SYMMETRIES (MS-E1997)

≈ INTRODUCTION TO REPRESENTATION THEORY

Practicalities

Mon	12–14	lecture,	M3
Wed	12–14	exercises,	Y346
Thu	14–16	lecture,	M3

5 credits

grading: EITHER $\underbrace{\tfrac{2}{3}\text{ exam} + \tfrac{1}{3}\text{ written exercises OR only exam}}_{\text{whichever is higher}}$

exam: we will agree on a date during week 43 (October 19–23)

written exercise solutions (to problems marked ) are to be returned to course TA Alex Karrila by Wednesdays at 12.

feedback lunch: Thu, September 24, at 12

Essentially all information on the course web page MyCourses ⇝ MS-E1997

Textbook:

Fulton & Harris: "Representation theory: a first course"

(covers most of the contents of this course plus of course much more, in a very concrete style!)

typically: a group

Symmetry = a set (*) of transformations which leave some relevant property unchanged

(of some object)

"invariant"

(*): Transformations of the object itself, or more generally and commonly, of some space associated to the object.

Examples:

- symmetry group of a regular polyhedron acts by transformations of the set(s) of vertices / edges / faces of the polyhedron
- on any reasonable space of functions of n variables, the group of permutations of the variables acts naturally
  - special case: functions of two variables with values on a vector space

    transposition τ acts on functions by
$$(\tau . f)(x_1, x_2) = f(x_2, x_1)$$
    Now we could consider e.g.

    symmetric functions $f(x_2, x_1) = f(x_1, x_2)$

    antisymmetric functions $f(x_2, x_1) = -f(x_1, x_2)$

    and note that any function decomposes as the sum of its symmetric and antisymmetric parts.

    Think about generalizations to n variables of this decomposition!

Recap:

Symmetry of some object
$$\|$$
a group of transformations of some space associated with the object, which leaves properties of interest invariant

Representation theory is concerned with the case when the space is a vector space. Why?

- ▶ This is often naturally the case, e.g.
  - a space of $\mathbb{R}$-valued or $\mathbb{C}$-valued functions defined on the object
  - physical states of quantum mechanical systems are vectors in a Hilbert space.
  - topological / geometric information encoded in vector spaces such as homology, cohomology, ...
  - transition probabilities of Markov processes encoded in a matrix acting on a vector space
- ▶ With the vector space structure one has a rich and powerful theory with many applications!

Two concrete examples of the most common way representation theory helps in solving problems:

Example 1: Let D be a translation invariant differential operator on the torus $\mathbb{T} = (\mathbb{R}/\mathbb{Z}) \times (\mathbb{R}/\mathbb{Z})$,
e.g. $D = \text{Laplacian} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$.

Translation invariance is a symmetry!

For any $a, b \in \mathbb{R}$ translations by (a,b) act on functions $f: \mathbb{T} \to \mathbb{C}$ by

$$(\theta_{a,b} f)(x,y) = f(x-a, y-b).$$

D translation invariant means that D commutes with $\theta_{a,b}$ for any $a,b \in \mathbb{R}$

$\Longrightarrow$ D and $\theta_{a,b}$ can be simultaneously diagonalized

Diagonalization of translations $\theta_{a,b}$? Fourier transform

$$f_{n,m}(x,y) = e^{i2\pi nx + i2\pi my}$$

$$\theta_{a,b} f_{n,m} = \lambda_{n,m}(a,b) \, f_{n,m} \qquad \lambda_{n,m}(a,b) = e^{-i2\pi na - i2\pi mb}$$

Thus $f_{n,m}$ are also eigenvectors of D
– we don't need to look any further!

e.g. $D = \text{Laplacian} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$

$$D f_{n,m} = -4\pi^2(n^2+m^2) \cdot f_{n,m}$$

Note: Translations form an abelian symmetry group
$\rightsquigarrow$ a very easy case.

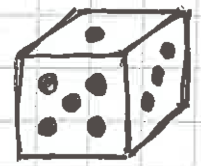
In some sense representation theory is the generalization of Fourier transform for more complicated symmetries.

Example 2: ~~Shuffling a deck of cards~~ (symmetric group S_{52})

too complicated, let's simplify

Shuffling a cubic dice:

which face is on top?



At time $t=0$ face ⚀ is on top.

At time $t+1$ apply a shuffling operation on the position of the dice at time t:
turn the dice randomly to one of its four adjacent faces (not the same, and not the opposite)

Markov process $(X_t)_{t \in \mathbb{Z}_{\geq 0}}$

$$X_t \in \{ ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ \} = F = \{\text{faces of the dice}\}$$

Transition probabilities $a, b \in F$

$$P_{a,b} = \begin{cases} \frac{1}{4} & \text{if faces } a,b \text{ are adjacent} \\ 0 & \text{if faces } a,b \text{ are either equal or opposite} \end{cases}$$

Probabilities of value a at time t

$$\mu_t(a) = \sum_{a_1, \ldots, a_{t-1} \in F} \mu_0(a_0) P_{a_0,a_1} P_{a_1,a_2} \cdots P_{a_{t-1}a_t} \qquad \left(\mu_t = \mu_0 P^t \right)$$

μ_t is a function on F

$$\mu_t \in \mathbb{C}^F = \{ \varphi : F \to \mathbb{C} \} \quad \longleftarrow \text{vector space}$$

P is a linear operator on $\mathbb{C}^F$

μ_t for large t governed by leading eigenvectors of P

symmetry: P commutes with the action of the group of symmetries of the cube.

