2) &= Y_2 \otimes I + I \otimes \overline{Y_2}. \end{aligned}$$

Also, recall the following fact: if v is a weight vector with weight μ and z_α is a root vector with root α, then either $\pi(z_\alpha)v = 0$ or $\pi(z_\alpha)v$ is a weight vector with weight $\mu + \alpha$; recall also that Y_1, Y_2 are the root vectors with roots $(-2,1)$ and $(1,-2)$ respectively.

Now let's apply $\pi(Y_1)$ and $\pi(Y_2)$ repeatedly, starting from $e_1 \otimes f_1$ till we get the zero vector. We get the following diagram:



In the above diagram, the tensor signs were omitted so as to save spaces. Left arrows denote applying $\pi(Y_1)$ and the right arrows $\pi(Y_2)$. Then we get

weights	weight vectors
$(1,1)$	$e_1 \otimes f_1$
$(-1,2)$	$e_2 \otimes f_1$
$(2,-1)$	$e_1 \otimes f_2$
$(0,0)$	$e_3 \otimes f_1 + e_2 \otimes f_2$ and $e_2 \otimes f_2 + e_1 \otimes f_3$
$(-2,1)$	$e_2 \otimes f_3$
$(1,-2)$	$e_3 \otimes f_2$
$(-1,-1)$	$e_3 \otimes f_3$.

So the dimension of the $(1,1)$ representation is 8, with 7 distinct weights and 6 of them have one-dimensional weight spaces, while one of them has a two-dimensional height space. But then $V_1 \otimes V_2 \cong \mathbb{C}^3 \otimes \mathbb{C}^3 \cong \mathbb{C}^9$, so $V_1 \otimes V_2$ is not irreducible. Also, since $\mathfrak{sl}(3,\mathbb{C})$ has the CRP, it follows that $V_1 \otimes V_2 = U \oplus V'$ where V' is some complementary invariant subspace. Since $\dim(U) = 8$, it follows that $\dim(V') = 1$. So $(1,0) \otimes (0,1) = (1,1) \oplus (0,0)$.

Remark 18.2. In Assignment #5, you will figure out M such that $(1,0) \otimes (1,0) = (2,0) \oplus M$.

Remark 18.3. The $(1,1)$ representation is isomorphic to the adjoint representation. Recall that the adjoint representation is defined as follows. If $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ with $V = \mathfrak{g} = \mathfrak{sl}(3,\mathbb{C})$ defined as $\mathrm{ad}_X(Y) = [X,Y]$ is is a representation of degree eight (i.e., $\dim(\mathfrak{sl}(3,\mathbb{C})) = 8$). One can check from the commutation relations of $\mathfrak{sl}(3,\mathbb{C})$ that this is true. One can find an isomorphism of representations between $(1,1)$ representation we constructed and the adjoint representation.

18.1. *Weyl group of* $\mathfrak{sl}(3,\mathbb{C})$ *(Weyl group of* $\mathrm{SU}(3)$*)*

Let $\pi : \mathfrak{sl}(3,\mathbb{C}) \to \mathfrak{gl}(V)$ be a finite-dimensional finite-dimensional representation of $\mathfrak{sl}(3,\mathbb{C})$. We know that, since $\mathrm{SU}(3)$ is simply connected, that $\pi = \Pi_*$ for some finite-dimensional complex representation $\Pi : \mathrm{SU}(3) \to \mathrm{GL}(V)$ of $\mathrm{SU}(3)$. Let $A \in \mathrm{SU}(3)$. WE can then construct a new representation $\pi_A : \mathfrak{sl}(3,\mathbb{C}) \to \mathfrak{gl}(V)$ acting on the *same* space V as follows:

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

Recall that $\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra automorphism since $\mathrm{Ad}_A[X,Y] = [\mathrm{Ad}_A X, \mathrm{Ad}_A Y]$. So π_A is indeed a representation of $\mathfrak{sl}(3,\mathbb{C})$. Note also that

$$\begin{aligned}\pi_A(X) &= \pi(AXA^{-1}) = \pi(\mathrm{Ad}_A X)\\ &= \Pi_*(\mathrm{Ad}_A X) = \mathrm{Ad}_{\Pi(A)}(\pi(X)) = \Pi(A)\pi(X)\Pi(A)^{-1}.\end{aligned}$$

Let $\mathrm{GL}(V) \ni T = \Pi(A) : V \cong V$. Then $\Pi(A)\pi(X) = \pi_A(X)\Pi(A)$ for all $X \in \mathfrak{sl}(3,\mathbb{C})$. Hence T is an isomorphism of representations, so π_A is isomorphic to π for all $A \in \mathrm{SU}(3)$. Given any $A \in \mathrm{SU}(3)$, we get an action on each isomorphism class of finite-dimensional complex representations of $\mathfrak{sl}(3,\mathbb{C})$.

Define $\mathfrak{h} := \mathrm{span}\{H_1, H_2\} \subseteq \mathfrak{sl}(3,\mathbb{C})$. This is a two-dimensional Lie subalgebra of $\mathfrak{sl}(3,\mathbb{C})$ on which the Lie bracket vanishes identically. In fact, $\mathfrak{h}$ is a maximal commutative subalgebra.

Definition 18.3. Such $\mathfrak{h}$ as defined above is called a *Cartan subalgebra* for $\mathfrak{sl}(3,\mathbb{C})$.

For any $A \in \mathrm{SU}(3)$, the map $\mathrm{Ad}_A : \mathfrak{sl}(3,\mathbb{C}) \to \mathfrak{sl}(3,\mathbb{C})$ is a Lie algebra automorphism. But there is no reason for Ad_A to preserve $\mathfrak{h}$. Now we are ready to define what a Weyl group is:

Definition 18.4. Let $Z \subseteq \mathrm{SU}(3)$ be the subgroup of $\mathrm{SU}(3)$ whose elements satisfy $\mathrm{Ad}_A(H) = H$ for all $H \in \mathfrak{h}$ (i.e., elements that fix $\mathfrak{h}$ pointwise).

Now let $N \subseteq \mathrm{SU}(3)$ be the subgroup of elements such that $\mathrm{Ad}_A(H) \in \mathfrak{h}$ for all $H \in \mathfrak{h}$ (i.e., the elements that preserve $\mathfrak{h}$).

Remark 18.4. Note that Z is a subgroup of $\mathrm{SU}(3)$ since $\mathrm{Ad}_{AB} = \mathrm{Ad}_A \circ \mathrm{Ad}_B$ implies that $A, B \in Z \Rightarrow AB \in \mathbb{Z}$, and $\mathrm{Ad}_{A^{-1}} = (\mathrm{Ad}_A)^{-1}$.

Remark 18.5. Note that $Z \triangleleft N$. We need to verify if $A \in Z$ and $B \in N$ then $BAB^{-1} \in Z$. Let $H \in \mathfrak{h}$. Then

$$\begin{aligned}\mathrm{Ad}_{BAB^{-1}} H &= \mathrm{Ad}_B \mathrm{Ad}_A (\mathrm{Ad}_B)^{-1}(H)\\ &= \mathrm{Ad}_B (\mathrm{Ad}_B)^{-1}(H) \quad (\mathrm{Ad}_A \text{ fixes } \mathfrak{h})\\ &= H.\end{aligned}$$

So $BAB^{-1} \in Z$.

Therefore we can take the quotient N/Z.

Definition 18.5. The *Weyl group of* $\mathfrak{sl}(3,\mathbb{C})$ (or equivalently the *Weyl group of* $\mathrm{SU}(3)$) is $W := N/Z$.

We have an action of W on $\mathfrak{h}$. Namely, if $[A] \in W$, define $[A]H = \mathrm{Ad}_A H$ where A is any representation of the equivalence class $[A]$, where the equivalence is defined as follows: if $[A] = [B]$,then there exists $C \in Z$ such that $A = BC$.

Theorem 18.6. *Let N and Z be the subgroups as defined in Definition 18.4.*

(1) Z consists of diagonal matrices in $\mathrm{SU}(3)$. *That is,*

$$Z = \left\{ \begin{bmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{i\varphi} & 0 \\ 0 & 0 & e^{-i(\theta+\varphi)} \end{bmatrix} \right\}$$

(2) N consists of the matrices $A \in \mathrm{SU}(3)$ such that for each $k = 1, 2, 3$ there exists $l \in \{1, 2, 3\}$ such that $Ae_k = e^{i\theta_k} e_k$ for some $e^{i\theta_k}$ (those that permute the standard basis vectors $\{e_1, e_2, e_3\}$ of $\mathbb{C}^2$ up to a complex phase)
(3) $W = N/Z \cong S_3$

Proof of Theorem 18.6(1). Let $A \in Z$. Then $AHA^{-1} = H$ for all $H \in \mathfrak{h}$, hence $AH = HA$. Thus $[A, H] = 0$ for all $H \in \mathfrak{h}$. Note that H_1 has three distinct eigenvalues with eigenvectors e_1, e_2, e_3, so A preserves the eigenspaces of H_1. If $Hv = \lambda v$, then $\lambda Av = HAv$ so A preserves the eigenspaces of H_1. So $Ae_1 = t_1 e_1, Ae_2 = t_2 e_2, Ae_3 = t_3 e_3$. Since $A \in \mathrm{SU}(3)$, we have $t_1 = e^{i\theta}, t_2 = e^{i\varphi}, t_3 = e^{-i(\varphi+\theta)}$. Conversely, every diagonal matrix in $\mathrm{SU}(3)$ commutes with every $H \in \mathfrak{h}$ since the H's are diagonal. Thus Z consists of diagonal matrices in $\mathrm{SU}(3)$. □

19. March 11

Proof of Theorem 18.6(2). Let $A \in N$. Then $AH_1A^{-1} \in \mathfrak{h}$, so it must be diagonal. Also, $H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ has three distinct eigenvalues $1, -1, 0$ with eigenvectors e_1, e_2, e_3. So AH_1A^{-1} is also diagonal, hence has e_1, e_2, e_3 as eigenvectors. So we have $(AH_1A^{-1})(Ae_k) = AH_1 e_k = \lambda_k Ae_k$. Therefore eigenvectors of AH_1A^{-1} are Ae_1, Ae_2, Ae_3 with distinct eigenvalues. Since $A \in \mathrm{SU}(3)$, we have $|Ae_k| = |e_k| = 1$. Hence, for all $k = 1, 2, 3$ there exists $l \in \{1, 2, 3\}, t_k \in \mathbb{C}, |t_k| = 1$ such that $Ae_k = t_k e_l$.

Conversely, suppose that $A \in \mathrm{SU}(3)$ so that for all k we have $Ae_k = t_l e_k$ for some l with $|t_k| = 1$, with eigenvectros of AH_1A^{-1} still e_1, e_2, e_3 (but in possibly a different order). We have $\mathrm{tr}(H_1) = 0$, so $\mathrm{tr}(AH_1A^{-1}) = 0$. Thus $AH_1A^{-1} \in \mathfrak{h}$. Similarly, $AH_2A^{-1} \in \mathfrak{h}$ so $A \in N$. □

Proof of Theorem 18.6(3). Let $A \in N$. consider $\mathrm{Ad}_A : \mathfrak{h} \cong \mathfrak{h}$. Let $\sigma \in S_3$ such that $Ae_k = t_k e_{\sigma(k)}$. Let $H \in \mathfrak{h}$ such that $He_k = \lambda_k e_k$. So it follows

$$AHA^{-1} e_{\sigma(k)} = AHt_k^{-1} e_k = A\lambda_k t_k^{-1} e_k = \lambda_k \frac{t_k}{t_k} e_{\sigma(k)} = \lambda_k e_{\sigma(k)}$$

if e_k is an eigenvector of H with eigenvector λ_k. Then $e_{\sigma(k)}$ is an eigenvector of $\mathrm{Ad}_A H$ with eigenvector λ_k. Hence the diagonal entries of H are permuted by the permutation σ, i.e.,

$$H = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \Rightarrow \mathrm{Ad}_A H = \begin{bmatrix} \lambda_{\sigma^{-1}(1)} & 0 & 0 \\ 0 & \lambda_{\sigma^{-1}(2)} & 0 \\ 0 & 0 & \lambda_{\sigma^{-1}(3)} \end{bmatrix}.$$

λ_k is now in $\sigma(k)$-th column of $\mathrm{Ad}_A H$. Hence $W \cong S_3$. If $[A] = [B] \in W$ then $A = BC$ with $C \in Z$. Thus $\mathrm{Ad}_A = \mathrm{Ad}_B \mathrm{Ad}_C$, and Ad_C acts as I on $\mathfrak{h}$. □

Why should we care about W? Note that W permutes the weights of any finite-dimensional representation of $\mathfrak{sl}(3, \mathbb{C})$ (not necessarily irreducible). To make this precise, we need to reformulate the notion of a weight in a basis-independent manner. Recall that $(m_1, m_2) \in \mathbb{C}^2$ is a weight of a representation (V, π) of $\mathfrak{sl}(3, \mathbb{C})$ if there is non-zero $v \in V$ such that $\pi(H_1)v = m_1 v$ and $\pi(H_2)v = m_2 v$. We also showed that $m_k \in \mathbb{Z}$. Since $\pi(a_1 H_1 + a_2 H_2)v = a_1 \pi(H_1)v + a_2 \pi(H_2)v = (a_1 m_1 + a_2 m_2)v$, it follows that v is an eigenvector for $\pi(H)$ for all $H \in \mathfrak{h}$ and the eigenvalue $\pi(H) \in \mathbb{C}$ of $\pi(H)$ with eigenvector v is a linear functional on $\mathfrak{h}$.

Definition 19.1. Define $\mathfrak{h} := \mathrm{span}\{H_1, H_2\} \subseteq \mathfrak{sl}(3,\mathbb{C})$. Let (V,π) be a finite-dimensional representation of $\mathfrak{sl}(3,\mathbb{C})$. A linear functional $\mu \in \mathfrak{h}^*$ is called a *weight* for π if there exists a non-zero $v \in V$ sic that $\pi(H)v = \mu(H)v$ for all $H \in \mathfrak{h}$, and v is called a *weight vector of* π with weight $\mu \in \mathfrak{h}^*$. Also, μ is uniquely determined by $\mu(H_1) = m_1$ and $\mu(H_2) = m_2$ so these notions are equivalent.

The Weyl group W acts on $\mathfrak{h}$. Denote it by $w \in W, H \in \mathfrak{h} \Rightarrow w \cdot H \in \mathfrak{h}$. So if $w = [A] \in N/Z$ with $A \in N$ then $w \cdot H = \mathrm{Ad}_A(H)$. Thus W defines on induced action on $\mathfrak{h}^*$, by $w \in W$ and $\mu \in \mathfrak{h}^*$:

$$\underbrace{(w \cdot \mu)}_{\in \mathfrak{h}^*}(H) := \mu(w^{-1} \cdot H).$$

With this definition, we can say $(w_1 w_2) \cdot \mu = w_1 \cdot (w_2 \cdot \mu)$ so this is an action.

Theorem 19.2. *Let π be a finite-dimensional representation of $\mathfrak{sl}(3,\mathbb{C})$ and let $\mu \in \mathfrak{h}^*$ be a weight of π. Then for all $w \in W$, $w \cdot \mu$ is a weight for π and the multiplicity of $w \cdot \mu$ is equal to the multiplicity of μ, which is the dimension of this weight space.*

Proof. Let Π be the associated representation of SU(3). Note that we know such thing exists since SU(3) is simply connected. Let μ be a weight for π, with weight vector $v \neq 0$. If $A \in N$, then

$$\begin{aligned}\pi(H)\Pi(A)v &= \Pi(A)\Pi(A^{-1})\pi(H)\Pi(A)v = \Pi(A)\pi(A^{-1}HA)v \\ &= \Pi(A)\mu(A^{-1}HA)v = \mu(A^{-1}HA)\Pi(A)v.\end{aligned}$$

Also, $A^{-1}HA = \mathrm{Ad}_{A^{-1}} H - w^{-1} \cdot H$ where $w = [A] \in W$. Hence,

$$\pi(H)\Pi(A)v = \pi(A^{-1}HA)\Pi(A)v = \mu(w^{-1} \cdot H)\Pi(A)v = (w \cdot \mu)(H)\Pi(A)v.$$

Therefore $\Pi(A)$ is a weight vector for π with weight $w \cdot \mu$. Note that $\Pi(A)W_\mu \subseteq W_{w\cdot\mu}$ and $\Pi(A^{-1})W_{w\cdot\mu} \subseteq W_\mu$, and $\Pi(A)$ is invertible hence it is an isomorphism from W_μ to $W_{w\cdot\mu}$. Therefore μ and $w \cdot \mu$ have the same multiplicity. ☐

Remark 19.1. Recall that the rtoos of $\mathfrak{sl}(3,\mathbb{C})$ are the non-zero weights of the adjoint representation. If $w \in W$ then w takes non-zero weights to non-zero weights. If $\mu = 0$ then $(w \cdot \mu)(H) = \mu(w^{-1} \cdot H) = 0$ for all $H \in \mathfrak{h}$. Then $w \cdot \mu = 0$. Thus W permutes the roots of $\mathfrak{sl}(3,\mathbb{C})$.

We want a geometric picture of the symmetry of W on the set of roots of $\mathfrak{sl}(3,\mathbb{C})$. We need one more change of point of view for this though. SInce $\mathfrak{h} \subseteq \mathfrak{sl}(3,\mathbb{C}) \subseteq \mathfrak{gl}(3,\mathbb{C}) \cong \mathbb{C}^9$, we have a standard Hermitian inner product on $\mathfrak{gl}(3,\mathbb{C})$, defined by

$$\langle A, B\rangle = \mathrm{tr}(A^*B) = \sum_{i,j=1}^{3} \overline{A_{ij}} B_{ij}.$$

Recall that W permutes the diagonal entries of an element $H \in \mathfrak{h}$. Hence if $H, J \in \mathfrak{h}, w \in W$, then $\langle w \cdot H, w \cdot J\rangle = \langle H, J\rangle$. Thus the inner product $\langle \_, \_\rangle$ restricted to $\mathfrak{h}$ is invariant under action of W. In other words, each $w \in W$ is a unitary operator on $(\mathfrak{h}, \langle \_, \_\rangle|_{\mathfrak{h}})$. So $\langle \_, \_\rangle$ induces an identification (bijective correspondence) between $\mathfrak{h}$ and $\mathfrak{h}^*$ via $H \mapsto \langle H, \_\rangle \in \mathfrak{h}^*$, which is a conjugate-linear isomorphism (for any $\alpha \in \mathfrak{h}^*$, there exists a unique $H \in \mathfrak{h}$ such that $\alpha(J) = \langle H, J\rangle$ for all $J \in \mathfrak{h}$).

We now can say that a weight for a representation (V, π) is an element $\alpha \in \mathfrak{h}$ such that $\pi(H)v = \langle \alpha, H \rangle v = \mu(H)v$ for all $H \in \mathfrak{h}$ for some non-zero $v \in V$.

Claim. Under this identification of $\mathfrak{h}$ with $\mathfrak{h}^*$, the action of W on $\mathfrak{h}^*$ coincides with the usual action of W on $\mathfrak{h}$.

Proof. $\mu(H) = \langle \alpha, H \rangle$, so $(w \cdot \mu)(H) = \mu(w^{-1} \cdot H) = \langle \alpha, w^{-1} \cdot H \rangle = \langle w \cdot \alpha, H \rangle$. □

We will now try to describe the roots of $\mathfrak{sl}(3, \mathbb{C})$ (the non-zero weights of the adjoin representation) as elements of $\mathfrak{h}$.

Claim. α_1 and α_2 coincide with H_1 and H_2 respectively. That is,

$$\alpha_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Proof. Verify that $\langle \alpha_1, H_1 \rangle = 2, \langle \alpha_1, H_2 \rangle = -1, \langle \alpha_2, H_1 \rangle = -1, \langle \alpha_2, H_2 \rangle = 2$. □

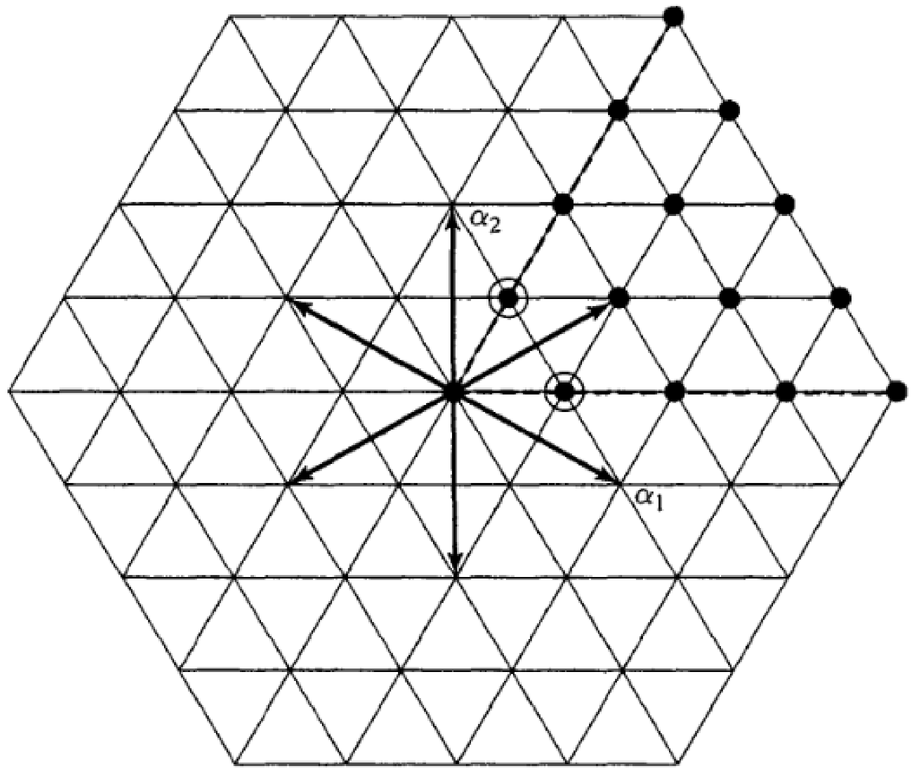
Note $\|\alpha_1\|^2 = \|\alpha_2\|^2 = 2$ and $\langle \alpha_1, \alpha_2 \rangle = -1 = \|\alpha_1\| \|\alpha_2\| \cos\theta = 2\cos\theta$. Hence $\theta = 2\pi/3$. (We shall visualize the two-dimensional real subspace of $\mathfrak{h}$ spanned by $\alpha_1, \alpha_2 \in \mathfrak{h}$.)

Definition 19.3. Let $\mu = (m_1, m_2)$ where m_1, m_2 are non-negative integers (that is, μ is a dominant integral element). These are exactly the highest weights of irreducible representations of $\mathfrak{sl}(3, \mathbb{C})$. Let $\mu_1 \leftrightarrow (1, 0)$ and $\mu_2 \leftrightarrow (0, 1)$. These two are said to be the *fundamental weights*.

Then $\mu_1, \mu_2 \in \mathfrak{h}$, and we have $\langle \mu_1, H_1 \rangle = \langle \mu_2, H_2 \rangle = 1$ and $\langle \mu_1, H_2 \rangle = \langle \mu_2, H_1 \rangle = 0$, where $\mu_1 = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$ and $\mu_2 = \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2$. As elements of $\mathfrak{h}$,

$$\mu_1 = \begin{bmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{bmatrix}.$$

So $\|\mu_1\| = \|\mu_2\| = \frac{\sqrt{6}}{3}$. But $\langle \mu_1, \mu_2 \rangle = \frac{3}{9} = \frac{6}{9}\cos\theta$ so $\theta = \pi/3$. The set of dominant integral elements is the set of non-negative integer linear combinations of μ_1, μ_2. So α_1, α_2 are length $\sqrt{2}$ with angle $2\pi/3$ and μ_1, μ_2 are length $\sqrt{6}/3$ with angle $\pi/3$. This gives us the following weight diagram for the adjoint representation (="root diagram"):



with each vector on the diagram having length $\sqrt{2}$ and the sides of each small triangle having length $\sqrt{6}/3$ and each dot representing a dominant integral element.

Now consider the action of the Weyl group $W = S_3$. Let, for instance, $\sigma = (123)$. Then since

$$\alpha_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \alpha_1 + \alpha_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

it follows that

$$\sigma \cdot \alpha_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \alpha_2$$

$$\sigma \cdot \alpha_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = -\alpha_1 - \alpha_2.$$

So $\sigma = (123)$ is a counterclockwise rotation by $2\pi/3$. As for $\tau = (12)$, we see that

$$\tau \cdot \alpha_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = -\alpha_1$$

$$\tau \cdot \alpha_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \alpha_1 + \alpha_2.$$

Hence $\tau = (12)$ represents a reflection across the line orthogonal to $\text{span}\{\alpha_1\}$. Hence, $S_3 \cong W \cong$ symmetry group of an equilateral triangle.

20. March 16: Structure theory for complex semisimple Lie algebras

Recall that a compact (matrix) Lie group has the complete reducibility property (CRP), which we stated without proof. So the Lie algebra $\mathfrak{g}$ of a compact, *simply connected* matrix group G has the CRP (because there exists a bijective correspondence between the representations preserving irreducibility); in this case, $\mathfrak{g}_{\mathbb{C}}$ also has the CRP. For instance, $\mathfrak{sl}(n, \mathbb{C}) = \mathfrak{su}(n)_{\mathbb{C}}$ and $\mathrm{SU}(n)$ is compact and simply connected, so $\mathrm{SU}(n)$ also has the CRP.

Definition 20.1. Let $\mathfrak{g}$ be a complex Lie algebra. An *ideal* $\mathfrak{h}$ of $\mathfrak{g}$ is a complex (Lie) subalgebra such that if $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$ then $[X, Y] \in \mathfrak{h}$.

Definition 20.2. A complex Lie algebra $\mathfrak{g}$ is called *simple* if $\dim \mathfrak{g} \geq 2$ and contains no non-trivial ideals. A complex Lie algebra $\mathfrak{g}$ is called *semisimple* if it is (isomorphic to) a direct sum of simple Lie algebras (as a Lie algebra). That is, if $\mathfrak{g}$ is semisimple, then there exist $\mathfrak{g}_i$ Lie algebras such that $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_n$, where the Lie bracket is defined to be

$$[X_1 \oplus X_2 \oplus \cdots \oplus X_n, Y_1 \oplus Y_2 \oplus \cdots \oplus Y_n] := [X_1, Y_1] \oplus \cdots \oplus [X_n, Y_n],$$

i.e., $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ whenever $i \neq j$.

A complex Lie algebra is called *reductive* if it is (isomorphic to) $\mathfrak{g} \cong \mathfrak{g}_a \oplus \mathfrak{g}_s$ where $\mathfrak{g}_\alpha$ is abelian and $\mathfrak{g}_s$ is semisimple.

Lemma 20.3. *A subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is an ideal if and only if $\mathfrak{h}$ is an invariant subspace for the adjoint representation of $\mathfrak{g}$.*

Proof. $\mathrm{ad}_X(\mathfrak{h}) \subseteq \mathfrak{h} \Leftrightarrow [X, H] \in \mathfrak{h}$ for all $H \in \mathfrak{h}, X \in \mathfrak{g}$. $\square$

Proposition 20.4. *A complex Lie algebra $\mathfrak{g}$ is reductive if and only if the adjoint representation is completely reducible.*

Proof. ($\Leftarrow$) Suppose that $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is completely reducible. Hence, as a vector space, $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_m$ where each $\mathfrak{g}_i$ is an irreducible invariant subspace of ad. Hence each $\mathfrak{g}_i$ is an ideal of $\mathfrak{g}$ by Lemma 20.3. But since $\mathfrak{g}_i$ is irreducible, $\mathfrak{g}_i$ contains no ideals of $\mathfrak{g}$, other than $\mathfrak{g}_i$ and $\{0\}$. Let $X \in \mathfrak{g}_k, Y \in \mathfrak{g}_l$ with $k \neq l$. Then $[X, Y] \in \mathfrak{g}_k \cap \mathfrak{g}_l = \{0\}$, where the membership follows from the fact that $\mathfrak{g}_k$ and $\mathfrak{g}_l$ are ideals and the equality follows from the fact that $\mathfrak{g}$ is a direct sum. Hence $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \mathfrak{g}_m$ as Lie algebra. We claim that each $\mathfrak{g}_i$ has no non-trivial ideals of $\mathfrak{g}$. If $\mathfrak{h} \subseteq \mathfrak{g}_i$ is an ideal of $\mathfrak{g}_i$ then $[X, Y] \in \mathfrak{h}$ for all $X \in \mathfrak{g}, Y \in \mathfrak{h}$. hence $[X, Y] = 0 \in \mathfrak{h}$ for all $X \in \mathfrak{g}_j, Y \in \mathfrak{h}$ and $j \neq i$. So $[X, Y] \in \mathfrak{h}$ for all $X \in \mathfrak{g}, Y \in \mathfrak{h}$. Hence $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ contained in $\mathfrak{g}_i$ so $\mathfrak{h} = \{0\}$ or $\mathfrak{g}_i$. So $\mathfrak{g}_i$ contains no non-trivial ideals of $\mathfrak{g}_i$.

Let

$$\mathfrak{g}_a = \bigoplus_{\substack{k=1,\ldots,m \\ \dim(\mathfrak{g}_k)=1}} \mathfrak{g}_k, \quad \mathfrak{g}_s = \bigoplus_{\substack{k=1,\ldots,m \\ \dim(\mathfrak{g}_k)\geq 2}} \mathfrak{g}_k.$$

Then $\mathfrak{g}_a$ is the abelian portion, and $\mathfrak{g}_s$ is the semisimple portion, as desired.

($\Rightarrow$) Conversely, suppose that $\mathfrak{g}$ is reductive. Thus,

$$\mathfrak{g} = \mathfrak{g}_s \oplus \mathfrak{g}_a = \left(\bigoplus_{\substack{k=1,\ldots,m \\ \dim(\mathfrak{g}_k)\geq 2}} \mathfrak{g}_k \right) \oplus \left(\bigoplus_{\substack{k=1,\ldots,m \\ \dim(\mathfrak{g}_k)=1}} \mathfrak{g}_k \right).$$

This is a direct sum of irreducible invariant subspaces for the adjoint representation. □

Corollary 20.5. *The complexification of the Lie algebra of a connected complex matrix Lie group is reductive.*

Proof. Every compact group has CRP, so in particular the adjoint (Ad) representation of G is completely reducible. Since G is connected, ad is completely reducible if and only if Ad is completely reducible. Hence $\mathrm{ad} : \mathfrak{g}_{\mathbb{C}} \to \mathfrak{gl}(\mathfrak{g}_{\mathbb{C}})$ is completely reducible also. □

Theorem 20.6. *A complex Lie algebra $\mathfrak{g}$ is semisimple if and only if it is (isomorphic to) the complexification of the Lie algebra of a* simply connected compact *matrix Lie group.*

Proof. ($\Rightarrow$) this direction is beyond the scope of this course – we need some deep structure theory of complex Lie algebras.

($\Leftarrow$) This one is easier, but we need some results from algebraic topology, which we are not assuming. So we will discuss underlying ideas. Start with K, a compact and simply connected matrix Lie group. Let $\mathfrak{k}$ be the Lie algebra of K. Let $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$ be its complexification. We already know that $\mathfrak{g} = \mathfrak{g}_a \oplus \mathfrak{g}_s$ is reductive, by the above corollary. One can show that $\mathfrak{k} = \mathfrak{k}_a \oplus \mathfrak{k}_s$ as Lie algebras. Then $K = K_a \times K_s$ where K_a and K_s are compact, simply connected matrix groups such that $\mathrm{Lie}(K_a) = \mathfrak{k}_a$ and $\mathrm{Lie}(K_s) = \mathfrak{k}_s$. Thus K_a is a commutative Lie group. But a result from algebraic topology says that the only simply connected abelian matrix group is isomorphic to $\mathbb{R}^n$, which is non-compact if $n \geq 1$. Therefore $n = 0$, so $\mathfrak{k}_a = \{0\}$ so $\mathfrak{g} = \mathfrak{g}_s$ as required. □

Definition 20.7. Let $\mathfrak{g}$ be a complex Lie algebra. Then a *compact real form* of $\mathfrak{g}$ is a real subalgebra $\mathfrak{k}$ such that:

(1) $\mathfrak{k}_{\mathbb{C}} = \mathfrak{g}$
(2) there exists a compact simply connected matrix Lie group K_1 such that the $\mathfrak{k}_1 = \mathrm{Lie}(K_1) \cong \mathfrak{k}$ (isomorphic, *not necessarily* equal).

Remark 20.1. The previous theorem says that a complex semisimple Lie algebra has a compact real form. The compact real from of a complex semisimple Lie algebra is unique up to conjugation.

Example 20.8. $\mathfrak{sl}(n, \mathbb{C})$ has a compact real form $\mathfrak{su}(n)$ since $\mathrm{SU}(n)$ is compact, simply connected, and $\mathfrak{su}(n)_{\mathbb{C}} = \mathfrak{sl}(n, \mathbb{C})$.

Proposition 20.9. *Let $\mathfrak{g} \subseteq \mathfrak{gl}(n, \mathbb{C})$ be a complex semisimple Lie algebra with compact real form $\mathfrak{k}$. Let K be the analytic subgroup of* $\mathrm{GL}(n, \mathbb{C})$ *whose Lie algebra is $\mathfrak{k}$. Then K is compact (but not necessarily simply connected).*

Proof. By the definition of a compact real form, there exists a compact simply connected matrix group K_1 whose Lie algebra $\mathfrak{k}_1$ is isomorphic to $\mathfrak{k}$. Let $\phi : \mathfrak{k}_1 \to \mathfrak{k} \subseteq \mathfrak{gl}(n, \mathbb{C})$ be the Lie algebra isomorphism. Because K_1 is simply connected, there exists a lift $\Phi : K_1 \to \mathrm{GL}(n, \mathbb{C})$

which is a Lie group homomorphism and $\Phi_* = \phi$. Let $K = \Phi(K_1)$. This is compact. We claim that K is the analytic subgroup of $\mathrm{GL}(n,\mathbb{C})$ with Lie algebra $\mathfrak{k}$. Let $\widetilde{K}$ be the analytic subgroup of $\mathrm{GL}(n,\mathbb{C})$ with Lie algebra $\mathfrak{k}$. Since K_1 is connected, every $A \in K_1$ is of the form

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

for each $X_i \in \mathfrak{k}_i$. It follows that

$$\Phi(A) = \Phi(e^{X_1}e^{X_2}\cdots e^{X_m}) = e^{\phi(X_1)}e^{\phi(X_2)}\cdots e^{\phi(X_m)} \in \widetilde{K},$$

since $\widetilde{K} = \{e^{Y_1}e^{Y_2}\cdots e^{Y_m} : Y_i \in \mathfrak{k}\}$. Hence $K = \Phi(K_1) \subseteq \widetilde{K}$.

Conversely, let $B \in \widetilde{K}$. Then $B = e^{Y_1}e^{Y_2}\cdots e^{Y_m}$ with $Y_i \in \mathfrak{k}$. Let $X_i \in \mathfrak{k}_i$ be such that $\phi(X_i) = Y_i$. Then $B = e^{\phi(X_1)}\cdots e^{\phi(X_m)} = \Phi(e^{X_1}e^{X_2}\cdots e^{X_m}) \subseteq \Phi(K_1)$. Hence $\widetilde{K} \subseteq \Phi(K_1) = K$. ☐

Remark 20.2. K need not be simply connected. For instance, take $\mathfrak{g} = \mathfrak{so}(3,\mathbb{C}) \subseteq \mathfrak{gl}(3,\mathbb{C})$ is the complexification of $\mathfrak{k} = \mathfrak{so}(3,\mathbb{R})$ which is isomorphic to the Lie algebra of a complex, simply connected group, because $\mathfrak{so}(3,\mathbb{R}) \cong \mathfrak{su}(2) = \mathfrak{k}_1$ hence $\mathrm{SU}(2) = K_1$. Hence $\mathfrak{so}(3,\mathbb{R})$ is the compact real form of $\mathfrak{so}(3,\mathbb{C})$ but the analytic subgroup of $\mathrm{GL}(3,\mathbb{C})$ whose Lie algebra is $\mathfrak{so}(3,\mathbb{R})$ is $K = \mathrm{SO}(3)$ which we know is not simply connected.