Rep. theory gives eigenvectors and multiplicities of different eigenvalues! Compare with Fourier transform.

Description of a few applications of representation theory

- **Quantum mechanics and chemistry**

A physical system, e.g. a molecule, described by the time-dependent state $\psi(t)$ in some Hilbert space V. Time-dependency governed by Schrödinger eq. $i\hbar \frac{\partial}{\partial t}\psi(t) = H\psi(t)$ where H is the Hamiltonian ("energy operator") of the system. A symmetry of the system (dynamics) is a representation of the symmetry group G on the space V, which commutes with H. The stationary physical states are eigenvectors of H. The eigenspaces of H are subrepresentations. One can conclude about stationary states, multiplicities in spectrum, conserved quantum numbers etc. by using representation theory of G.

- **Number theory** The following result of Dirichlet is a simple application of representation theory.

Thm (Dirichlet) Let $q \in \mathbb{Z}_{>0}$ and $a \in \mathbb{Z}$ be coprime with q. Then there exist infinitely many primes p of the form $p \equiv a \pmod q$.

- **Group theory** Representation theory can be used to establish various properties of groups. Here is one classical application (proof without rep. th. much harder)

Thm (Burnside) Let G be a group whose order is divisible by at most two primes. Then G is solvable.

(Recall: G solvable means $\exists$ subgroups
$$1 \triangleleft G_1 \triangleleft G_2 \triangleleft \cdots \triangleleft G_{n-1} \triangleleft G_n = G$$
s.t. G_{k+1}/G_k are abelian.)

- **Geometry** E.g. the classifications of Riemannian symmetric spaces and singularities of Kleinian surfaces use rep. th.

I REPRESENTATIONS OF FINITE GROUPS

BASIC DEFINITIONS AND BACKGROUND

Def. A group is a set G equipped with a binary operation $* : G \times G \longrightarrow G$ s.t.

- $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3 \quad \forall g_1, g_2, g_3 \in G$ "associativity"
- $\exists e \in G$ s.t. $\forall g \in G$
  $g * e = g = e * g$ "neutral element"
- $\forall g \in G \quad \exists g^{-1} \in G$ s.t.
  $g * g^{-1} = e = g^{-1} * g$ "inverse element"

Remark For abelian groups often $g_1 * g_2$ is denoted by $g_1 + g_2$ and for general groups we usually omit $*$ and denote $g_1 * g_2$ by $g_1 g_2$.

Examples

- Cyclic group $\mathbb{Z}/n\mathbb{Z}$ (abelian), $\#(\mathbb{Z}/n\mathbb{Z}) = n$.
- Symmetric group X a set "permutation group of X" "symmetric group on X"

  $S(X) = \{ \varphi : X \to X \text{ bijection} \}$

  group operation: composition of functions

  $\varphi_1 * \varphi_2 = \varphi_1 \circ \varphi_2 \qquad \big(\varphi_1 \circ \varphi_2 (x) = \varphi_1(\varphi_2(x)) \big)$

  Special case: If $X = \{1, 2, \dots, n\}$ denote $S(X) = S_n$.

  $\# S_n = n! = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$

- General linear group $\mathbb{K}$ a field (say $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$)

  $GL_n(\mathbb{K}) = \{ M \in \mathbb{K}^{n \times n} \mid \det(M) \neq 0 \}$
  = the set of invertible $n \times n$ - matrices with entries in $\mathbb{K}$.

  group operation: matrix multiplication

Corresponding coordinate invariant description:

V a $\mathbb{K}$-vector space

$Aut(V) = \{ L : V \to V \text{ invertible linear map} \}$

group operation: composition of (linear) maps

If $\dim_{\mathbb{K}}(V) = n$ then $Aut(V) \cong GL_n(\mathbb{K})$.

- <u>Dihedral group</u>

Symmetry group of regular n-gon

generators r, m

$r =$ rotation by $\frac{2\pi}{n}$ $m =$ reflection

relations $r^n = e$, $mrm = r^{-1}$, $m^2 = e$

$G = \{ e, r, r^2, \dots, r^{n-1}, m, rm, r^2 m, \dots, r^{n-1} m \}$

$\# G = 2n$

Terminology: the number of elements, $\# G$, is called the <u>order</u> of G.

G is said to be a finite group if $\# G < \infty$.

<u>Def</u>: If G and $\tilde{G}$ are two groups (group operations $*$ and $\tilde{*}$, respectively) then a mapping $f : G \to \tilde{G}$ is said to be a (group) <u>homomorphism</u> if $\forall g, h \in G$: $f(g * h) = f(g) \,\tilde{*}\, f(h)$.

<u>Examples</u>:
- Determinant $\det : GL_n(\mathbb{K}) \longrightarrow \mathbb{K}^\times = \mathbb{K} \setminus \{0\}$ is a homomorphism from general linear group to the multiplicative group of invertible scalars.
- The signature of a permutation (parity of the number of transpositions in any expression) $sgn : S_n \to \mathbb{Z}/2\mathbb{Z}$ is a homomorphism from the symmetric group to the two element group $\mathbb{Z}/2\mathbb{Z} \cong \{\pm 1\} \subset \mathbb{R}^\times$

You should be familiar with notions of subgroup, normal subgroup, quotient group, kernel, isomorphism, ...