Corollary 20.10. *Every complex semisimple Lie algebra has the complete reducibility property.*

Proof. There is a bijective correspondence between the representations of $\mathfrak{g}$ and those of $\mathfrak{k}$; and there is also a bijective correspondence between the representations of $\mathfrak{k}$ and those of $\mathfrak{k}_1$, since $\mathfrak{k} \cong \mathfrak{k}_1$. These are erectly the representations of K_1 which is compact, hence has the CRP. ☐

Remark 20.3. Let's summarize what we have done so far. Namely, we showed that the following characterizations are equivalent:

(1) $\mathfrak{g}$ is a complex semisimple Lie algebra
(2) $\mathfrak{g}$ is a direct sum of of simple Lie algebras (as a Lie algebra)
(3) $\mathfrak{g}$ is the complexification of a real LIe algebra $\mathfrak{k}$ which is isomorphic to the Lie algebra of a compact simply connected matrix group
(4) $\mathfrak{g}$ has the complete reducibility property.

20.1. *Examples of complex Lie algebras*

$\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$	compact real form $\mathfrak{k}$	K_1
$\mathfrak{sl}(n,\mathbb{C}), n \geq 2$ (semisimple)	$\mathfrak{su}(n)$	$\mathrm{SU}(n)$
$\mathfrak{so}(n,\mathbb{C}), n \geq 3$ (semisimple)	$\mathfrak{so}(n)$	$\mathrm{Spin}(n)$
$\mathfrak{so}(2,\mathbb{C})$ (reductive but *not semisimple*)	N/A	N/A
$\mathfrak{gl}(n,\mathbb{C})$ (reductive but *not semisimple*)	N/A	N/A
$\mathfrak{sp}(n,\mathbb{C})$ (semisimple)	$\mathfrak{sp}(n)$	$\mathrm{Sp}(n)$

A few things about the table above:

(1) $\mathrm{Spin}(n)$ denotes the simply connected complex matrix group with Lie algebra isomorphic to $\mathfrak{so}(n)$.
(2) $\mathfrak{so}(2,\mathbb{C})$ is not semisimple since $\mathfrak{so}(2) = \mathfrak{so}(2,\mathbb{R}) \ncong$ Lie algebra of a simply connected, compact group.
(3) $\mathfrak{gl}(n,\mathbb{C})$ is not semisimple since $\mathfrak{gl}(n,\mathbb{C}) = \mathfrak{u}(n)_{\mathbb{C}}$ and $\mathrm{U}(n)$ is not simply connected.

Remark 20.4. The semisimple ones $\mathfrak{sl}(n,\mathbb{C})(n \geq 2), \mathfrak{so}(n,\mathbb{C})(n \geq 3), \mathfrak{sp}(n,\mathbb{C})(n \geq 1)$ are in fact all simple, except for $\mathfrak{so}(4,\mathbb{C}) \cong \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})$.

Theorem 20.11 (Classification theorem for complex simple Lie algebras)**.** *If $\mathfrak{g}$ is a complex simple Lie algebra, then $\mathfrak{g}$ is either one of*

$$\begin{array}{c|c|l} \mathfrak{sl}(n,\mathbb{C}) & A_{n-1} & n \geq 2 \\ \mathfrak{so}(2n,\mathbb{C}) & D_n & n \geq 3 \\ \mathfrak{so}(2n+1,\mathbb{C}) & B_n & n \geq 1 \\ \mathfrak{sp}(n,\mathbb{C}) & C_n & n \geq 1 \end{array}$$

or one of the five exceptional *Lie algebras:* $\mathfrak{g}_2, \mathfrak{f}_4, \mathfrak{e}_6, \mathfrak{e}_7, \mathfrak{e}_8$.

21. March 18

21.1. *Cartan subalgebras*

Definition 21.1. Let $\mathfrak{g}$ be a complex Lie algebra. A *Cartan subalgebra* $\mathfrak{h}$ is a vector subspace of $\mathfrak{g}$ such that:

- $[H_1, H_2] = 0$ for all $H_1, H_2 \in \mathfrak{h}$;
- if $X \in \mathfrak{g}$ and $[X, H] = 0$ for all $H \in \mathfrak{h}$ then $X \in \mathfrak{h}$;
- for all $H \in \mathfrak{h}$, the map $\mathrm{ad}_H : \mathfrak{g} \to \mathfrak{g}$ is diagonalizable.

Remark 21.1. If $H_1, H_2 \in \mathfrak{h}$ then $[\mathrm{ad}_{H_1}, \mathrm{ad}_{H_2}] = \mathrm{ad}_{[H_1,H_2]} = \mathrm{ad}_0 = 0$ so all ad_H's commute for $H \in \mathfrak{h}$. Hence, if $\mathfrak{h}$ is a Cartan subalgebra, then all the ad_H's are *simultaneously diagonalizable* for all $H \in \mathfrak{h}$.

Proposition 21.2. *If $\mathfrak{g} \subseteq \mathfrak{gl}(n,\mathbb{C})$ is a complex semisimple Lie algebra, then a Cartan subalgebra* always *exists.*

Proof. Let $\mathfrak{k}$ be a compact real fem of $\mathfrak{g}$. Let $\mathfrak{t}$ be a maximal abelian Lie subalgebra of $\mathfrak{k}$ (start with any one-dimensional subalgebra and keep "increasing" until maximal). Define $\mathfrak{h} = \mathfrak{t}_{\mathbb{C}} = \mathfrak{t} + i\mathfrak{t}$. Then $\mathfrak{h}$ is an abelian subalgebra of $\mathfrak{g}$.

Let $X \in \mathfrak{g}$. Suppose that $[X, H] = 0$ for all $H \in \mathfrak{h}$. Then $[X, H] = 0$ for all $H \in \mathfrak{t}$. Write $X = X_1 + iX_2$ where $X_1, X_2 \in \mathfrak{k}$. Then $[X_1 + iX_2, H] = [X_1, H] + i[X_2, H] = 0$ for all $H \in \mathfrak{t}$. By maximality of $\mathfrak{t}$ in $\mathfrak{k}$, we get $X_1, X_2 \in \mathfrak{t}$ so $X \in \mathfrak{h}$. Therefore $\mathfrak{h}$ is maximal. It still remains to show that every ad_H is diagonalizable for all $H \in \mathfrak{h}$. Let K be the analytic subgroup of $\mathrm{GL}(n,\mathbb{C})$ whose Lie algebra is $\mathfrak{k}$. We showed last time that K is compact.

Claim. There exists a real-valued positive-definite inner product on $\mathfrak{k}$ that is invariant under the adjoint action of K. That is, there exists a positive-definite inner product $\cdot : \mathfrak{k} \times \mathfrak{k} \to \mathbb{R}$ such that fro all $A \in K$, the following identity is satisfied:

$$(\mathrm{Ad}_A X) \cdot (\mathrm{Ad}_A Y) = X \cdot Y,$$

where $\mathrm{Ad}_A X = AXA^{-1}$ as usual.

How do we prove this claim? The idea is to choose any inner product $\cdot_o$ on K, and define

$$X \cdot Y := \int_K (\mathrm{Ad}_A X) \cdot_o (\mathrm{Ad}_A Y) \ \mathrm{vol}_K .$$

(We can integrate over compact groups with a left-invariant volume form.) We can extend this inner product on $\mathfrak{k}$ to a Hermitian positive-definite inner product on $\mathfrak{k}_{\mathbb{C}} = \mathfrak{g}$ in the usual way. Let $\langle \_, \_ \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ such that $\langle X_1 + iX_2, Y_1 + iY_2 \rangle = (X_i - iX_2) \cdot (Y_1 + iY_2)$. Then by extending $\cdot$ to $\mathfrak{k}_{\mathbb{C}}$ by complex bilinearity, we have

$$\langle X_i + iX_2, Y_1 + iY_2 \rangle = (X_1 - iX_2) \cdot (Y_1 + iY_2) = (X_1 \cdot Y_1 + X_2 \cdot Y_2) + i(X_1 \cdot Y_2 - X_2 \cdot Y_1).$$

It is a straightforward verification to check if $\langle \_, \_ \rangle$ is a positive-definite Hermitian inner product. So $\langle \operatorname{Ad}_A X, \operatorname{Ad}_A Y \rangle = (\operatorname{Ad}_A X) \cdot (\operatorname{Ad}_A Y) = X \cdot Y = \langle X, Y \rangle$ for all $A \in K$ and $X, Y \in \mathfrak{g}$. Moreover, $\langle \_, \_ \rangle$ takes real values on $\mathfrak{k}$. Let $A \in K$ with $\langle \operatorname{Ad}_A X, \operatorname{Ad}_A Y \rangle = \langle X, Y \rangle$ for all $X, Y \in \mathfrak{g}$. Let $z \in \mathfrak{k}$. Then $e^{tZ} \in K$ and $\operatorname{Ad}_{e^{tZ}} = e^{t \operatorname{ad}_Z}$. Therefore $\langle e^{t \operatorname{ad}_Z} X, e^{t \operatorname{ad}_Z} Y \rangle = \langle X, Y \rangle$ for all $t \in \mathbb{R}$. Take the derivative on both sides at $t = 0$ to get

$$\langle \operatorname{ad}_Z X, Y \rangle + \langle X, \operatorname{ad}_Z Y \rangle = 0$$

if $X, Y \in \mathfrak{g}$ and $Z \in \mathfrak{k}$. So $\operatorname{ad}_Z : \mathfrak{g} \to \mathfrak{g}$ is skew-Hermitian for all $Z \in \mathfrak{k}$. In particular, ad_H is diagonalizable for all $H \in \mathfrak{t}$, since $\mathfrak{t} \in \mathfrak{k}$. Let $H = H_1 + iH_2 + \mathfrak{h}$ where $H_i \in \mathfrak{t}$. Then $\operatorname{ad}_{H_1}, \operatorname{ad}_{H_2}$ commute and are both diagonalizable; hence $\operatorname{ad}_{H_1}, \operatorname{ad}_{H_2}$ are simultaneously diagonalizable. So $\operatorname{ad}_H = \operatorname{ad}_{H_1} + i \operatorname{ad}_{H_2}$ is also diagonalizable. □

Remark 21.2. If $\mathfrak{g}$ is not semisimple, then there may not exist a Cartan subalgebra (see Assignment #6 for more).

Remark 21.3. If $\mathfrak{g}$ is complex semisimple, then *every* Cartan subalgebra of $\mathfrak{g}$ arises in this way, and they are all conjugate to each other (hence all isomorphic). This involves some deep structure theory which is beyond the scope of this course, so we will not prove this fact.

Definition 21.3. Let $\mathfrak{g}$ be a complex semisimple Lie algebra. Then the *rank of* $\mathfrak{g}$ is the dimension of any Cartan subalgebra.

Example 21.4. $\mathfrak{sl}(n, \mathbb{C}) = A_{n-1}$ has rank $n - 1$. Therefore $\mathfrak{sl}(2, \mathbb{C})$ has rank 1, since $\mathfrak{h} = \operatorname{span}\{H\}$; $\mathfrak{sl}(3, \mathbb{C})$ has rank 2 as $\mathfrak{h} = \operatorname{span}\{H_1, H_2\}$. $\mathfrak{so}(2n, \mathbb{C}) = D_n, \mathfrak{so}(2n+1, \mathbb{C}) = B_n, \mathfrak{sp}(n, \mathbb{C}) = C_n$ all have rank n. Similarly, $\mathfrak{g}_2, \mathfrak{f}_4, \mathfrak{e}_6, \mathfrak{e}_7, \mathfrak{e}_8$ have rank $2, 4, 6, 7, 8$ respectively.

For the rest of the course until stated otherwise, we will assume that:

(1) $\mathfrak{g}$ is a complex semisimple Lie algebra
(2) $\mathfrak{k}$ is a compact real form of $\mathfrak{g}$ with maximal abelian subalgebra $\mathfrak{t}$
(3) $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t} = \mathfrak{t}_{\mathbb{C}}$ is a Cartan subalgebra of $\mathfrak{g}$
(4) We have chosen a positive-definite Hermitian inner product $\langle \_, \_ \rangle$ on $\mathfrak{g}$ that is Ad_A-invariant for all $A \in K$, and takes real values on $\mathfrak{k}$.

21.2. *Roots and root spaces*

Definition 21.5. A *root of* $\mathfrak{g}$ (relative to the Cartan subalgebra $\mathfrak{h}$) is a non-zero linear functional $\alpha \in \mathfrak{h}^*$ such that there exists $X \neq 0$ such that $[H, X] = \operatorname{ad}_H X = \alpha(H)X$ for all $H \in \mathfrak{h}$. In other words, a root is a collection of simultaneous eigenvalues of $\operatorname{ad}_H : \mathfrak{g} \to \mathfrak{g}$ for all $H \in \mathfrak{h}$.

Remark 21.4. Note that α has to be linear:

$$\begin{aligned} \alpha(a_1 H_1 + a_2 H_2)X &= \operatorname{ad}_{a_1 H_1 + a_2 H_2}(X) = [a_1 H_1 + a_2 H_2, X] \\ &= a_1 \operatorname{ad}_{H_1} X + a_2 \operatorname{ad}_{H_2} X = (a_1 \alpha(H_1) + a_2 \alpha(H_2))X. \end{aligned}$$

Let R be the set of roots of $\mathfrak{g}$.

Proposition 21.6. *If $\alpha \in R$, then $\alpha(H) \in i\mathbb{R}$ for all $H \in \mathfrak{t}$.*

Proof. We showed that ad_H is skew-adjoint for all $H \in Z$, and the eigenvalues of a skew-adjoint operator are all purely imaginary. □

Definition 21.7. Let α be a root of $\mathfrak{g}$. Then the root space $\mathfrak{g}_\alpha$ is defined to be

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} : [H, X] = \alpha(H)X \text{ for all } H \in \mathfrak{h}\},$$

i.e., $\mathfrak{g}$ is the set of all simultaneous eigenvalues of all ad_H's with eigenvalue $\alpha(H)$ plus the zero vector.

In fact, $\mathfrak{g}_\alpha$ is a *subspace* of $\mathfrak{g}$. In general, for any $\alpha \in \mathfrak{h}^*$, then

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} : [H, X] = \alpha(H)X \text{ for all } H \in \mathfrak{h}\}.$$

Thus if $\alpha \neq 0$, then $\mathfrak{g}_\alpha = \{0\}$ unless α is a root. Meanwhlie, if $\alpha = 0$, then

$$\mathfrak{g}_0 = \{X \in \mathfrak{g} : [H, X] = 0 \text{ for all } H \in \mathfrak{h}\} = \mathfrak{h}$$

by maximality of a Cartan subalgebra. Since all the ad_H's are simultaneously diagonalizable, there exists a basis of $\mathfrak{g}$ consisting of simultaneous eigenvectors of all ad_H's (i.e., root vectors or elements of $\mathfrak{h}$). Hence, *as a vector space*, there exists a decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \left(\bigoplus_{\alpha \in R} \mathfrak{g}_\alpha \right),$$

so every element of $\mathfrak{g}$ is expressible uniquely as an element of $\mathfrak{h}$ plus one root vector from each root space.

Proposition 21.8. *Let $\alpha, \beta \in \mathfrak{h}^*$. Then $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$ i.e., if $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$, then $[X, Y] \in \mathfrak{g}_{\alpha+\beta}$.*

Proof. Suppose $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$. Then $[H, X] = \alpha(H)X$ and $[H, Y] = \beta(H)Y$ for all $H \in \mathfrak{h}$. Recall the Jacobi identity: $[H, [X, Y]] = [[H, X], Y] + [X, [H, Y]]$, or equivalently, $\mathrm{ad}_H[X, Y] = [\mathrm{ad}_H X, Y] + [X, \mathrm{ad}_H Y]$, i.e., ad_H is a Lie algebra derivation.

IN this case, for all $H \in \mathfrak{h}$, we have $[H, [X, Y]] = \alpha(H)[X, Y] + \beta(H)[X, Y] = (\alpha + \beta)(H)[X, Y]$. Hence $[X, Y] \in \mathfrak{g}_{\alpha+\beta}$. □

Corollary 21.9. $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{h} = \mathfrak{g}_0$.

Corollary 21.10. *If $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$ and $\alpha + \beta \neq 0$ and $\alpha + \beta$ is not a root, then* $[X, Y] = 0$.

Next, we shall investigate restrictions on the set R of roots of $\mathfrak{g}$.

Proposition 21.11. *If α is a root, then so is $-\alpha$. Also, the roots span $\mathfrak{h}^*$.*

Proof. Since α is a root, then there exists $X \neq 0$ in $\mathfrak{g}$ such that $[H, X] = \alpha(H)X$ for all $H \in \mathfrak{h}$. Hence $H \in \mathfrak{t}$. Write $X = X_1 + iX_2$ and $X_j \in \mathfrak{k}$. Let $H \in \mathfrak{t}$. Then $[H, X_1] + i[H, X_2] = \alpha(H)(X_1 + iX_2)$. We have seen that $\alpha(H) = i\theta$ for some $\theta \in \mathbb{R}$ since $H \in \mathfrak{t}$. So we have $[H, X_1] + i[H, X_2] = i\theta(X_1 + iX_2) = -\theta X_2 + i\theta X_1$. Therefore $[H, X_1] = -\theta X_2$ and $[H, X_2] = \theta X_1$. Define $\overline{X} := X_1 - iX_2 \neq 0$. Then, for all $H \in \mathfrak{t}$,

$$[H, \overline{X}] = [H, X_1] - i[H, X_2] = -\theta X_2 - i\theta X_1 = -i\theta(X_1 - iX_2) = -i\theta\overline{X} = -\alpha(H)\overline{X}.$$

Thus $-\alpha$ is a root of $\mathfrak{g}$, and $\overline{X}$ is a root vector in $\mathfrak{g}_{-\alpha}$. This completes the proof of the first part.

As for the second part, start by assuming that roots do not span $\mathfrak{h}^*$. Choose a maximal linearly independent subset $\alpha_1, \ldots, \alpha_k$ of R and complete it to a basis $\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_r$ of $\mathfrak{h}^*$ where $r = \operatorname{rank}(\mathfrak{g})$. Let $Z_1, \ldots, Z_k, Z_{k+1}, \ldots, Z_r$ be the dual basis of $\mathfrak{h}$. Note that $Z_{k+1} \neq 0$ since it is a basis element. Note also that $\alpha(Z_{k+1}) = 0$ for all $\alpha \in R$ because every root is a linear combination of $\alpha_1, \ldots, \alpha_k$. Therefore $[H, Z_{k+1}] = 0$ for all $H \in \mathfrak{h}$ since $Z_{k+1} \in \mathfrak{h}$ and $[Z_{k+1}, X_\alpha] = \alpha(Z_{k+1})X_\alpha = 0$ for all $\alpha \in R$ and $X_\alpha \in \mathfrak{g}_\alpha$. Hence, since

$$\mathfrak{g} = \mathfrak{h} \oplus \left(\bigoplus_{\alpha \in R} \mathfrak{g}_\alpha \right),$$

we get $[Z_{k+1}, X] = 0$ for all $X \in \mathfrak{g}$. So Z_{k+1} is in the centre of $\mathfrak{g}$. But since $\mathfrak{g}$ is semisimple, the centre of $\mathfrak{g}$ is trivial. Therefore $Z_{k+1} = 0$ which is a contradiction. □

Lemma 21.12. *Let $X \in \mathfrak{g}_\alpha, Y \in \mathfrak{g}_{-\alpha}, H \in \mathfrak{h}$. Then $[X, Y] \in \mathfrak{h}$, and $\langle [X, Y], H \rangle = \alpha(H)\langle Y, -\overline{X} \rangle$, where $X = X_1 + iX_2$ and $-\overline{X} = -X_1 + iX_2$, where $X_j \in \mathfrak{k}$.*

Proof. If $X \in \mathfrak{k}$ we showed that ad_X is skew-Hermitian, i.e., $\langle \operatorname{ad}_X Y, Z \rangle = -\langle Y, \operatorname{ad}_X Z \rangle$ for all $x \in \mathfrak{k}$ and $Y, Z \in \mathfrak{g}$. Write $X = X_i + iX_2$ with $X_j \in \mathfrak{k}$, and compute $\langle \operatorname{ad}_X Y, Z \rangle$:

$$\begin{aligned} \langle \operatorname{ad}_X Y, Z \rangle = \langle \operatorname{ad}_{X_1 + iX_2} Y, Z \rangle &= \langle \operatorname{ad}_{X_1} Y + i \operatorname{ad}_{X_2} Y, Z \rangle = \langle \operatorname{ad}_{X_1} Y, Z \rangle - i \langle \operatorname{ad}_{X_2} Y, Z \rangle \\ &= -\langle Y, \operatorname{ad}_{X_1} Z \rangle + \langle Y, i \operatorname{ad}_{X_2} Z \rangle = \langle Y, \operatorname{ad}_{-X_1 + iX_2} Z \rangle = \langle Y, \operatorname{ad}_{-\overline{X}} Z \rangle. \end{aligned}$$

Therefore $\langle \operatorname{ad}_X Y, Z \rangle = \langle Y, \operatorname{ad}_{-\overline{X}} Z \rangle$ for all $X, Y, Z \in \mathfrak{g}$. Hence

$$\begin{aligned} \langle [X, Y], H \rangle = \langle \operatorname{ad}_X Y, H \rangle &= \langle Y, \operatorname{ad}_{-\overline{X}} H \rangle = \langle Y, [-\overline{X}, H] \rangle \\ &= \langle Y, [H, \overline{X}] \rangle = -\alpha(H)\langle Y, \overline{X} \rangle = \alpha(H)\langle Y, -\overline{X} \rangle. \end{aligned}$$ □

22. March 23

Our main goal today is to prove the following theorem:

Theorem 22.1 (Main theorem on root spaces)**.** *The following statements are true for root spaces:*

(1) If α is a root, then the only *multiples of α that are roots are α and $-\alpha$.*
(2) If α is a root, then g_α is one-dimensional.
(3) For every root $\alpha \in R$, there exists $0 \neq X_\alpha \in \mathfrak{g}, 0 \neq Y_\alpha \in \mathfrak{g}_{-\alpha}, 0 \neq H_\alpha \in \mathfrak{h}$ such that $[H_\alpha, X_\alpha] = 2X_\alpha, [H_\alpha, Y_\alpha] = -2Y_\alpha, [X_\alpha, Y_\alpha] = H_\alpha$; and H_α is unique*, independent of X_α and Y_α.*

Proof. Let $\alpha \in R$ be a *fixed root.*

Claim (Claim 1). If $x \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$ then $[X, Y] \in \mathfrak{h}$ and $[X, Y]$ is orthogonal to all elements $H \in \mathfrak{h}$ such that $\alpha(H) = 0$.

Proof of Claim 1. Let $\ker(\alpha) = \{H \in \mathfrak{h} : \alpha(H) = 0\}$. We can write $\mathfrak{h} = \ker(\alpha) + \ker(\alpha)^\perp$ (as vector spaces; $\ker(\alpha)^\perp$ is the orthogonal complement in $\mathfrak{h}$ with respect to $\langle \_, \_ \rangle|_{\mathfrak{h}}$). Let $\sigma = \operatorname{rank}(\mathfrak{g}) = \dim(\mathfrak{h})$. Since $\alpha \neq 0$ by definition, $\ker(\alpha)$ has dimension $r - 1$. Hence $\dim(\ker(\alpha)^\perp) = 1$. By Claim 1, we have

$$[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \ker(\alpha)^\perp. \tag{†}$$

Let X_α be non-zero in $\mathfrak{g}_\alpha$. Then $-\overline{X} \in \mathfrak{g}_{-\alpha}$, and $\overline{X} \neq 0$ (by Lemma 21.12). □

Claim (Claim 2). $[X, -\overline{X}] \neq 0$ and $\alpha([X, -\overline{X}])$ is real and *strictly positive*.

Proof of Claim 2. Take $Y = -\overline{X}$ in Lemma 21.12. Then $\langle [X, -\overline{X}], H \rangle = \alpha(H) | -\overline{X}||^2$, and $| -\overline{X}|^2$ is positive. Choose $H \in \mathfrak{h}$ such that $\alpha(H) \neq 0$. So $[X, -\overline{X}] \neq 0$.

Let $H = [X, -\overline{X}] \in \mathfrak{h}$. Note that $\|[X, -\overline{X}]\|^2 = \alpha([X, -\overline{X}]) \| -\overline{X}\|^2$, so $\alpha([X, -\overline{X}])$ is real and positive. □

Let $Y = -\overline{X}$. Then $H = [X, Y] = [X, -\overline{X}]$.

Since $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, we have

$$[H, X] = \alpha(H)X$$
$$[H, Y] = \alpha(H)Y.$$

Therefore $\alpha(H) = \alpha([X, -\overline{X}]) > 0$ by Claim 2. Define

$$X_\alpha := \sqrt{\frac{2}{\alpha(H)}} X$$
$$Y_\alpha := \sqrt{\frac{2}{\alpha(H)}} Y$$
$$H_\alpha := \frac{2}{\alpha(H)} H.$$

Thus we have

$$[H_\alpha, X_\alpha] = \left(\frac{2}{\alpha(H)}\right)^{3/2} [H, X] = \left(\frac{2}{\alpha(H)}\right)^{3/2} \alpha(H)X = 2X_\alpha$$
$$[H_\alpha, Y_\alpha] = -2Y_\alpha$$
$$[X_\alpha, Y_\alpha] = H_\alpha.$$

Notice that $[H_\alpha, X_\alpha] = 2X_\alpha = \alpha(H_\alpha)X_\alpha$. Since $X_\alpha \neq 0$, it follows $\alpha(H_\alpha) = 2$.

Claim (Claim 3). If $\beta = k\alpha$ is a root of $\mathfrak{g}$ for some $k \in \mathbb{C}$, then $k = \frac{m}{2}$ for some integer $m \in \mathbb{Z}$.

Proof of Claim 3. Let $S_\alpha := \operatorname{span}\{X_\alpha, Y_\alpha, H_\alpha\} \cong \mathfrak{sl}(2, \mathbb{C})$. Let V_α be the subspace of $\mathfrak{g}$ spanned by $\ker(\alpha)^\perp$ and the root spaces $\mathfrak{g}_\beta$ where $\beta = k\alpha$ for some $k \in \mathbb{C}$. We shall show that V_α is invariant under S_α, with respect to the adjoint action. Hence it is isomorphic to a finite-dimensional complex representation of $\mathfrak{sl}(2, \mathbb{C})$. In other words, $\operatorname{ad}|_{S_\alpha}$ has V_α as an invariant subspace.

We need to show that $[S_\alpha, V_\alpha] \subseteq V_\alpha$ if $Z \in \ker(\alpha)^\perp$. Then $Z = \lambda H_\alpha$ for some $\lambda \in \mathbb{C}$ (because $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \ker(\alpha)^\perp$ by Claim 1). Therefore $\ker(\alpha)^\perp = \operatorname{span}\{H_\alpha\}$. Note that $\operatorname{ad}_{H_\alpha} Z = [H_\alpha, \lambda H_\alpha] = 0$. If $Z \in \mathfrak{g}_\beta$, we have $\operatorname{ad}_{H_\alpha} Z = \beta(H_\alpha)Z = 2kZ$. Therefore $\operatorname{ad}_{H_\alpha}(V_\alpha) \subseteq V_\alpha$. Also,

$$[X_\alpha, H_\alpha] = -2X_\alpha \in \mathfrak{g}_\alpha \subseteq V_\alpha$$
$$[Y_\alpha, H_\alpha] = 2Y_\alpha \in \mathfrak{g}_{-\alpha} \subseteq V_\alpha.$$

Let $Z \in \mathfrak{g}_\beta$.Then $[X_\alpha, Z] \in \mathfrak{g}_{\alpha+\beta} = \mathfrak{g}_{(k+1)\alpha} \subseteq V_\alpha$ and $[Y_\alpha, Z] = \mathfrak{g}_{-\alpha+\beta} = \mathfrak{g}_{(k-1)\alpha} \subseteq V_\alpha$. Therefore $\mathrm{ad}_{X_\alpha}, \mathrm{ad}_{Y_\alpha}, \mathrm{ad}_{H_\alpha}$ leave V_α invariant, so V_α is a finite-dimensional representation of S_α. From what we know about finite-dimensional representations of $\mathfrak{sl}(2, \mathbb{C})$, the eigenvalues of H_α must be integers. If $0 \neq Z \in \mathfrak{g}_\beta$ with $\beta = k\alpha$, then

$$\mathrm{ad}_{H_\alpha}(Z) = [H_\alpha, Z] = \beta(H_\alpha)Z = 2kZ.$$

Hence $2k$ must be integer, so $k \in \frac{1}{2}\mathbb{Z}$ as desired. □

Claim (Claim 4). If α is a root, then 2α is *not* a root.

Proof of Claim 4. Recall $S_\alpha \subseteq V_\alpha$, so by the complete reducibility property (CRP) of $\mathfrak{sl}(2, \mathbb{C})$, V_α decomposes into

$$V_\alpha = S_\alpha \oplus U_1 \oplus \cdots U_p$$

as representations of $S_\alpha \cong \mathfrak{sl}(2, \mathbb{C})$, with each U_i irreducible representations of $\mathfrak{sl}(2, \mathbb{C})$. Suppose that 2α is a root of $\mathfrak{g}$. Then there exists a non-zero $Z \in \mathfrak{g}_{2\alpha}$ such that $[H_\alpha, Z] = (2\alpha)(H_\alpha)Z = 4Z$. Because weight vectors correspond to distinct weights are independent and $\lambda = 4$ occurs as an eigenvalue of ad_{H_α} in V_α, it must occur in one of the factors. Since S_α cannot have eigenvalue 4 (S_α has $-2, 0, 2$ as eigenvalues), there exists some $U_i \neq 0$ such that $\lambda = 4$ occurs as an eigenvalue. By our knowledge of irreducible representations of $\mathfrak{sl}(2, \mathbb{C})$, if m occurs as an eigenvalue of $\pi(H_\alpha)$, then so does $m-2, m-4, \cdots, -m$. If $\lambda = 4$ occurs in U_i, so does $\lambda = 0$. So suppose that $W \in U_i$ is a non-zero vector such that $[H_\alpha, W] = 0$. Hence $W \in \ker(\alpha)^\perp = \mathrm{span}\{H_\alpha\}$. Why? note that $V_\alpha = \ker(\alpha)^\perp \oplus \left(\bigoplus_{\beta = k\alpha} \mathfrak{g}_\beta \right)$. Recall that $\ker(\alpha)^\perp$ is a weight space of ad_{H_α} with height 0, with the remaining portion being weight spaces of ad_{H_α} with weight $\beta = k\alpha \neq 0$. But $U_i \cap \ker(\alpha)^\perp \subseteq U_i \cap S_a = \{0\}$, so this contradicts the fact that W is non-zero. Therefore 2α is not a root. □

Claim (Claim 5). The only multiples of α that are roots are $\pm\alpha$.

Proof of Claim 5. Let α be a root, and $\beta = k\alpha$ a root of $k \in \mathbb{C}^*$. By Claim 3, we have $k = m/2$ with $m \in \mathbb{Z}$. Without loss of generality, let $k > 0$ because otherwise we can replace α with $-\alpha$. So by Claim 4,

$$k \in \left\{ \frac{1}{2}, 1, \frac{3}{2} . \frac{5}{2}, \frac{6}{2}, \cdots \right\} =: N.$$

And $\alpha = \frac{1}{k}\beta$ are both roots, meaning that $k^{-1} \in N$, so $k = 1$. □

Part (a) of the theorem follows from the five claims.

Claim (Claim 6). The root spaces $\mathfrak{g}_\alpha$ are one-dimensional.

Proof of Claim 6. Suppose otherwise. Then there exists X' independent from X_α such that $\mathrm{ad}_{H_\alpha}(X') = [H_\alpha, X'] = \alpha(H_\alpha)X' = 2X'$. So there exists eigenvalue 2 in V_α not coming from S_α. Hence there exists eigenvalue 0 of ad_{H_α} in V_α not coming from S_α. But this contradicts Claim 4, so g_α is one-dimensional. □

So part (b) follows.

Finally, since $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are both one-dimensional, $H_\alpha = [X_\alpha, Y_\alpha]$ is unique up to a scale. The scaling is fixed by the requirement that $[H_\alpha, X_\alpha] = 2X_\alpha$ (i.e., $\alpha(H_\alpha) = 2$). This proves uniqueness. □

Remark 22.1. So what is this good for? If $\mathfrak{g}$ is a complex semisimple Lie algebra, then the possible roots $\alpha \in \mathfrak{h}^*$ of $\mathfrak{g}$ are severely restricted, namely of the form

$$\mathfrak{g} = \mathfrak{h} \oplus \left(\bigoplus_{\alpha \in R} \mathfrak{g}_\alpha \right).$$

This gives hope that one can classify complex simple Lie algebras (reduce to classifying possible "root systems").

Now we will try to give a more "geometric picture" of the roots.

Proposition 22.2. *Let α, β be roots of $\mathfrak{g}$. Let H_α be the co-root associated to α. Then $\beta(H_\alpha) \in \mathbb{Z}$.*

Proof. Let $S_\alpha = \operatorname{span}\{X_\alpha, Y_\alpha, H_\alpha\}$ as before. If $0 \neq X_\beta \in \mathfrak{g}_\beta$ then $[H_\alpha, X_\beta] = \beta(H_\alpha)X_\beta$ so $\beta(H_\alpha)$ is an eigenvalue of $\operatorname{ad}_{H_\alpha}$ on V_α. Therefore $\beta(H_\alpha) \in \mathbb{Z}$. ☐

Recall that we use the positive-definite inner product $\langle \_, \_ \rangle$ on $\mathfrak{g}$ that is ad_A-invariant for all $A \in K$ restricted to $\mathfrak{h}$ to identify $\mathfrak{h}$ with $\mathfrak{h}^*$. So $\# : \mathfrak{h} \to \mathfrak{h}^*$ defined by $H \mapsto H^\#$ where $H^\#(J) = \langle H, J \rangle$. So for any $\alpha \in \mathfrak{h}^*$, there exists a unique $\alpha^\flat$ such that $(\alpha^\flat)^\# = \alpha$, i.e., $\langle \alpha^\flat, H \rangle = \alpha(H)$ for all $H \in \mathfrak{h}$. Using this identification, we can drop $\flat$ and $\#$ notation. That is, a root of $\mathfrak{g}$ is a non-zero element $\alpha \in \mathfrak{h}$ such that there exists a no-zero $X \in \mathfrak{g}$ with $[H, X] = \langle \alpha, H \rangle X$ for all $H \in \mathfrak{h}$. Note that $\langle \alpha, H \rangle = \alpha(H) = \langle \alpha^\flat, H \rangle$.

Let R be the set of roots of $\mathfrak{g}$ (as a finite subset of $\mathfrak{h}$ this time). We have so far shown that:

(1) each root $\alpha \in R$ is contained in $i\mathfrak{t} \subseteq \mathfrak{h} = \mathfrak{t} + i\mathfrak{t}$ (because if $H \in \mathfrak{t}$ then $\alpha(H)$ is purely imaginary).
(2) The roots span $\mathfrak{h}$ (the map $\#$ taking spanning sets to other spanning sets).
(3) if α is a root, so is $-\alpha$ but no other multiples.

Proposition 22.3. *Let α be a root and let H_α be its co-root. Then*

$$H_\alpha = \frac{2\alpha}{\|\alpha\|^2}, \quad \alpha = \frac{2H_\alpha}{\|H_\alpha\|^2}.$$

In particular, α and H_α are positive real multiples of each other, and $\|\alpha\|^2 \|H_\alpha\|^2 = 4$.

Proof. $\operatorname{span}\{H_\alpha\} = \ker(\alpha)^\perp \in \mathfrak{h}$, and $\ker(\alpha) = \{H \in \mathfrak{h} : \langle \alpha, H \rangle = 0\} = \operatorname{span}(\alpha)^\perp$. So $\operatorname{span}\{H_\alpha\} = \ker(\alpha)^\perp = (\operatorname{span}\{\alpha\})^{\perp\perp} = \operatorname{span}\{\alpha\}$. Therefore $\operatorname{span}\{\alpha\} = \operatorname{span}\{H_\alpha\}$. $\alpha = \lambda H_\alpha$ for some $\lambda \in \mathbb{C}$, so

$$\|\alpha\|^2 = \lambda \langle \alpha, H_\alpha \rangle = \lambda \alpha(H_\alpha) = 2\lambda,$$

so $\lambda = \|\alpha\|^2/2$. ☐

Corollary 22.4. $\langle \beta, H_\alpha \rangle = \beta(H_\alpha\rangle \in \mathbb{Z}$.