For any algebraic structure, one of the first basic results is the isomorphism theorem. For groups:

Thm (Isomorphism theorem for groups)

Let G, H be groups and $f: G \to H$ a homomorphism.
Then
1°) $\mathrm{Im}(f) = f(G) \subset H$ is a subgroup
2°) $\mathrm{Ker}(f) = f^{-1}(\{e_H\}) \subset G$ is a normal subgroup
3°) the quotient $G/\mathrm{Ker}(f)$ is isomorphic to $\mathrm{Im}(f)$.

It's good to realize that isomorphism theorems are similar for all algebraic structures, once the suitable notions have been given. In this course, for example, we will use the following cases:

STRUCTURE	MORPHISM f	IMAGE $\mathrm{Im}(f)$	KERNEL $\mathrm{Ker}(f)$ (OR FACTOR TO BE DIVIDED BY)
group	homomorphism	subgroup	normal subgroup
vector space	linear map	vector subspace	vector subspace
Lie algebra	(Lie algebra) homomorphism	Lie subalgebra	ideal
representation	intertwining map	subrepresentation	subrepresentation
⋮	⋮	⋮	⋮

We can summarize the isomorphism theorem in the commutative diagram:

if $f: G \to H$ is a group homomorphism, then there exists a unique group isomorphism $\tilde{f}: G/\mathrm{Ker}(f) \to \mathrm{Im}(f) \subset H$ such that the following diagram commutes:

$$\begin{tikzcd} G \arrow[rr, "f"] \arrow[dr, "\text{projection}"'] & & H \\ & G/\mathrm{Ker}(f) \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

Def.: An action of a group G on a set X is a group homomorphism
$$\alpha : G \longrightarrow S(X)$$

So, concretely, $g \in G$ acts on $x \in X$ by
$$x \longmapsto (\alpha(g))(x) =: g.x \quad \text{(convenient notation)}$$
and the defining requirements of group action are
$$e.x = x \quad \forall x \in X \quad \text{and} \quad (gh).x = g.(h.x).$$

Examples
- Any group acts on itself by left multiplication
$(\alpha(g))(h) = g*h$ (or briefly $g.h = gh$) $(g, h \in G)$
- The symmetric group S_n acts on $\{1, 2, \dots, n\}$ in the obvious way: $\sigma \in S_n$ $x \in \{1, \dots, n\}$: $\sigma.x = \sigma(x)$.
- The dihedral group acts on the set of vertices of an n-gon in the obvious way.

Def.: A representation of a group G on a vector space V is a group homomorphism
$$\rho : G \longrightarrow \mathrm{Aut}(V)$$

So, concretely, $g \in G$ acts by a linear map $\rho(g)$ on vectors $v \in V$
$$v \longmapsto \rho(g)\, v = g.v$$ (convenient notation — V becomes a (left) G-module)

and the defining requirements of a representation are
$$g.(v_1 + v_2) = g.v_1 + g.v_2 \quad \forall g \in G \quad \forall v_1, v_2 \in V$$
$$g.(\lambda v) = \lambda\, g.v \quad \forall g \in G \quad \forall v \in V \quad \forall \lambda \in K.$$
$$e.v = v \quad \forall v \in V$$
$$(gh).v = g.(h.v) \quad \forall g, h \in G \quad \forall v \in V.$$

Def.: A subspace $V' \subset V$ of a representation $\rho : G \to \mathrm{Aut}(V)$ is a <u>subrepresentation</u> if $\forall g \in G$ $\rho(g)\, V' \subset V'$.

Then the restrictions $g \mapsto \rho(g)|_{V'}$ define a representation on V'.

Remark: We almost always consider complex vector spaces!

Examples

- trivial representation

  $\rho(g) = id_V \quad \forall g \in G$ trivial representation on V

  The case $V = \mathbb{C}$ with $\rho(g) = id_{\mathbb{C}} \quad \forall g \in G$ is called the trivial representation of G.

- fundamental representation of the dihedral group D_4

  D_4: generators r, m, relations $r^4 = e$, $mrm = r^{-1}$

  Representation ρ on $V = \mathbb{R}^2$:

  $$r \mapsto R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad m \mapsto M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

  Observation 0: It is sufficient to give the images of the generators.

  Observation 1: R and M satisfy the relations

  $$R^4 = \mathbb{1}_{2\times 2} \qquad M^2 = \mathbb{1}_{2\times 2} \qquad MRM = R^{-1}$$

  so indeed there exists a homom. $\rho: D_4 \to Aut(\mathbb{R}^2)$ defined by these.

  Observation 2: One can concretely check that $Im(\rho)$ consists of 8 different matrices. This is basically the easiest way to prove that $\# D_4 \geq 8$. (To show $\# D_4 \leq 8$ one just uses the relations to reduce any word to r^n or mr^n, $n = 0,1,2,3$)

REPRESENTATIONS AND THEIR EQUIVALENCES

Recall: A representation of a group G on a vector space V is a homomorphism

$$\rho : G \longrightarrow \mathrm{Aut}(V) = \{T : V \to V \text{ bijective linear map}\}.$$

Convenient notation: $\rho(g)v = g.v$ for $g \in G$, $v \in V$

$\rightsquigarrow$ V becomes a left G-module

$\therefore$ representations of G $\longleftrightarrow$ left G-modules

Maps between representations which preserve the representation structure are called alternatively either intertwining maps or G-module maps.