Proof. Since $H_\alpha = \frac{2\alpha}{\|\alpha\|^2}$. Hence $\frac{2\langle \beta, \alpha \rangle}{\|\alpha\|^2} \in \mathbb{Z}$ whenever $\alpha, \beta \in R$. Similarly we have $\frac{2\langle \alpha, \beta \rangle}{\|\beta\|^2} \in \mathbb{Z}$. Also,

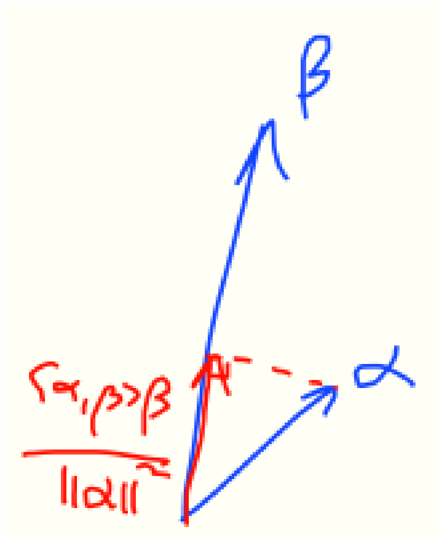
$$\beta = \frac{2H_\beta}{\|H_\beta\|^2},$$

so $\beta(H_\alpha) = \langle \beta, H_\alpha \rangle = \frac{2\langle H_\beta, H_\alpha \rangle}{\|H_\beta\|^2} \in \mathbb{Z}$. ☐

Remark 22.2 (Summary). If $\alpha, \beta \in R$ then

$$\frac{2\langle \alpha, \beta \rangle}{\|\alpha\|^2} = \frac{2\langle H_\alpha, H_\beta \rangle}{\|H_\beta\|^2} \in \mathbb{Z},$$

and both equal $\beta(H_\alpha)$, the eigenvalue of ad_{H_α} on $\mathfrak{g}_\beta$. Geometrically speaking, if $\alpha, \beta \in R$ then the orthogonal projection of α onto β is an integer or half integer multiple of β and vice versa.



23. March 25: The Weyl group

Recall that K is the compact (but not necessarily simply connected) analytic subgroup of $\mathrm{GL}(n, \mathbb{C})$ with Lie algebra $\mathfrak{k}$, which is a compact real form of $\mathfrak{g} \subseteq \mathfrak{gl}(n, \mathbb{C})$. Let $\mathfrak{t}$ be the maximal abelian subalgebra of $\mathfrak{k}$. Then $\mathfrak{h} = \mathfrak{t}_\mathbb{C} = \mathfrak{t} + i\mathfrak{t}$ is a Cartan subalgebra of $\mathfrak{g}$.

Definition 23.1. Define

$$Z(\mathfrak{t}) = \{A \in K : \mathrm{Ad}_A(H) = H \text{ for all } H \in \mathfrak{t}\}$$
$$N(\mathfrak{t}) = \{A \in K : \mathrm{Ad}_A(H) \in \mathfrak{t} \text{ for all } H \in \mathfrak{t}\}.$$

It is easy to check that both $Z(\mathfrak{t}), N(\mathfrak{t})$ are subgroups of K and $Z(\mathfrak{t})$ is a normal subgroup of $N(\mathfrak{t})$.

Definition 23.2. The *Weyl group of* $\mathfrak{g}$ is $W := N(\mathfrak{t})/Z(\mathfrak{t})$.

Define an action of W on $\mathfrak{t}$ by $[A] = w \in W, A \in N(\mathfrak{t})$ as follows: $w \cdot H = \mathrm{Ad}_A(H)$ for all $H \in \mathfrak{t}$. Since $\mathfrak{h} = \mathfrak{t}_\mathbb{C}$, this action extends (complex-linearly) to an action on $\mathfrak{h}$. (That is, W consists of invertible complex-linear operators on $\mathfrak{h}$).

Proposition 23.3. *The following hold regarding the Weyl group W:*

(1) The inner product $\langle \_, \_ \rangle$ on $\mathfrak{h}$ is invariant under the action of W.
(2) The set $R \subseteq \mathfrak{h}$ of roots is invariant under W.
(3) The set of co-roots is invariant under W, and $w \cdot H_\alpha = H_{w \cdot \alpha}$.
(4) W is a finite *group.*

Proof. The first statement is immediate from the fact that $\langle \_, \_ \rangle$ is Ad_A-invariant fro all $A \in K$. As for the second statement, let $w \in W$ and $w = [A]$. Let $\alpha \in R$. Then there exists $X \neq 0$ in $\mathfrak{g}$ such that $[H, X] = \langle \alpha, H \rangle X$ for all $H \in \mathfrak{h}$, and $\mathrm{Ad}_A(X) \neq 0$ in $\mathfrak{g}$. Note that $[H, \mathrm{Ad}_A X] = \mathrm{Ad}_A([\mathrm{Ad}_A^{-1} H, X])$, since Ad_A is a Lie algebra automorphism; and $\mathrm{Ad}_{A^{-1}} H \in \mathfrak{h}$ because $A^{-1} \in N(\mathfrak{t})$. Hence

$$[H, \mathrm{Ad}_A X] = \mathrm{Ad}_A(\langle \alpha, \mathrm{Ad}_{A^{-1}} H \rangle X) = \langle \mathrm{Ad}_A \alpha, H \rangle \mathrm{Ad}_A X = \langle w \cdot \alpha, H \rangle \mathrm{Ad}_A X$$

for all $H \in \mathfrak{h}$. Hence, $w \cdot \alpha$ is a root, with root vector $\mathrm{Ad}_A X$.

As for the third part, write $H_\alpha = 2\alpha/\|\alpha\|^2$. Then $w \cdot H_\alpha = 2(w \cdot \alpha)/\|\alpha\|^2 = 2(w \cdot \alpha)/\|w \cdot \alpha\|^2 = H_{w \cdot \alpha}$, with the second-to-last inequality following from the first part of the proposition.

For the last part, we begin by noting that the action of $w \in W$ is completely determined by what it does to the roots, but we know from the second part that w ail permute the roots. So W is a subgroup of the permutation group of the finite set R of roots. □

Example 23.4. Let $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$. Then $\mathfrak{k} = \mathfrak{su}(2)$, the 2×2 traceless Skew-Hermitian matrices. Then $K = \mathrm{SU}(2) \subseteq \mathrm{GL}(2, \mathbb{C})$. Then

$$\mathfrak{t} = \{ \begin{bmatrix} ia & 0 \\ 0 & -ia \end{bmatrix} : a \in \mathbb{R} \}$$

is a maximal abelian subalgebra. We claim that

$$Z(\mathfrak{t}) = \left\{ \begin{bmatrix} e^{ia} & 0 \\ 0 & e^{-ia} \end{bmatrix} : a \in \mathbb{R} \right\}$$

$$N(\mathfrak{t}) = Z(\mathfrak{t}) \sqcup \left\{ \begin{bmatrix} 0 & e^{ia} \\ -e^{-ia} & 0 \end{bmatrix} : a \in \mathbb{R} \right\}$$

Proof. Take

$$H_0 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \in \mathfrak{t}.$$

If $A \in Z(\mathfrak{t})$ then $\mathrm{Ad}_A H_0 = H_0$ iff $AH_0 = H_0 A$. Note that if $H_0 v = \lambda v$ then $AH_0 v = H_0 A v = \lambda A v$. Thus A preserves eigenspaces of H_0. Suppose that eigenspaces are spanned by $\{e_1\}$ and $\{e_2\}$ respectively. Hence for some $c, d \in \mathbb{C}$ we have $Ae_1 = ce_1$ and $Ae_2 = de_2$. Therefore

$$A = \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} \in \mathrm{SU}(2),$$

so

$$A = \begin{bmatrix} e^{ia} & 0 \\ 0 & e^{-ia} \end{bmatrix}.$$

Conversely, it's trivial to check that anything of the form $\begin{bmatrix} e^{ia} & 0 \\ 0 & e^{-ia} \end{bmatrix}$ is in $Z(\mathfrak{t})$.

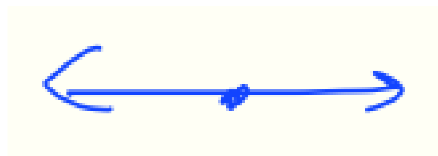
If $A \in N(\mathfrak{t})$, then $AH_0A^{-1} \in \mathfrak{t}$, so it is diagonal, and AH_0A^{-1} has eigenvectors Ae_1, Ae_2 with eigenvalues $\pm i$ respectively, i.e., $AH_0A^{-1}(Ae_i) = AH_0 e_i = \lambda_i A e_i$ where $\lambda_1 = i$ and $\lambda_2 = -i$. But every element on $\mathfrak{t}$ has eigenvectors e_1 and e_2 since it is diagonal. So there exist c and d such that $Ae_1 = ce_1, Ae_2 = de_2$ or $Ae_1 = de_1, Ae_2 = ce_2$. The former is the $Z(\mathfrak{t})$ case, while the latter is the $\begin{bmatrix} 0 & e^{ia} \\ -e^{-ia} & 0 \end{bmatrix}$ case. Conversely, anything of the form $\begin{bmatrix} 0 & e^{ia} \\ -e^{-ia} & 0 \end{bmatrix}$ is in $N(\mathfrak{t})$.

Then $W = N(\mathfrak{t})/Z(\mathfrak{t})$. Let $w = [A]$, $A \in N(\mathfrak{t})$. Note that $w \cdot H = \operatorname{Ad}_A H$. If $A \in Z(\mathfrak{t})$, then w is the identity on $\mathfrak{h}$ so $\operatorname{Ad}_A H = H$ for all $H \in \mathfrak{h}$. On the other hand, if $A \in N(\mathfrak{t}) \setminus Z(\mathfrak{t})$, then

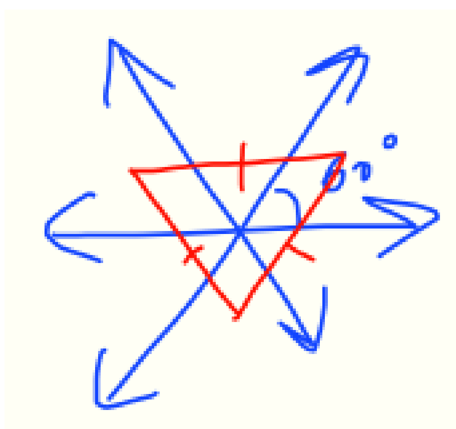
$$\operatorname{Ad}_A \begin{bmatrix} ib & 0 \\ 0 & -ib \end{bmatrix} = \begin{bmatrix} 0 & e^{ia} \\ -e^{-ia} & 0 \end{bmatrix} \begin{bmatrix} ib & 0 \\ 0 & -ib \end{bmatrix} \begin{bmatrix} 0 & -e^{ia} \\ e^{-ia} & 0 \end{bmatrix} = \begin{bmatrix} -ib & 0 \\ 0 & ib \end{bmatrix}.$$

So $w = [A]$ acts as $-\operatorname{id}$ on $\mathfrak{h}$. So the Weyl group of $\mathfrak{sl}(2, \mathbb{C}) \cong \mathbb{Z}/2\mathbb{Z} = \{\pm \operatorname{id}\}$. □

Recall that the root diagram of $\mathfrak{sl}(2, \mathbb{C})$ is



so the Weyl group is the permutation group of R. But in the case of $\mathfrak{sl}(3, \mathbb{C})$, then Weyl group is isomorphic to S_3 (symmetries of equilateral triangles) hence Weyl group is a proper subset of the permutation group of R.

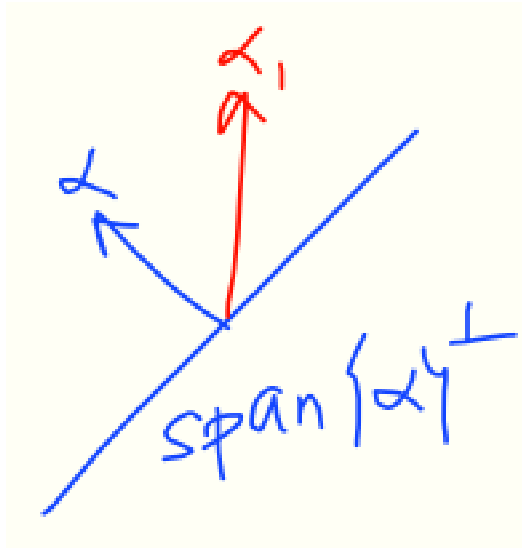


Going back to the general complex semisimple algebra, the following theorem holds:

Theorem 23.5. *For every root $\alpha \in R$, there exists an element $w_\alpha \in W$ such that*

$$w_\alpha \cdot \alpha = -\alpha$$
$$w_\alpha \cdot H = H$$

for all $H \in \mathfrak{h}$ such that $\langle H, \alpha \rangle = 0$. That is, for every root α, there exists an element $w_\alpha \in W$ that is reflection across the codimension one hyperplane $(\operatorname{span}\{\alpha\})^\perp$.



Remark 23.1. If $\beta \in \mathfrak{h}$, then

$$\beta = \frac{\langle \alpha, \beta \rangle \alpha}{\|\alpha\|^2} + \underbrace{\left(\beta - \frac{\langle \alpha, \beta \rangle \alpha}{\|\alpha\|^2} \right)}_{\in \operatorname{span}\{\alpha\}^\perp}$$

So

$$\begin{aligned} w_\alpha \cdot \beta &= -\frac{\langle \alpha, \beta \rangle \alpha}{\|\alpha\|^2} + \left(\beta - \frac{\langle \alpha, \beta \rangle \alpha}{\|\alpha\|^2} \right) \\ &= \beta - \frac{2\langle \alpha, \beta \rangle \alpha}{\|\alpha\|^2} \\ \beta - w_\alpha \cdot \beta &= 2\frac{\langle \alpha, \beta \rangle \alpha}{\|\alpha\|^2}. \end{aligned}$$

Hence if $\alpha, \beta \in R$ then $\beta - w_\alpha \cdot \beta$ is an integer multiple of α.

Proof. Let $X_\alpha, Y_\alpha, H_\alpha$ be as before, where $Y_\alpha = -\overline{X_\alpha}$. We have $X_\alpha = X_1 + iX_2$ with $X_i \in \mathfrak{k}$. So $Y_\alpha = -X_1 + iX_2$, so $X_\alpha - Y_\alpha = 2X_1 \in \mathfrak{k}$. Let $A_\alpha := \exp\left(\frac{\pi}{2}(X_\alpha - Y_\alpha)\right) \in K$. If $H \in \mathfrak{h}$ with $\langle \alpha, H \rangle = 0$ then $[H, X_\alpha] = \langle \alpha, H \rangle X_\alpha = 0$ and $[H, Y_\alpha] = -\langle \alpha, H \rangle Y_\alpha = 0$. So $Y_\alpha \in \mathfrak{g}_{-\alpha}$. So X_α, Y_α commute with any $H \in (\operatorname{span}\{\alpha\})^\perp$.

Claim. $\operatorname{Ad}_{A_\alpha}(H) = H$ for all $H \in (\operatorname{span}\{\alpha\})^\perp$ and $\operatorname{Ad}_{A_\alpha}(\alpha) = -\alpha$.

Proof of Claim. Note that

$$\begin{aligned} \operatorname{Ad}_{A_\alpha} = \operatorname{Ad}_{\exp\left(\frac{\pi}{2}(X_\alpha - Y_\alpha)\right)} &= \exp\left(\operatorname{ad}\left(\frac{\pi}{2}(X_\alpha - Y_\alpha)\right)\right) \\ &= \exp\left(\frac{\pi}{2}\left(\operatorname{ad}_{X_\alpha} - \operatorname{ad}_{Y_\alpha}\right)\right) \end{aligned}$$

if $\langle \alpha, H \rangle = 0$, then $\operatorname{ad}_{X_\alpha} H = \operatorname{ad}_{Y_\alpha} H = 0$. So

$$\operatorname{Ad}_{A_\alpha} H = \exp\left(\frac{\pi}{2}(\operatorname{ad}_{X_\alpha} - \operatorname{ad}_{Y_\alpha})\right) H = H,$$

since the only constant term in power series contributes.

If $H = \alpha\ H_\alpha$ is a positive multiple of α, then

$$\mathrm{Ad}_{A_\alpha} H_\alpha = \exp\left(\frac{\pi}{2}(\mathrm{ad}_{X_\alpha} - \mathrm{ad}_{Y_\alpha})\right) H_\alpha,$$

and the RHS only involves brackets of $X_\alpha, Y_\alpha, H_\alpha$ so only depends on commutative relations $\cong \mathfrak{sl}(2,\mathbb{C})$ with basis X, Y, H. In $\mathfrak{sl}(2,\mathbb{C})$, where

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

we have

$$\frac{\pi}{2}(X - Y) = \begin{bmatrix} 0 & \frac{\pi}{2} \\ -\frac{\pi}{2} & 0 \end{bmatrix},$$

so

$$\exp\left(\frac{\pi}{2}(X - Y)\right) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in N(\mathfrak{t}) \setminus Z(\mathfrak{t}).$$

Therefore $\mathrm{Ad}_{\exp\left(\frac{\pi}{2}(X-Y)\right)} H_\alpha = -H_\alpha = \mathrm{Ad}_{\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]} H_\alpha$, as desired. □

This completes the proof of our original theorem. □

Remark 23.2. In fact, one can prove that the Weyl group W is the subgroup generated by such reflections $\{w_\alpha : \alpha \in R\}$. So this fact imposes more constraints on what sets R can be roots of a complex semisimple Lie algebra. Hence it is useful for classification.

We have almost all the language to state the theorem of the highest weight for a complex semisimple Lie algebra $\mathfrak{g}$.

Theorem 23.6. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra, and assume that every representation is complex and finite-dimensional. Then:*

1. *Every irreducible representation of $\mathfrak{g}$ has a (unique) highest weight.*
2. *Two irreducible representations of $\mathfrak{g}$ are isomorphic if and only if they have the same highest weight.*
3. *The highest weight of an irreducible representation is a dominant integral element.*
4. *Every dominant integral element* is *the highest weight of an irreducible representation.*

We first need to define weight, highest weight, integral element, and dominant integral element.

Definition 23.7. A *base* for R is a subset $\{\alpha_1, \dots, \alpha_r\}$ of R such that:

1. it is a basis of $\mathfrak{h}$ ($r = \dim(\mathfrak{h}) = \mathrm{rank}(\mathfrak{g})$)
2. every $\alpha \in R$ can be written as $\alpha = c_1\alpha_1 + \cdots + c_r\alpha_r$ where each $c_j \in \mathbb{Z}$, and all c_j's are either non-negative or non-positive.