Def: Let $\rho_1 : G \to \mathrm{Aut}(V_1)$ and $\rho_2 : G \to \mathrm{Aut}(V_2)$ be two representations of a group G. A linear map $f : V_1 \longrightarrow V_2$ is called an intertwining map if

$$\forall g \in G : \quad f \circ \rho_1(g) = \rho_2(g) \circ f.$$

i.e. $\forall g \in G \ \forall v \in V \quad f(g.v) = g.f(v)$

A bijective intertwining map is called either an isomorphism or an equivalence (of representations).

One fundamental question in representation theory is:

- Can we classify all representations (possibly all finite-dim.) of a given group up to isomorphism?

As for other algebraic structures, one has an isomorphism theorem for representations (of a given G).

Exercise State and prove the isomorphism theorem for representations.

It is common to denote the space of intertwining maps between two representations V_1, V_2 of G by $\mathrm{Hom}_G(V_1, V_2)$.

OPERATIONS ON REPRESENTATIONS

Recall: A representation of a group G is a homomorphism $\rho: G \longrightarrow \mathrm{Aut}(V)$ from G to the group of invertible linear maps of a vector space V.

We often denote briefly $\rho(g)v = g.v$ for $g \in G$, $v \in V$.

We will next show how to construct new representations from given ones by

- direct sums $V_1 \oplus V_2$
- tensor products $V_1 \otimes V_2$
- spaces of linear maps $\mathrm{Hom}(V_1, V_2)$ and in particular duals $V^* = \mathrm{Hom}(V, \mathbb{K})$
- invariants

In the following three constructions we assume that $\rho_1 : G \to \mathrm{Aut}(V_1)$ and $\rho_2 : G \to \mathrm{Aut}(V_2)$ are two representations.

Direct sum

Recall that $V_1 \oplus V_2$ is the vector space of pairs (v_1, v_2) of vectors $v_1 \in V_1$, $v_2 \in V_2$. (Usually denoted $v_1 + v_2$)

This becomes a representation by setting

$$\rho(g)\,(v_1, v_2) = (\rho_1(g)v_1\,,\ \rho_2(g)v_2).$$

i.e. $g.(v_1 + v_2) = g.v_1 + g.v_2$ in the module notation.

Tensor product

Recall the construction of the tensor product vector space $V_1 \otimes V_2$ (spanned by $v_1 \otimes v_2$ with $v_1 \in V_1$, $v_2 \in V_2$).

This becomes a representation by setting

$$\rho(g)\,(v_1 \otimes v_2) = \rho_1(g)v_1 \otimes \rho_2(g)v_2$$

and extending linearly.

i.e. $g.(v_1 \otimes v_2) = g.v_1 \otimes g.v_2$

Linear maps

The space of linear maps from V_1 to V_2 is denoted by $\mathrm{Hom}(V_1, V_2)$ and it is itself a vector space.

This becomes a representation by setting, for any linear map $T: V_1 \to V_2$ and $g \in G$

$$\rho(g)\,T = \rho_2(g) \circ T \circ \rho_1(g^{-1}) : V_1 \to V_2$$

i.e. $(g.T)(v) = g.(T(g^{-1}.v)) \quad \forall v \in V_1$

Check: Clearly $\rho(g): \mathrm{Hom}(V_1, V_2) \to \mathrm{Hom}(V_1, V_2)$ is linear. We thus only need to observe that for any $g, h \in G$

$$\begin{aligned}
\rho(g)(\rho(h)T) &= \rho_2(g) \circ (\rho(h)T) \circ \rho_1(g^{-1}) \\
&= \rho_2(g) \circ \rho_2(h) \circ T \circ \rho_1(h^{-1}) \circ \rho_1(g^{-1}) \\
&= \rho_2(gh) \circ T \circ \rho_1((gh)^{-1}) \\
&= \rho(gh)\,T
\end{aligned}$$

or perhaps more concretely

$$\begin{aligned}
(g.(h.T))(v) &= g.((h.T)(g^{-1}.v)) \\
&= g.(h.(T(h^{-1}.(g^{-1}.v)))) \\
&= gh.\,T((gh)^{-1}.v) \\
&= ((gh).T)(v)
\end{aligned}$$

Duals

If $\rho: G \to \mathrm{Aut}(V)$ is a representation, then as a special case of the previous construction, the dual $V^* = \mathrm{Hom}(V, \mathbb{K}) = \{\varphi : V \to \mathbb{K} \text{ linear}\}$ becomes a representation. Note that on the 1-dimensional vector space $\mathbb{K}$ we use the trivial representation — all $g \in G$ act as identity on $\mathbb{K}$.
Concretely, for $g \in G$, $\varphi \in V^*$, $v \in V$ we have

$$\langle g.\varphi, v \rangle = \langle \varphi, g^{-1}.v \rangle .$$

Exercise If V_1 and V_2 are finite dimensional, then we can identify the vector spaces

$$\mathrm{Hom}(V_1, V_2) \cong V_2 \otimes V_1^* .$$

The constructions of dual representation and tensor product representation makes the right hand side a representation, whereas the left hand side was also made a representation. Show that the two constructions coincide in the sense that the identification is an isomorphism of representations.

SUBREPRESENTATIONS, IRREDUCIBILITY AND COMPLETE REDUCIBILITY

Def: A subrepresentation of a representation $\rho : G \to \mathrm{Aut}(V)$ is a vector subspace $\tilde{V} \subset V$ such that $\forall g \in G$: $\rho(g)\,\tilde{V} \subset \tilde{V}$. ("invariant subspace")

Note that $\tilde{V}$ is itself naturally a representation, with $\tilde{\rho} : G \to \mathrm{Aut}(\tilde{V})$ defined by restriction

$$\tilde{\rho}(g) = \rho(g)\big|_{\tilde{V}} \qquad \forall g \in G.$$

Examples

- The zero subspace $\{0\} \subset V$ and the full space $V \subset V$ are obviously always subrepresentations.
- If $V = V_1 \oplus V_2$ is the direct sum of two representations, then $V_1 \subset V$ and $V_2 \subset V$ are subrepresentations.

It turns out that the last example above is in fact the general case — any subrepresentation is a direct summand in the full representation (Well, we need the group G to be finite and the field $\mathbb{K}$ to have characteristic 0.)

Proposition: Let G be a finite group, and assume that the characteristic of the field $\mathbb{K}$ does not divide the order $\#G$ of the group. If V is a representation of G, and $V' \subset V$ is a subrepresentation, then there exists another subrepresentation $V'' \subset V$ such that $V = V' \oplus V''$.

(This V'' is called a "complementary subrepresentation")

Proof: We can first choose a complementary vector subspace $U \subset V$, i.e., the equality

$$V = V' \oplus U$$ holds as vector spaces.

Let $\pi' : V \to V'$ be the projection to first summand, i.e. $\pi'(v'+u) = v'$ if $v' \in V' \subset V$ and $u \in U \subset V$.

Define $\pi : V \to V$ by

$$\pi(v) = \frac{1}{\#G} \sum_{g \in G} g.\pi'(g^{-1}.v).$$

(Recalling Hom(V,V) representation, this says $\pi = \frac{1}{\#G} \sum_g (g.\pi')$)

We will show that this π is a projection to V' as well.

First check that $\mathrm{Im}(\pi) \subset V'$:

recall that $\pi'(w) \in V'$ for any $w \in V$, and $g.V' \subset V'$ for all $g \in G$, so each term in the sum is in V'.

Then check that the restiction of π to V' is the identity: if $v' \in V' \subset V$ then $g^{-1}.v' \in V'$ and thus $\pi'(g^{-1}.v') = g^{-1}.v'$ and finally $g.\pi'(g^{-1}.v') = g.g^{-1}.v' = v'$, which shows that $\pi(v') = \frac{1}{\#G} \sum_{g \in G} v' = v'$.

The final observation is that π is an intertwining map: indeed, if $h \in G$, we have

$$\pi(h.v) = \frac{1}{\#G} \sum_{g \in G} g.\pi'(g^{-1}h.v)$$

(Change of variables $\tilde{g} = h^{-1}g$ in the sum)

$$= \frac{1}{\#G} \sum_{\tilde{g} \in G} (h\tilde{g}).\pi'(\tilde{g}^{-1}.v)$$

$$= \frac{1}{\#G} \sum_{\tilde{g} \in G} h.\left(\tilde{g}.\pi'(\tilde{g}^{-1}.v)\right) = h.\pi(v).$$

Now $V'' = \mathrm{Ker}(\pi)$ is a subrepresentation, and we have $V = V' \oplus V''$. □

So whenever we find a subrepresentation, we are able to decompose the representation as a direct sum of two pieces. This motivates:

Def. A representation V is called *irreducible* if $V \neq \{0\}$ and V has no other subrepresentations but $\{0\}$ and V.

The idea is that we can decompose any representation to a direct sum of irreducible pieces.

Theorem Let G be a finite group, and assume that the characteristic of K does not divide $\# G$. Any finite dimensional representation V of G can be written as $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$, where $V_1, V_2, \dots, V_n$ are irreducible subrepresentations, and this decomposition is unique up to permutations of the summands.

Proof: Easy induction on $\dim(V)$. $\square$

The task of classifying of all (finite-dimensional) representations of G is thus reduced to the classification of all irreducible representations.

Schur's lemma

There is very little freedom for constructing intertwining maps between irreducible representations — and this turns out to be really crucial in all representation theory:

Theorem (Schur's lemma)

Let V and W be irreducible representations of a group G, and let $f: V \to W$ be an intertwining map. Then either $f \equiv 0$ or f is an isomorphism.

Proof: If $\mathrm{Ker}(f) \neq \{0\}$, then by irreducibility of V we have $\mathrm{Ker}(f) = V$ and so $f \equiv 0$.
If $\mathrm{Ker}(f) = \{0\}$ then f is injective, and so $\mathrm{Im}(f) \neq \{0\}$, and thus by irreducibility of W we have $\mathrm{Im}(f) = W$. $\square$

Let us now assume that $\mathbb{K}$ is algebraically closed (for most of this course we take $\mathbb{K} = \mathbb{C}$).
Then we can conclude:

Theorem (also called Schur's lemma).