Also, if all the c_j's above are all non-negative, then these roots are called *positive with respect to the given basis*; if c_j's are all non-positive, then these roots are called *negative with respect to the same basis*. Furthermore, $\{\alpha_1, \dots, \alpha_r\}$ are called *positive simple roots*.

Remark 23.3. A base *always* exists.

Definition 23.8. An element $\mu \in \mathfrak{h}$ is an *integral element* if $\langle \mu, H_\alpha \rangle \in \mathbb{Z}$ for all $\alpha \in R$ and H_α a co-root (equivalent to saying that $\frac{2\langle \mu, \alpha \rangle}{\|\alpha\|^2} \in \mathbb{Z}$ for all $\alpha \in R$. A *dominant integral element* is an integral element $\mu \in \mathfrak{h}$ such that $\langle \mu, H_\alpha \rangle \geq 0$ for every positive simple root α.

Definition 23.9. Let (V, π) be a finite-dimensional complex representation of $\mathfrak{g}$. An element $\mu \in \mathfrak{h}$ is a *weight* of π if there exists $v \neq 0$ in V such that $\pi(H)v = \langle \mu, H \rangle v$ for all $H \in \mathfrak{h}$.

Definition 23.10. If $\mu_1, \mu_2 \in \mathfrak{h}$, then we say that μ_1 is *higher than* μ_2 if $\mu_1 - \mu_2 = c_1\alpha_1 + \cdots + c_r\alpha_r$ with all $c_i \geq 0$ (real numbers). If this is the case, then we write $\mu_1 \succeq \mu_2$.

As for the proof of the theorem of the highest weight, the proofs for the first three statements are identical to the $\mathfrak{sl}(3, \mathbb{C})$ case. The fourth one is much harder. There are three possible approaches: Verma modules (Lie algebra approach), Peter-Weyl theorem (compact group approach), and Borel-Weil construction (complex group approach). Theorem means that irreducible representations (up to isomorphism) are "parametrized" (bijective correspondence) with dominant integral elements.

24. March 30: Construction of Spin(n)

Today, we will explicitly construct a matrix Lie group $\mathrm{Spin}(n)$ for all $n \geq 3$, and a two-to-one homomorphism $\pi : \mathrm{Spin}(n) \to \mathrm{SO}(n)$ such that $\mathrm{Spin}(n)$ is compact, simply connected, and $\mathrm{Lie}(\mathrm{Spin}(n)) \cong \mathfrak{so}(n) = \mathrm{Lie}(\mathrm{SO}(n))$. This is a generalization of $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{SO}(3)$. Particularly, we will see that $\mathrm{SU}(2) \cong \mathrm{Spin}(3)$. This is all more general than what we are going to do today.

Consider $V = (\mathbb{R}^n, \langle \_, \_ \rangle)$ with the standard Euclidean positive-definite inner product. Let $\{e_1, \ldots, e_n\}$ be the standard basis of V. This is orthonormal.

Definition 24.1. We define $\mathrm{Cl}(V)$ to be the *Clifford algebra* of $V = (\mathbb{R}^n, \langle \_, \_ \rangle)$, the associative algebra over $\mathbb{R}$ with identity generated by $\{e_1, \ldots, e_n\}$ subject to the relations

- $e_i \cdot e_j = -e_j \cdot e_i$ if $i \neq j$
- $e_i^2 = e_i \cdot e_i = -1$ for all $i = 1, 2, \ldots, n$.

Formally speaking, the Clifford algebra is

$$\mathrm{Cl}(V) = \left(\bigoplus_{k=0}^{\infty} (V^{\otimes k}) \right) / I,$$

where I is the two-sided ideal generated by $\{v \otimes v + \langle v, v \rangle 1 : v \in V\}$.

Remark 24.1. From the two relations defined above, it follows that $v \cdot w + w \cdot v = -\langle v, w \rangle 1$ for all $v, w \in V$.

Basis for $\mathrm{Cl}(V)$ is all the elements that have one of the following forms 1, e_i, $e_i \cdot e_j$ $(i < j)$, $e_i \cdot e_j \cdot e_k$ $(i < j < k), \ldots, e_{i_1} \cdot \cdots \cdot e_{i_l}$ $(i_1 < i_2 < \cdots < i_l), \ldots, e_1 \cdot e_2 \cdot \cdots \cdot e_n$. Thus,

$$\dim(\mathrm{Cl}(V)) = \sum_{l=0}^{n} \binom{n}{l} = (1+1)^n = 2^n.$$

Remark 24.2. As $\mathbb{R}$-vector spaces, we have

$$\mathrm{Cl}(V) \cong \Lambda^{\bullet}(V) = \bigoplus_{k=0}^{n} \Lambda^k(V),$$

but *not as algebras*.

Remark 24.3. $V \subseteq \mathrm{Cl}(V)$ as a subspace, and similarly $\mathbb{R} \subseteq \mathrm{Cl}(V)$ as a subspace, where $\mathbb{R} = \mathrm{span}\{1\}$.

Definition 24.2. A *Clifford automorphism* $P \in \mathrm{Aut}(\mathrm{Cl}(V))$ is an automorphism of the underlying real algebra that also maps V into V satisfying

- $P(a+b) = P(a) + P(b)$
- $P(ta) = tP(a)$ for all $t \in \mathbb{R}$
- $P(a \cdot b) = P(a) \cdot P(b)$
- $P(v) \in V \subseteq \mathrm{Cl}(V)$ for all $v \in V$.

Remark 24.4. Note that the first three indicate that P is an algebra homomorphism. So if P is an algebra homomorphism satisfying the fourth condition, then P becomes a Clifford automorphism.

Theorem 24.3. $\mathrm{Aut}(\mathrm{Cl}(V)) = \mathrm{O}(n) = \mathrm{O}(n, \mathbb{R})$.

Proof. Let $P \in \mathrm{Aut}(\mathrm{Cl}(V))$ a Clifford automorphism. Let $v \in V$. Then we have $P(v) \in V$. By the fundamental defining relation ($v \cdot v = -|v|^2 1$) we have

$$\begin{aligned}\langle P(v), P(v)\rangle 1 &= -P(v) \cdot P(v) = -P(v, v) = P(-v \cdot v) \\ &= P(\langle v, v\rangle 1) = \langle v, v\rangle P(1) = \langle v, v\rangle 1.\end{aligned}$$

Thus $\langle P(v), P(v)\rangle = \langle v, v\rangle$ for all $v \in V$ so $P|_V \in \mathrm{O}(n)$.

Now suppose $P \in \mathrm{O}(n)$. Extend P to $\mathrm{Cl}(V)$ in the obvious way to make it an algebra homomorphism. Then one can easily verify that this is well-defined and is an algebra automorphism. ☐

Definition 24.4. A *Clifford anti-automorphism* is an algebra *anti-automorphism*, i.e., $P(ab) = P(b) \cdot P(a)$ for all $a, b \in \mathrm{Cl}(V)$ such that $P(v) \in V$ for all $v \in V$.

Remark 24.5. $\mathrm{Cl}(V)$ has a canonical Clifford automorphism and two Clifford anti-automorphisms.

Definition 24.5. The *isometry* $v \mapsto -v$ of V extend to an automorphism of $\mathrm{Cl}(V)$, i.e., $P = -I \in \mathrm{O}(n)$. Denote this automorphism by $\sim\colon \mathrm{Cl}(V) \to \mathrm{Cl}(V)$.

Remark 24.6. $\widetilde{\widetilde{\alpha}} = \alpha$ for all $\alpha \in \mathrm{Cl}(V)$ since $(-I)^2 = I$. Therefore $\sim$ is an involution.

Definition 24.6. We are now ready to define *even Clifford algebras* and *odd Clifford algebras*:

$$\begin{aligned}\mathrm{Cl}(V)^{\mathrm{even}} &= \{\alpha \in \mathrm{Cl}(V) : \widetilde{\alpha} = \alpha\} \\ \mathrm{Cl}(V)^{\mathrm{odd}} &= \{\alpha \in \mathrm{Cl}(V) : \widetilde{\alpha} = -\alpha\}.\end{aligned}$$

Remark 24.7. Both are subspaces of $\mathrm{Cl}(V)$, and in particular $\mathrm{Cl}(V)^{\mathrm{even}}$ is a *subalgebra*. Also, since

$$\alpha = \frac{\alpha + \widetilde{\alpha}}{2} + \frac{\alpha - \widetilde{\alpha}}{2},$$

and $(\alpha + \widetilde{\alpha})/2 \in \mathrm{Cl}(V)^{\mathrm{even}}$ and $(\alpha - \widetilde{\alpha})/2 \in \mathrm{Cl}(V)^{\mathrm{odd}}$, we have $\mathrm{Cl}(V) = \mathrm{Cl}(V)^{\mathrm{even}} \oplus \mathrm{Cl}(V)^{\mathrm{odd}}$ as *real vector spaces*.

Define an anti-automorphism $\vee : \mathrm{Cl}(V) \to \mathrm{Cl}(V)$ by defining it on a basis (let $i_1 < i_2 < \cdots < i_k$):

$$\begin{aligned}(e_{i_1} \cdot e_{i_2} \cdot \cdots \cdot e_{i_k})^\vee &= e_{i_k} \cdot e_{i_{k-1}} \cdot \cdots \cdot e_{i_1} \\ &= (-1)^{(k-1)+(k-2)+\cdots+1} e_{i_1} \cdot \cdots \cdot e_{i_k} = (-1)^{k(k-1)/2} e_{i_1} \cdot \cdots \cdot e_{i_k}\end{aligned}$$

(i.e., reverse order of factors). Note that $(\alpha \cdot \beta)^\vee = \beta^\vee \cdot \alpha^\vee$, and $(\alpha^\vee)^\vee = \alpha$ so $\vee$ is an involution. Furthermore, notice that, if $k = 4m + l$ with $l = 0, 1, 2, 3$ then

$$\frac{k(k-1)}{2} \equiv \frac{l(l-1)}{2} \pmod 4,$$

so if $k \equiv 0, 1 \pmod 4$ then $(-1)^{k(k-1)/2}$ is 1 while $(-1)^{k(k-1)/2}$ is negative if $k \equiv 2, 3 \pmod 4$.

Remark 24.8. $\vee\circ \sim \circ\vee$ is an automorphism of $\mathrm{Cl}(V)$ that maps $V \hookrightarrow V$, since $(\vee \sim \vee)(e_i) = \vee(\sim(e_i)) = \vee(-e_i) = -e_i = \sim e_i$. Therefore $\vee \sim \vee = \sim$ on V. Hence $\vee \sim \vee = \sim$ on $\mathrm{Cl}(V)$. Hence $\sim \vee = \vee \sim$ (i.e., they commute).

Definition 24.7. $\wedge : \mathrm{Cl}(V) \to \mathrm{Cl}(V)$ is said to be the Clifford anti-automorphism $\wedge := \sim \vee = \vee \sim$.

For the record, we will record the signs of each map:

$k \bmod 4$	0	1	2	3
$\sim$	+	−	+	−
$\vee$	+	+	−	−
$\wedge$	+	−	−	+

24.1. *Main symmetry of Clifford algebras*

Proposition 24.8. $\mathrm{Cl}(\mathbb{R}^n) \cong (\mathrm{Cl}(\mathbb{R}^{n+1}))^{\text{even}}$.

Proof. Let $e_1, \dots, e_n$ be the standard basis of $\mathbb{R}^n$ and $E_0, \dots, E_n$ the standard basis of $\mathbb{R}^{n+1}$. Define a map $\phi : \mathrm{Cl}(\mathbb{R}^n) \to \mathrm{Cl}(\mathbb{R}^{n+1})$ such that $\phi(e_i) = E_0 \cdot E_i \in (\mathrm{Cl}(\mathbb{R}^{n+1}))^{\text{even}}$ extend to make ϕ to make an algebra homomorphism:

$$\begin{aligned}\phi(e_i)\phi(e_j) &= E_0E_iE_0E_j = -E_0E_0E_iE_j \\ &= -E_0^2E_iE_j = E_iE_j = -E_jE_i = -\phi(e_j)\phi(e_i).\end{aligned}$$

for all $i \neq j$. If $i = j$, then $\phi(e_i)\phi(e_i) = E_i^2 = -1 = \phi(e_i^2)$. Therefore $\phi : \mathrm{Cl}(\mathbb{R}^n) \to (\mathrm{Cl}(\mathbb{R}^{n+1}))^{\text{even}}$, and note that both have dimension 2^n, hence is an isomorphism of real algebras (check the details) that preserves the Clifford relation. □

Example 24.9. $\mathrm{Cl}(\mathbb{R}^0) = \mathbb{R}^0 = \mathbb{R}$, and has dimension $2^0 = 1$, and $\sim, \vee, \wedge$ all identity.

Example 24.10. If $r = 1$, then $\mathrm{Cl}(\mathbb{R}^1) = \mathrm{span}_\mathbb{R}\{1\} \oplus \mathrm{span}_\mathbb{R}\{e_1\} \cong \mathbb{R}^2$ as real vector spaces. And every $\alpha \in \mathrm{Cl}(\mathbb{R}^1)$ can be written in the form $\alpha = a1 + be_1$ with $a, b \in \mathbb{R}$, where 1 is the multiplicative identity and $e_1^2 = -1$ by defining relations. Therefore $\mathrm{Cl}(\mathbb{R}^1) \cong \mathbb{C}$ as $\mathbb{R}$- algebras. And $\sim = \vee$ is the map $a1 + be_1 \mapsto a1 - be_1$, which is complex conjugate. Hence $\mathrm{Cl}(\mathbb{R}^1)^{\text{even}} = \mathrm{span}\{1\} = \mathbb{R}$ and $\mathrm{Cl}(\mathbb{R}^1)^{\text{odd}} = \mathrm{span}\{e_1\} \cong i\mathbb{R}$. It thus follows that $\mathrm{Cl}(\mathbb{R}^0) \cong \mathrm{Cl}(\mathbb{R}^1)^{\text{even}}$ and $\mathbb{R} = \mathfrak{Re}(\mathbb{C}) = \mathbb{C}^{\text{even}}$.

Example 24.11. What if $r = 2$? Note that

$$\mathrm{Cl}(\mathbb{R}^2) = \mathrm{span}\{1\} \oplus \mathrm{span}\{e_1, e_2\} \oplus \mathrm{span}\{e_1 \cdot e_2\}.$$

Then $\mathrm{span}\{e_1, e_2\} \cong V$ and $\mathrm{span}\{1\} \cong \mathbb{R}$. Recall that $e_1^2 = e_2^2 = -1$ and $e_1 \cdot e_2 = -e_2 \cdot e_1$. Notice that $(e_1 \cdot e_2) \cdot e_1 = -e_2 \cdot e_1 \cdot e_1 = e_2$, and similarly, $e_1 \cdot (e_1 \cdot e_2) = -e_2, e_2 \cdot (e_1 \cdot e_2) = -e_1 \cdot e_2 \cdot e_2 = e_1$, and $(e_1 \cdot e_2) \cdot e_2 = -e_1$. Letting $i = e_1 \cdot e_2, j = e_1, k = e_2$ we see that

$i^2 = j^2 = k^2 = -1, ij = k = -ji, jk = i = -kj, ki = j = -ik$. Therefore $\mathrm{Cl}(\mathbb{R}^2) \cong \mathbb{R}^{2^2} \cong \mathbb{R}^4$ as real vector spaces, and (as $\mathbb{R}$-algebras) $\mathrm{Cl}(\mathbb{R}^2) \cong \mathbb{H}$, the real algebra of quaternions. Then

$$(\mathrm{Cl}(\mathbb{R}^2))^{\text{even}} = \mathrm{span}\{1, e_1 \cdot e_2\} = \mathrm{span}\{1, i\} \cong \mathbb{C}$$
$$(\mathrm{Cl}(\mathbb{R}^2))^{\text{odd}} = \mathrm{span}\{e_1, e_2\} = \mathrm{span}\{j, k\},$$

so $\mathrm{Cl}(\mathbb{R}^2)^{\text{odd}}$ is a real subspace but not a subalgebra; and $\mathbb{H}$ has a real subalgebra isomorphic to $\mathbb{C}$. Hence $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}j$. The anti-automorphism is defined as $\wedge(1) = 1, \wedge(\pm i) = \mp i, \wedge(\pm j) = \mp j, \wedge(\pm k) = \mp k$. Therefore $\wedge$ corresponds to the quaternionic conjugation (recall that $\mathbb{H} = \mathbb{R} \oplus \mathfrak{Im}(\mathbb{H})$).

Note that $\mathbb{H}$ is a division ring (i.e., every element has a multiplicative inverse – but without commutativity). If $q \in \mathbb{H}$ is non-zero, then $q^{-1} = \bar{q}/|q|^2$ because $\bar{q}q = q\bar{q} = |q|^2$ (analogous to $z\bar{z} = |z|^2$ in $\mathbb{C}$).

Example 24.12. If $r = 3$, then $\mathrm{Cl}(\mathbb{R}^3) \cong \mathbb{R}^8$ as real vector spaces, but is *not the octonions* $\mathbb{O}$, since $\mathbb{O}$ is *not associative*. As real algebras, $\mathrm{Cl}(\mathbb{R}^3) \cong \mathbb{H} \oplus \mathbb{H}$, where even: (p, p) and odd: $(p, -p)$, where $(p_1, q_1)(p_2, q_2) = (p_1p_2, q_1q_2)$ with the multiplicative identity $1 = (1, 1)$.

Example 24.13. We can keep going, and we will see that:
- $\mathrm{Cl}(\mathbb{R}^4) = \mathrm{M}_{2\times 2}(\mathbb{H})$, which is 16-dimensional over reals.
- $\mathrm{Cl}(\mathbb{R}^5) = \mathrm{M}_{4\times 4}(\mathbb{C})$ (32-dimensional over reals)
- $\mathrm{Cl}(\mathbb{R}^6) = \mathrm{M}_{8\times 8}(\mathbb{R})$ (64-dimensional over $\mathbb{R}$)
- $\mathrm{Cl}(\mathbb{R}^7) = \mathrm{M}_{8\times 8}(\mathbb{R}) \oplus \mathrm{M}_{8\times 8}(\mathbb{R})$ (128-dimensional)
- $\mathrm{Cl}(\mathbb{R}^8) = \mathrm{M}_{16\times 16}(\mathbb{R})$ (256-dimensional)

But we can stop here actually, because the pattern will repeat. In other words,

n mod 8	$\mathrm{Cl}(\mathbb{R}^n)$	dimension $= 2^n$
0	$\mathrm{M}_{N\times N}(\mathbb{R})$	N^2
1	$\mathrm{M}_{N\times N}(\mathbb{C})$	$2N^2$
2	$\mathrm{M}_{N\times N}(\mathbb{H})$	$4N^2$
3	$\mathrm{M}_{N\times N}(\mathbb{H}) \oplus \mathrm{M}_{N\times N}(\mathbb{H})$	$8N^2$
4	$\mathrm{M}_{N\times N}(\mathbb{H})$	$4N^2$
5	$\mathrm{M}_{N\times N}(\mathbb{C})$	$2N^2$
6	$\mathrm{M}_{N\times N}(\mathbb{R})$	N^2
7	$\mathrm{M}_{N\times N}(\mathbb{R}) \otimes \mathrm{M}_{N\times N}(\mathbb{R})$	$2N^2$

24.2. *The Clifford centre and the twisted centre*

Definition 24.14. We define the *Clifford centre* $\mathrm{cent}(\mathrm{Cl}(V))$ and the *twisted centre* $\mathrm{twcent}(\mathrm{Cl}(V))$ as follows:

$$\begin{aligned}\mathrm{cent}(\mathrm{Cl}(V)) &= \{\alpha \in \mathrm{Cl}(V) : \alpha \cdot \beta = \beta \cdot \alpha \text{ for all } \beta \in \mathrm{Cl}(V)\}\\ &= \{\alpha \in \mathrm{Cl}(V) : \alpha \cdot v = v \cdot \alpha \text{ for all } v \in V\}\\ \mathrm{twcent}(\mathrm{Cl}(V)) &= \{\alpha \in \mathrm{Cl}(V) : \alpha \cdot \beta = -\beta \cdot \alpha \text{ for all} \beta \in \mathrm{Cl}(V)\}\\ &= \{\alpha \in \mathrm{Cl}(V) : \alpha \cdot v = -v \cdot \alpha \text{ for all } v \in V\}.\end{aligned}$$

Lemma 24.15. *Let $V = \mathbb{R}^n$.*

(1) If n is even, then $\mathrm{cent}(\mathrm{Cl}(V)) = \mathbb{R} = \mathrm{span}\{1\}$ *and* $\mathrm{twcent}(\mathrm{Cl}(V)) \cong \mathbb{R} = \mathrm{span}\{\mu\}$.

(2) If n is odd, then $\mathrm{cent}(\mathrm{Cl}(V)) = \mathrm{span}_n\{1, \mu\} \cong \mathbb{R}^2$ *and* $\mathrm{twcent}(\mathrm{Cl}(V)) = \{0\}$.

Therefore, $1 \in \text{cent}(\text{Cl}(V))$ *always. Also, only other basis vector that can be in* cent *or* twcent *is* μ*, where* $\mu \in \text{cent}$ *for* n *odd and* $\mu \in \text{twcent}$ *for* n *even.*

Proof. Let $e_i \in V$ be a basis vector. Let $\alpha \in \text{Cl}(V)$. Write $\alpha = \beta + e_i \cdot \gamma$, where β, γ do not involve e_i. Then $\alpha e_i = (\beta + e_i\gamma)e_i$, and $e_i\alpha = e_i(\beta + e_i\gamma) = e_i\beta - \gamma$. If α is even, then β is even and γ is odd. Hence $\alpha \cdot e_i = e_i\beta - e_i \cdot e_i \cdot \gamma = e_i\beta + \gamma$ so:

- $\alpha \cdot e_i = e_i \cdot \alpha \Leftrightarrow \gamma = 0 \Leftrightarrow \alpha$ *does not involve* e_i.
- $\alpha \cdot e_i = -e_i \cdot \alpha \Leftrightarrow \beta = 0 \Leftrightarrow \alpha$ *does involve* e_i.

If α is odd, then β is odd and γ is even. Since $\alpha \cdot e_i = -e_i\beta - \gamma$, so we have:

- $\alpha \cdot e_i = e_i \cdot \alpha \Leftrightarrow \beta = 0 \Leftrightarrow \alpha$ *does involve* e_i.
- $\alpha \cdot e_i = -e_i \cdot \alpha \Leftrightarrow \gamma = 0 \Leftrightarrow \alpha$ *does not involve* e_i.

Hence, we see that $\alpha \in \text{cent}(\text{Cl}(V))$ if and only if $\alpha^{\text{even}}, \alpha^{\text{odd}} \in \text{cent}$; similarly, $\alpha \in \text{twcent}(\text{Cl}(V))$ if and only if $\alpha^{\text{even}}, \alpha^{\text{odd}} \in \text{twcent}(\text{Cl}(V))$. □

25. April 1: Last lecture!

Recall that $\text{Cl}(\mathbb{R}^n) = \text{Cl}(\mathbb{R}^n)^{\text{even}} \oplus \text{Cl}(\mathbb{R}^n)^{\text{odd}}$, and that $\text{Cl}(\mathbb{R}^n)^{\text{even}}$ is a subalgebra and $\text{Cl}(\mathbb{R}^n)^{\text{odd}}$ is a subspace (but not a subalgebra). Last time, we defined the Clifford centre $\text{cent}(\text{Cl}(V))$ and the twisted centre $\text{twcent}(\text{Cl}(V))$:

$$\begin{aligned} \text{cent}(\text{Cl}(V)) &= \{\alpha \in \text{Cl}(V) : \alpha \cdot \beta = \beta \cdot \alpha \text{ for all } \beta \in \text{Cl}(V)\} \\ &= \{\alpha \in \text{Cl}(V) : \alpha \cdot v = v \cdot \alpha \text{ for all } v \in V\} \\ \text{twcent}(\text{Cl}(V)) &= \{\alpha \in \text{Cl}(V) : \alpha \cdot \beta = -\beta \cdot \alpha \text{ for all} \beta \in \text{Cl}(V)\} \\ &= \{\alpha \in \text{Cl}(V) : \alpha \cdot v = -v \cdot \alpha \text{ for all } v \in V\}. \end{aligned}$$

We also proved that if n is even, then cent $= \text{span}\{1\}$ and twcent $= \text{span}\{\mu\}$, where $\mu := e_1 \cdot e_2 \cdot \cdots \cdot e_n$. If n is odd, then $\text{cent}(\text{Cl}(V)) = \text{span}\{1, \mu\}$ and $\text{twcent}(\text{Cl}(V)) = \{0\}$.

25.1. *Self-duality and anti-self-duality*

We have $\mu = e_1 \cdot e_2 \cdot \cdots \cdot e_n \cdot e_1 \cdot e_2 \cdot \cdots \cdot e_n = (-1)^{n(n-1)/2} e_1 \cdot \cdots \cdot e_n \cdot e_n \cdot e_{n-1} \cdot \cdots \cdot e_1 = (-1)^{l(l+1)/2}$, where $n = 4k + l$. This gives us that

k mod 4	0	1	2	3
μ^2	1	-1	-1	1

Definition 25.1. Suppose that $n \equiv 0$ or 3 mod 4. Then $\alpha \in \text{Cl}(V)$ is called *self-dual* if $\mu \cdot \alpha = \alpha$ and *anti-self-dual* if $\mu \cdot \alpha = -\alpha$.

Remark 25.1. For any $\alpha \in \text{Cl}(V)$, we can write

$$\alpha = \frac{\alpha + \mu \cdot \alpha}{2} + \frac{\alpha - \mu \cdot \alpha}{2}.$$

Then $\frac{\alpha+\mu\cdot\alpha}{2}$ is self-dual and $\frac{\alpha-\mu\cdot\alpha}{2}$ is anti-self-dual. This gives us another splitting (distinct from the even/odd splitting):

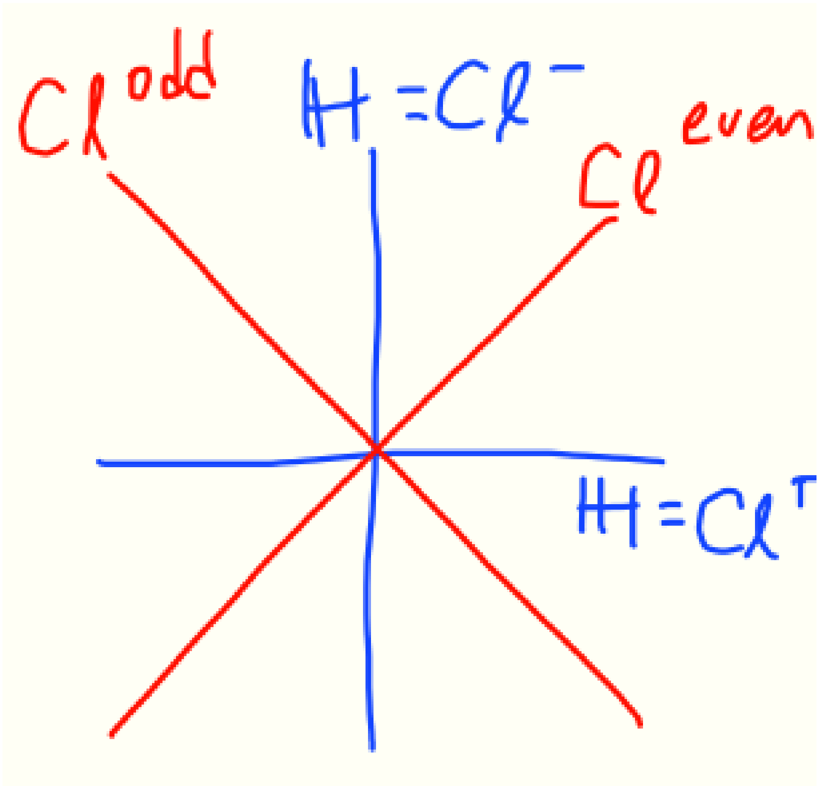
$$\text{Cl}(V) = \text{Cl}^+(V) \oplus \text{Cl}^-(V),$$

where $\text{Cl}^+(V)$ denotes the self-dual elements and $\text{Cl}^-(V)$ denotes anti-self-dual elements. Both are just subspaces, not subalgebras.

Remark 25.2. $\mu \cdot 1 = \mu$ so 1 is *not* self-dual. But $\frac{1+\mu}{2} \in \text{Cl}^+(V)$ and $\frac{1-\mu}{2} \in \text{Cl}^-(V)$.

Example 25.2. $\mathrm{Cl}(\mathbb{R}^3) = \mathbb{H} \oplus \mathbb{H} = \mathrm{Cl}^+(\mathbb{R}^3) \oplus \mathrm{Cl}^-(\mathbb{R}^3)$. Recall that $\mathrm{Cl}(\mathbb{R}^3)^{\text{even}} = \{(p,p) : p \in \mathbb{H}\}$ and $\mathrm{Cl}(\mathbb{R}^3)^{\text{odd}} = \{(p,-p) : p \in \mathbb{H}\}$. So we have

$$\mathrm{Cl}^+(\mathbb{R}^3) = \mathrm{span}\left\{\frac{1+\mu}{2}, \frac{e_i + e_j \cdot e_k}{2}\right\}$$
$$\mathrm{Cl}^-(\mathbb{R}^3) = \mathrm{span}\left\{\frac{1-\mu}{2}, \frac{e_i - e_j \cdot e_k}{2}\right\}.$$



25.2. *Pin groups and spin groups*

Let $v \in \mathbb{R}^n$ te non-zero. Then we have $v \cdot v = v^2 = -|v|^2 1 \equiv 0$. So v is invertible, and $v^{-1} = -v/|v|^2$. If $|v| = 1$, then $v^{-1} = -v$. Now we are ready to define pin groups.

Definition 25.3. The *Pin group* of $(\mathbb{R}^n, \langle \_, \_ \rangle)$ denoted $\mathrm{Pin}(n)$, is the subgroup of $\mathrm{Cl}(V)^*$ generated by unit vectors in $\mathbb{R}^n$, i.e.,

$$\mathrm{Pin}(n) = \{u_1 \cdot u_2 \cdot \cdots \cdot u_r : |u_i| = 1, u_i \in V\}.$$

Remark 25.3. This is clearly a group, and notice that $\pm 1 \in \mathrm{Pin}(n)$. Note that every $\alpha \in \mathrm{Pin}(n)$ is *either even or odd.*

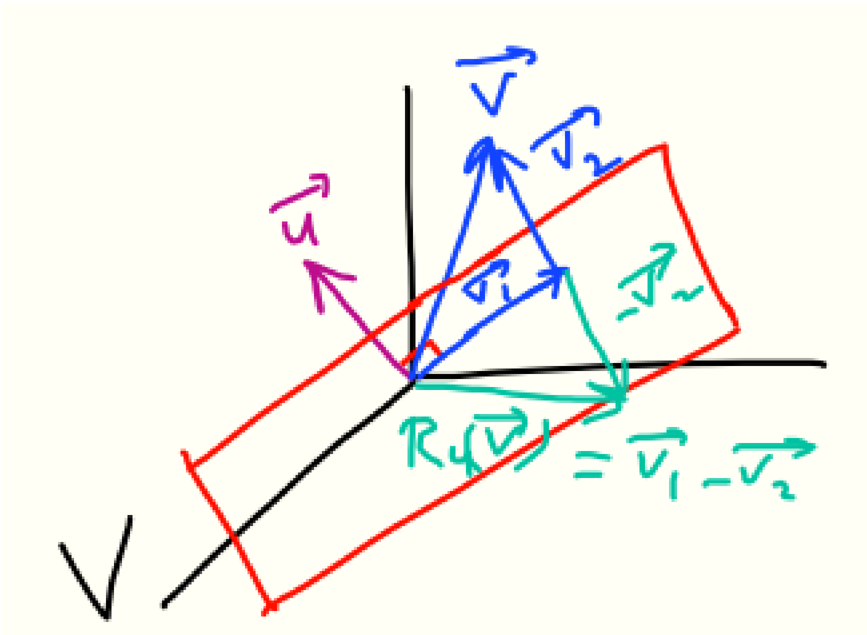
Definition 25.4. The *Spin group* of $(\mathbb{R}^n, \langle \_, \_ \rangle)$ denoted $\mathrm{Spin}(n)$ is

$$\mathrm{Spin}(n) := \mathrm{Pin}(n) \cap \mathrm{Cl}(V)^{\text{even}} = \{u_1 \cdot \cdots \cdot u_{2r} : u_i \in V, |u_i| = 1\}.$$

Remark 25.4. Clearly we have $\pm 1 \in \mathrm{Spin}(n)$ and $\mathrm{Spin}(n)$ is a subgroup of $\mathrm{Pin}(n)$.

Lemma 25.5 (Relation been reflections and the Clifford multiplication)**.** *Let $u \in V$ be non-zero. Let $R_u : V\ to\, \mathrm{Cl}(V)$ defined by $R_u(v) = -u \cdot v \cdot u^{-1}$. We know that $u^{-1} = -u$. Therefore*

$R_u(v) = u \cdot v \cdot u$ is linear in V. We also have $R_u(v) \in V$ for all $v \in V$, and $R_u : V \to V$ is reflection across the hyperplane orthogonal to span$\{u\}$.



Proof. R_u is linear, so there are two possible cases. First, suppose that $v = \lambda u$ where $\lambda \in \mathbb{R}$. Then $R_u(v) = uvu = \lambda u^3 = \lambda u(u^2) = -\lambda u = -v$ (recall that $u^2 = -|u|^2 1 = -1$). If $v \perp u$ (the second case), then in this case we have $u \cdot v = -v \cdot u$, so $R_u(v) = uvu = -vu^2 = -v(-1) = v$. ☐

This lemma motivates the following definition:

Definition 25.6. The *adjoint representation* Ad of $\mathrm{Cl}(V)^*$ on $\mathrm{Cl}(V)$ is the group homomorphism $\mathrm{Ad} : \mathrm{Cl}(V)^* \to \mathrm{GL}(\mathrm{Cl}(V))$ given by the isomorphism $\mathrm{Ad}_\alpha : \mathrm{Cl}(V) \cong \mathrm{Cl}(V)$ defined by $\beta \mapsto \alpha \cdot \beta \cdot \alpha^{-1}$ (obviously, $\alpha \in \mathrm{Cl}(V)^*$).

Remark 25.5. This is a representation since $\mathrm{Ad}_{\alpha^{-1}} = (\mathrm{Ad}_\alpha)^{-1}$ and $\mathrm{Ad}_\alpha \cdot \mathrm{Ad}_\beta = \mathrm{Ad}_{\alpha\beta}$.

Definition 25.7. The *twisted adjoint representation* $\widetilde{\mathrm{Ad}}$ is a representation of $\mathrm{Cl}(V)^*$ on $\mathrm{Cl}(V)$ given by

$$\widetilde{\mathrm{Ad}} : \mathrm{Cl}(V)^* \to \mathrm{GL}(\mathrm{Cl}(V))$$

where $\widetilde{\mathrm{Ad}}_\alpha \beta = \widetilde{\alpha} \cdot \beta \cdot \alpha^{-1}$ for all $\alpha \in \mathrm{Cl}(V)^*$.

Remark 25.6. $\widetilde{\mathrm{Ad}}$ is a representation because

$$\begin{aligned} \widetilde{\mathrm{Ad}}_{\alpha_1} \cdot \widetilde{\mathrm{Ad}}_{\alpha_2} \beta &= \widetilde{\alpha_1}(\widetilde{\alpha_2} \beta \alpha_2^{-1}) \alpha_1^{-1} \\ &= \widetilde{\alpha_1}\widetilde{\alpha_2} \beta (\alpha_1 \alpha_2)^{-1} \\ &= (\widetilde{\alpha_1 \alpha_2}) \beta (\alpha_1 \alpha_2)^{-1} \\ &= \widetilde{\mathrm{Ad}}_{\alpha_1 \alpha_2} \beta. \end{aligned}$$

Remark 25.7. If $\alpha \in V$ and $\alpha \in \mathrm{Cl}(V)^*$, then α is a non-zero vector in $\mathbb{R}^n$. We have $\widetilde{\alpha} = -\alpha$ and $\alpha^{-1} = -\alpha/\|\alpha\|^2$.