Let V be an irreducible representation of G and $f: V \to V$ an intertwining map. Then we have $f = \lambda \cdot \mathrm{id}_V$ for some scalar $\lambda \in \mathbb{K}$.

Proof: Pick one eigenvalue λ of f. Then $f - \lambda \cdot \mathrm{id}_V$ is an intertwining map, and $\mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V) \neq \{0\}$, so by irreducibility $\mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V) = V$. $\square$

Corollary (also called Schur's lemma).

Let V and W be irreducible representations of G.
We have
$$\dim\left(\mathrm{Hom}_G(V, W)\right) = \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{if } V \not\cong W. \end{cases}$$

Invariants

Let G be a group, and $\rho: G \longrightarrow Aut(V)$ a representation. Then the subspace (of "invariants")

$$V^G = \{ v \in V \mid \forall g \in G : \rho(g)v = v \} = \{ v \in V \mid g.v = v \ \forall g \in G \}$$

is obviously a subrepresentation.

There is one particularly important case of this: the invariants of the space of linear maps between two representations.

Proposition $\underbrace{Hom(V_1, V_2)^G}_{\text{invariants in the space of linear maps}} = \underbrace{Hom_G(V_1, V_2)}_{\text{intertwining maps}}$

Proof: "$\supset$": Suppose $f: V_1 \to V_2$ is an intertwining map. Then for any $g \in G$

$$(g.f)(v) \overset{(def)}{=} g.f(g^{-1}.v) \overset{(intertw.)}{=} g.g^{-1}.f(v) = f(v).$$

"$\subset$" Suppose $f: V_1 \longrightarrow V_2$ is an invariant in $Hom(V_1, V_2)$. Then for any $g \in G$ and all $w \in V_1$

$$g.f(g^{-1}.w) = f(w)$$

For a given $v \in V_1$ choose $w = g.v$ above to get $g.f(v) = f(g.v)$. $\square$

CHARACTER THEORY FOR REPRESENTATIONS OF FINITE GROUPS

Assume throughout this lecture:

- G a finite group
- $\mathbb{K} = \mathbb{C}$, all vector spaces are complex and linear maps complex-linear
- all representations of interest are finite-dimensional

Recall that under the above assumptions we have

- Any representation V of G is a direct sum $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$ of irreducible subrepresentations $V_1, V_2, \ldots, V_n \subset V$.
  "complete reducibility"
- If V and W are irreducible representations of G which are not isomorphic to each other, then there are no non-zero intertwining maps between them
  "Schur's lemma, part 1"
- If V is an irreducible representation of G, then any intertwining map $V \to V$ is a scalar multiple of the identity, $\lambda \cdot \mathrm{id}_V$, $\lambda \in \mathbb{C}$
  "Schur's lemma, part 2"

The last two properties can be concisely summarized:

$$\dim\left(\mathrm{Hom}_G(V,W)\right) = \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{if } V \ncong W \end{cases}$$

for V, W irreducible.

In this lecture we will show that there are only finitely many different irreducible representations of G (isomorphic ones are identified) and we show how various questions about representations of G can be turned into straightforward calculations with characters.

Def: The <u>character</u> χ_ρ of a representation $\rho: G \longrightarrow Aut(V)$ is the function $\chi_\rho: G \to \mathbb{C}$ given by $\chi_\rho(g) = tr_V(\rho(g)) \quad \forall g \in G$.

Remarks:

- The trace tr_V is well defined, independent of the basis.
- If $g_1, g_2 \in G$ are conjugate, $g_2 = h g_1 h^{-1}$, then the character values are equal:

$$\chi_\rho(g_2) = tr_V(\rho(g_2)) = tr_V(\rho(h g_1 h^{-1}))$$
$$= tr_V(\rho(h)\rho(g_1)\rho(h^{-1})) \underset{\substack{\uparrow \\ \text{cyclicity of trace:} \\ tr(ABC) = tr(BCA)}}{=} tr_V(\rho(g_1))$$
$$= \chi_\rho(g_1)$$

  Therefore characters are constants on each conjugacy class of G. Such functions are called <u>class functions</u>.
- At the neutral element $e \in G$ we have
$$\chi_\rho(e) = tr_V(\rho(e)) = tr_V(id_V) = \dim(V).$$
- The character of the trivial representation $\mathbb{C}$ ($\rho(g) = id_{\mathbb{C}} \; \forall g \in G$) is constant one: $\chi_{triv.}(g) = 1 \quad \forall g \in G$.

Example: Let us consider as an example the symmetric group on three letters, S_3.

In the exercises, you have found the conjugacy classes of S_3:

identity: $\{e\}$

transpositions: $\{(12), (13), (23)\}$

3-cycles: $\{(123), (132)\}$.