Let $\alpha = u \in V$. Then we have

$$\widetilde{\mathrm{Ad}}_u(\beta) = \widetilde{u}\beta u^{-1} = -u \cdot \beta u^{-1}.$$

Therefore $\widetilde{\mathrm{Ad}}_u|_V = R_u$.

Theorem 25.8. *The following sequences are short exact sequences:*

(1) $1 \to \mathbb{Z}/2\mathbb{Z} \xrightarrow{i} \mathrm{Pin}(n) \xrightarrow{\widetilde{\mathrm{Ad}}} \mathrm{O}(n) \to 1$

(2) $1 \to \mathbb{Z}/2\mathbb{Z} \xrightarrow{i} \mathrm{Spin}(n) \xrightarrow{\widetilde{\mathrm{Ad}}} \mathrm{SO}(n) \to 1.$

Proof. If $u \neq 0$ in V then $\widetilde{\mathrm{Ad}}_u = R_u \in \mathrm{O}(n)$. If $\alpha \in \mathrm{Pin}(n)$, then $\alpha = u_1 \cdot u_2 \cdot \dots \cdot u_r$ where $u_i \in V$ and $u_i \neq 0$ with $|u_i|| = 1$. Notice that we have

$$\widetilde{\mathrm{Ad}}_\alpha = \widetilde{\mathrm{Ad}}_{u_1 \cdot \dots \cdot u_r} = \widetilde{\mathrm{Ad}}_{u_1} \cdot \widetilde{\mathrm{Ad}}_{u_2} \cdot \dots \cdot \widetilde{\mathrm{Ad}}_{u_r} = R_{u_1} \cdot \dots \cdot R_{u_r} \in \mathrm{O}(n).$$

By Cartan-Dieudonné theorem, every $P \in \mathrm{O}(n)$ is a product of a finite number of replications. So $P = R_{u_1} \cdot \dots \cdot R_{u_r}$ for some $u_1, \dots, u_r \in V$, all of which are non-zero. Hence $\widetilde{\mathrm{Ad}} : \mathrm{Pin}(n) \to \mathrm{O}(n)$ is surjective.

Note that i is injective with image ± 1; and $\widetilde{\mathrm{Ad}}_{-1}\beta = \widetilde{-1} \cdot \beta(-1)^{-1} = (-1) \cdot \beta \cdot (-1) = \beta$. Hence $\widetilde{\mathrm{Ad}}_{-1} = 1$. Therefore $\mathrm{im}(i) \subseteq \ker(\widetilde{\mathrm{Ad}})$. We need to show that $\mathrm{im}(i) = \ker(\widetilde{\mathrm{Ad}})$.

Suppose that $\alpha \in \mathrm{Pin}(n)$ is in $\ker(\widetilde{\mathrm{Ad}})$. That is, $\widetilde{\mathrm{Ad}}_\alpha = 1$. Hence $\widetilde{\alpha} \cdot v\alpha^{-1} = v$ for all $v \in V$, or equivalently $\widetilde{\alpha}v = v\alpha$ for all $v \in V$.

If α is odd, then $\widetilde{\alpha} = -\alpha$. So we have $-\alpha v = v\alpha$ for all $v \in V$. Hence $\alpha \in \mathrm{twcent}(\mathrm{Cl}(V))$. But then recall that

	n even	n odd
cent	1	$\{1, \mu\}$
twcent	μ	$\{0\}$

Hence there are no non-zero elements in the twisted centre, so this is a contradiction. Therefore α must be even, i.e., $\widetilde{\alpha} = \alpha$. So $\alpha v = v\alpha$ for all $v \in V$. Hence $\alpha \in \mathrm{cent}(\mathrm{Cl}(V))$, so $\alpha \in \mathrm{span}\{1\}$. Thus $\alpha = t$ where $t \in \mathbb{R}$. One can easily verify that if $\alpha \in \mathrm{Pin}(n)$ then $t = \pm 1$, thereby proving (a).

For (b), it's the similar argument as (a), but here we use the fact that $\det(R_u) = -1$ for all $u \in V \setminus \{0\}$. Hence, the elements of $\mathrm{SO}(n)$ are the products of *even* number of reflections. Follow the reasoning outlined in the proof of (a). $\square$

Note that we still haven't explored why $\mathrm{Pin}(n)$ and $\mathrm{Spin}(n)$ are matrix groups. Why?

Definition 25.9. We define $\mathbb{P}$ to be the *space of pinors* and $\mathbb{S}$ to be the *space of spinors.*

The spaces $\mathbb{P}$ resp. $\mathbb{S}$ of pinors resp. spinors are finite-dimensional real vector spaces on which $\mathrm{Cl}(\mathbb{R}^n)$ resp. $\mathrm{Cl}(\mathbb{R}^n)^{\mathrm{even}}$ acts on.

Proposition 25.10. $\mathrm{M}_{N\times N}(\mathbb{F})$ *where* $\mathbb{F} = \mathbb{R}, \mathbb{C},$ *or* $\mathbb{H}$ *has* exactly one *irreducible representation (as a* real algebra*), which is* $\mathbb{F}^N$.

TABLE 1. Table for pinors

$n \bmod 8$	$\mathrm{Cl}(\mathbb{R}^n)$	dimension $= 2^n$	$\mathbb{P}$
0	$\mathrm{M}_{N\times N}(\mathbb{R})$	N^2	$\mathbb{R}^N$
1	$\mathrm{M}_{N\times N}(\mathbb{C})$	$2N^2$	$\mathbb{C}^N$
2	$\mathrm{M}_{N\times N}(\mathbb{H})$	$4N^2$	$\mathbb{H}^N$
3	$\mathrm{M}_{N\times N}(\mathbb{H}) \oplus \mathrm{M}_{N\times N}(\mathbb{H})$	$8N^2$	$\mathbb{H}^N \oplus \mathbb{H}^N$
4	$\mathrm{M}_{N\times N}(\mathbb{H})$	$4N^2$	$\mathbb{H}^N$
5	$\mathrm{M}_{N\times N}(\mathbb{C})$	$2N^2$	$\mathbb{C}^N$
6	$\mathrm{M}_{N\times N}(\mathbb{R})$	N^2	$\mathbb{R}^N$
7	$\mathrm{M}_{N\times N}(\mathbb{R}) \oplus \mathrm{M}_{N\times N}(\mathbb{R})$	$2N^2$	$\mathbb{R}^N \oplus \mathbb{R}^N$

TABLE 2. Table for spinors

$n \bmod 8$	$\mathrm{Cl}(\mathbb{R}^n)^{\text{even}} \cong \mathrm{Cl}(\mathbb{R}^{n-1})$	$\mathbb{S}$
0	$\mathrm{M}_{N\times N}(\mathbb{R}) \oplus \mathrm{M}_{N\times N}(\mathbb{R})$	$\mathbb{R}^N \oplus \mathbb{R}^N$
1	$\mathrm{M}_{N\times N}(\mathbb{R})$	$\mathbb{R}^N$
2	$\mathrm{M}_{N\times N}(\mathbb{C})$	$\mathbb{C}^N$
3	$\mathrm{M}_{N\times N}(\mathbb{H})$	$\mathbb{H}^N$
4	$\mathrm{M}_{N\times N}(\mathbb{H}) \oplus \mathrm{M}_{N\times N}(\mathbb{H})$	$\mathbb{H}^N \oplus \mathbb{H}^N$
5	$\mathrm{M}_{N\times N}(\mathbb{H})$	$\mathbb{H}^N$
6	$\mathrm{M}_{N\times N}(\mathbb{C})$	$\mathbb{C}^N$
7	$\mathrm{M}_{N\times N}(\mathbb{R})$	$\mathbb{R}^N$

Proposition 25.11. *Both* $\mathrm{Pin}(n)$ *and* $\mathrm{Spin}(n)$ *are canonically matrix groups. Namely,* $\mathrm{Spin}(n)$ *is the square matrices of size depending on n (in a rather complicated way), so* $\mathrm{Spin}(n)$ *is a matrix group such that* $\mathrm{Lie}(\mathrm{Spin}(n)) = \mathrm{Lie}(\mathrm{SO}(n)) = \mathfrak{so}(n)$*, by using the two-to-one homomorphism.*

Proposition 25.12. $\mathrm{Spin}(n)$ *is connected.* $\mathrm{Spin}(n)$ *is simply connected for all* $n \geq 3$.

Proof. Let $u_1, u_2 \in \mathbb{R}^n$ with $|u_1| = |u_2| = 1$ and $\langle u_1, u_2\rangle = 0$. Note that for any $t \in \mathbb{R}$,

$$\begin{aligned}(-\cos(t)u_1 - \sin(t)u_2)\cdot(\cos(t)u_1 - \sin(t)u_2) &= (\cos^2 t - \sin^2 t)1 + 2\sin(t)\cos(t)u_1\cdot u_2 \\ &= \cos(2t)1 + \sin(2t)(u_1\cdot u_2).\end{aligned}$$

This is a path from 1 to $u_1 \cdot u_2$. Therefore $\mathrm{Spin}(n)$ is connected. Hmm but this isn't enough to show that it is simply connected! □

Example 25.13. Both Spin(3) and Spin(4) can be explicitly realized using quaternions: $S^3 \cong \mathrm{Spin}(3)$ as manifolds, and $\mathrm{Spin}(3) \cong \mathrm{SU}(2) \cong \mathrm{Sp}(1)$ as matrix Lie groups. Also, note that $S^3 \subseteq \mathbb{H} =\cong \mathbb{R}^4$, so S^3 can be thought of as unit quaternions.

Consider the map $\mathrm{Spin}(3) \to \mathrm{SO}(3)$ defined by $p \mapsto \mathrm{Ad}_p$. Then $\mathrm{Ad}_p(v) = pvp^{-1} = pv\bar{p}$ is a two-to-one Lie group homomorphism. And $|\mathrm{Ad}_p(v)| = |pv\bar{p}| = |v|$ so $\mathrm{Ad}_p \in \mathrm{O}(3)$ by continuity. Hence $\mathrm{Ad}_p \in \mathrm{SO}(3)$ for all $p \in \mathrm{Spin}(3)$, with $\ker(\mathrm{Ad}) = \pm 1$. One can show that this map is surjective and S^3 is simply connected. So $\mathrm{Spin}(3) \cong S^3$, which can be viewed as the set of unit quaternions. Similarly, $\mathrm{Spin}(4) = S^3 \times S^3 = \mathrm{Spin}(3) \times \mathrm{Spin}(3)$. Let $\rho : S^3 \times S^3 \to \mathrm{SO}(4)$ such that $(p,q) \mapsto \rho_{p,q}$ where $\rho_{p,q}(v) := p\sqrt{q}$. We see that $\ker(\rho) = \pm 1$.

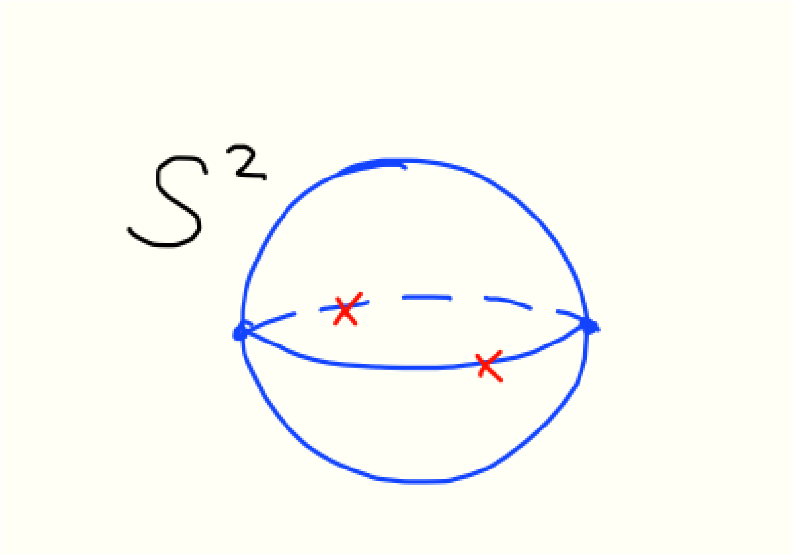
The consequence is that $\mathrm{Lie}(\mathrm{Spin}(4)) = \mathfrak{so}(4) = \mathfrak{so}(3) \oplus \mathfrak{so}(3)$. For the second equality, note that

$$\mathfrak{so}(4) = \wedge^2(\mathbb{R}^4) = \wedge^2_+(\mathbb{R}^4) \oplus \wedge^2_-(\mathbb{R}^4),$$

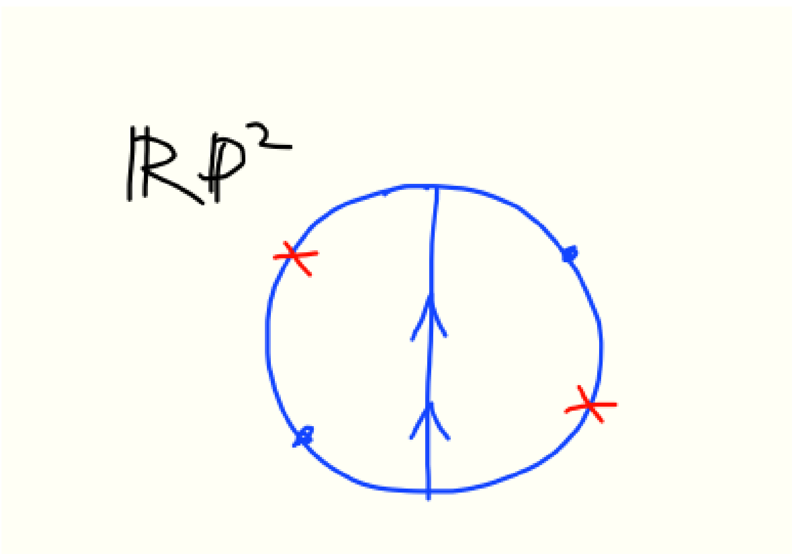
and $\wedge^2_+(\mathbb{R}^4) \cong \mathbb{R}^3 \cong \wedge^2(\mathbb{R}^3) \cong \mathfrak{so}(3)$ and similarly $\wedge^2_-(\mathbb{R}^4) \cong \mathbb{R}^3 \cong \wedge^2(\mathbb{R}^3) \cong \mathfrak{so}(3)$.

Remark 25.8. Note that $\mathrm{Spin}(3) = S^3$, and by the two-to-one map from Spin(3) to SO(3) and the corresponding map S^3 to $\mathbb{RP}^3$ (identify antipodal points), we have $\pi_1(\mathrm{SO}(3)) = \mathbb{Z}/2\mathbb{Z}$ and $\pi_1(\mathrm{Spin}(3)) = \{1\}$. So there is a non-contractible loop in $\mathbb{RP}^3$, but when doubled it does become contractible.

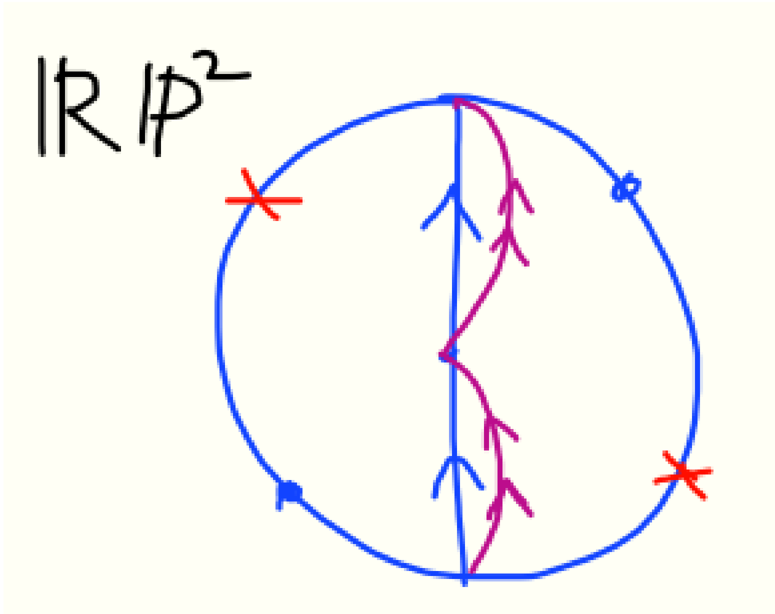
As an example, we will consider the case S^2 and $\mathbb{RP}^2$. Notice that you can deform S^2 into $\mathbb{RP}^2$ from



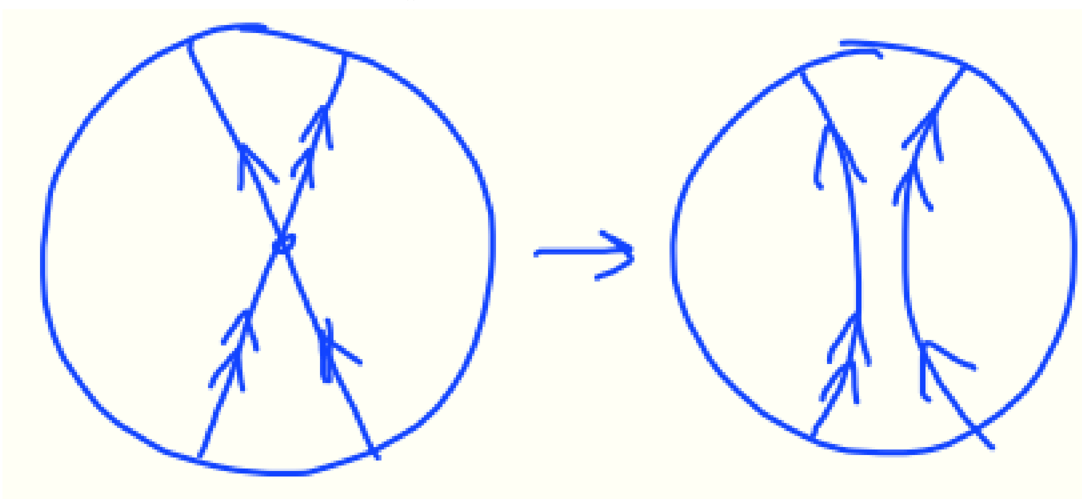
to



Note that the above loop is *not* contractible. But if we add two more segments (purple-coloured lines), then we do have contractible loops:



Let's deform a bit more for clearer pictures. Indeed,



This marks the completion of the lectures and this lecture note.

Department of Pure Mathematics, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1

E-mail address: `hsyang@uwaterloo.ca`