Let us consider the following three representations:

- trivial rep $U = \mathbb{C}$, $\rho_U(\sigma) = \mathrm{id}_{\mathbb{C}}$ $\forall \sigma \in S_3$
- alternating rep. $U' = \mathbb{C}$, $\rho_{U'}(\sigma) = \mathrm{sgn}(\sigma) \cdot \mathrm{id}_{\mathbb{C}}$ $\forall \sigma \in S_3$
- a two-dimensional rep. V:

  By realizing that S_3 is isomorphic to the dihedral group D_3 of order 6, we translate the defining representation of D_3 to the following representation of S_3:

$$\rho_V((12)) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \rho_V((123)) = \begin{bmatrix} \cos(\frac{2\pi}{3}) & -\sin(\frac{2\pi}{3}) \\ \sin(\frac{2\pi}{3}) & \cos(\frac{2\pi}{3}) \end{bmatrix}$$

$$= \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}$$

The character values on representatives e, (12), (123) of the conjugacy classes are now

	e	(12)	(123)
χ_U	1	1	1
$\chi_{U'}$	1	-1	1
χ_V	2	0	-1

We next check what the operations that we can perform on representations (direct sum, tensor product, dual, etc.) do to characters.

First observe the following:

<u>Lemma</u> If G is a (finite) group, and $\rho : G \to \mathrm{Aut}(V)$ a (finite-dimensional complex) representation of G, then for each $g \in G$, the linear map $\rho(g): V \to V$ is diagonalizable, and all eigenvalues λ of $\rho(g)$ are roots of unity:

$$\lambda^n = 1 \qquad \text{where} \quad n = \# G \quad \text{is the order of } G.$$

<u>Proof</u>: The order m of $g \in G$ divides the order $n = \# G$ of the group. Now since $g^m = e$, we have

$$\rho(g)^m = \rho(g^m) = \rho(e) = \mathrm{id}_V.$$

Therefore the minimal polynomial of $\rho(g)$ divides the polynomial $x^m - 1$. But the roots of $x^m - 1$ are simple, so the roots of the minimal polynomial are also simple, and therefore $\rho(g)$ is diagonalizable.
It follows also that eigenvalues λ of $\rho(g)$ satisfy $\lambda^m = 1$ and since $m \mid n$, also $\lambda^n = 1$. $\square$

<u>Theorem</u> Let $\rho_V : G \to \mathrm{Aut}(V)$ and $\rho_W : G \to \mathrm{Aut}(W)$ be two representations of G, and χ_V and χ_W their characters. Then we have, $\forall g \in G$

- $\chi_{V^*}(g) = \overline{\chi_V(g)}$ (character of dual rep.)
- $\chi_{V \oplus W}(g) = \chi_V(g) + \chi_W(g)$ (character of direct sum rep.)
- $\chi_{V \otimes W}(g) = \chi_V(g)\chi_W(g)$ (character of tensor product rep.)

<u>Remark</u>: By the Lemma, $\chi_\rho(g)$ is the sum of eigenvalues of $\rho(g): V \to V$ (with multiplicity).

Proof: Denote the dimensions by $n = \dim(V)$, $m = \dim(W)$.

Fix $g \in G$, and diagonalize $\rho_V(g) : V \to V$ by a basis $v_1, \ldots, v_n$ of eigenvectors, with respective eigenvalues $\lambda_1, \ldots, \lambda_n$, so that

$$\rho_V(g)\, v_i = \lambda_i v_i \qquad \forall i = 1, \ldots, n.$$

Let then $v_1^*, \ldots, v_n^*$ be the dual basis of V^*, so that

$$\langle v_j^*, v_i \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Note that by definition of the dual representation

$$\begin{aligned} \langle \rho_{V^*}(g)\, v_j^*, v_i \rangle &= \langle v_j^*, \rho_V(g^{-1}) v_i \rangle = \lambda_i^{-1} \langle v_j^*, v_i \rangle \\ &= \lambda_i^{-1} \delta_{ij} = \lambda_j^{-1} \delta_{ij} \end{aligned}$$

for all i, so we have

$$\rho_{V^*}(g)\, v_j^* = \lambda_j^{-1} v_j^*.$$

But <u>recall</u> that λ_j is a root of unity, so $\lambda_j^{-1} = \bar{\lambda}_j$. Therefore $v_1^*, \ldots, v_n^*$ is a basis of eigenvectors of $\rho_{V^*}(g)$, with respective eigenvalues $\bar{\lambda}_1, \ldots, \bar{\lambda}_n$. Now

$$\chi_{V^*}(g) = \bar{\lambda}_1 + \cdots + \bar{\lambda}_n = \overline{(\lambda_1 + \cdots + \lambda_n)} = \overline{\chi_V(g)}.$$

Similarly, let $w_1, \ldots, w_m$ be a basis of eigenvectors of $\rho_W(g) : W \to W$, with resp. eigenvalues $\mu_1, \ldots, \mu_m$.

Now $v_1, \ldots, v_n, w_1, \ldots, w_m$ is a basis of eigenvectors of $\rho_{V \oplus W}(g)$ on $V \oplus W$, and we get

$$\chi_{V \oplus W}(g) = \lambda_1 + \cdots + \lambda_n + \mu_1 + \cdots + \mu_m = \chi_V(g) + \chi_W(g).$$

Similarly, $(v_i \otimes w_k)_{i=1,\ldots,n;\ k=1,\ldots,m}$ is a basis of $V \otimes W$ and

$$\begin{aligned} \rho_{V \otimes W}(g)\,(v_i \otimes w_k) &= \rho_V(g) v_i \otimes \rho_W(g) w_k \\ &= \lambda_i \mu_k \; v_i \otimes w_k \end{aligned}$$

so $\chi_{V \otimes W}(g) = \sum_{i=1}^{n} \sum_{k=1}^{m} \lambda_i \mu_k = (\sum_i \lambda_i)(\sum_k \mu_k) = \chi_V(g)\, \chi_W(g).$ $\square$

Aside: How to pick the invariants in a representation?

V a representation of G

$V^G = \{v \in V \mid g.v = v \quad \forall g \in G\}$ subrepr. of "invariants".

Define a linear map φ on V by

$$\varphi(v) = \frac{1}{\# G} \sum_{g \in G} g.v \qquad (v \in V).$$

Lemma: The map φ is a projection $V \to V^G$.

Proof: If $v \in V^G$ then $\varphi(v) = \frac{1}{\# G} \sum_{g \in G} v = v$, so $\varphi|_{V^G} = \mathrm{id}_{V^G}$.

Let $h \in G$, $v \in V$. Then with change of var. $k = hg$

$$h.\varphi(v) = \frac{1}{\# G} \sum_{g \in G} hg.v = \frac{1}{\# G} \sum_{k \in G} k.v = \varphi(v),$$

so $\mathrm{Im}(\varphi) \subset V^G$. $\square$

Corollary $\dim(V^G) = \frac{1}{\# G} \sum_{g \in G} \chi_V(g)$

Proof: Evaluate $\mathrm{tr}(\varphi)$ either by projection property (LHS) or directly from definition. $\square$

Proposition For V and W representations of G we have $\dim(\mathrm{Hom}_G(V,W)) = \frac{1}{\# G} \sum_{g \in G} \overline{\chi_V(g)} \cdot \chi_W(g)$

Proof: Recall that $\mathrm{Hom}(V,W) \cong W \otimes V^*$

$\Rightarrow \chi_{\mathrm{Hom}(V,W)}(g) = \chi_W(g)\,\overline{\chi_V(g)}$ by properties of characters.

Recall also that $\mathrm{Hom}_G(V,W) = \mathrm{Hom}(V,W)^G$.

Therefore we get from the previous corollary

$$\dim(\mathrm{Hom}_G(V,W)) = \frac{1}{\# G} \sum_{g \in G} \chi_{\mathrm{Hom}(V,W)}(g) = \frac{1}{\# G} \sum_{g \in G} \overline{\chi_V(g)}\, \chi_W(g).$$

$\square$

Let us agree to use the following natural inner product on the space of functions $G \to \mathbb{C}$:

$$(\varphi, \phi) = \frac{1}{\# G} \sum_{g \in G} \overline{\varphi(g)} \cdot \phi(g) .$$

Then we can rewrite: $\dim(\mathrm{Hom}_G(V,W)) = (\chi_V, \chi_W)$.

We also easily get the following powerful theorem:

<u>Theorem</u>

(i) If V and W are irreducible representations of G then $(\chi_V, \chi_W) = \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{if } V \ncong W \end{cases}$.

(ii) The characters of (non-isomorphic) irreducible representations are linearly independent.

(iii) The number of different (isomorphism classes of) irreducible representations is at most the number of conjugacy classes of G.

<u>Proof</u>: (i) follows from the above observation with Schur's lemma.

(ii) follows from (i), which guarantees orthogonality of irreducible characters.

(iii) follows from linear independence (ii) and the dimension of the space of class functions. $\square$

We proceed with further consequences.

Fix the finite group G, and let $(W_\alpha)_{\alpha=1}^{s}$ be the (finite) collection of all mutually non-isomorphic irreducible representations of G.

Let V be a representation of G, and for all α denote by $m_\alpha \in \mathbb{Z}_{\geq 0}$ the multiplicity of W_α in the decomposition $V = \bigoplus_{\alpha=1}^{s} m_\alpha W_\alpha$ given by complete reducibility.

Theorem

(i) $m_\alpha = (\chi_{W_\alpha}, \chi_V)$ for all $\alpha = 1, \dots, s$.

(ii) The character χ_V determines V up to isomorphism.

(iii) We have $(\chi_V, \chi_V) = \sum_{\alpha=1}^{s} m_\alpha^2$.

(iv) V is irreducible if and only if $(\chi_V, \chi_V) = 1$.

Proof: The character of $V \cong \bigoplus_\alpha m_\alpha W_\alpha$ is $\chi_V = \sum_\alpha m_\alpha \chi_{W_\alpha}$. Now for any $\beta = 1, \dots, s$, note that

$$(\chi_{W_\beta}, \chi_V) = \sum_\alpha m_\alpha \underbrace{(\chi_{W_\beta}, \chi_{W_\alpha})}_{= \delta_{\alpha,\beta} \text{ by previous Thm}} = m_\beta ,$$

which proves (i).

By (i), the character χ_V determines the multiplicity of any irreducible in the decomposition of V, and therefore the isomorphism type of V.

The formula (iii) follows by calculating

$$(\chi_V, \chi_V) = \Big(\sum_\alpha m_\alpha \chi_{W_\alpha}, \sum_\beta m_\beta \chi_{W_\beta}\Big) = \sum_{\alpha,\beta} m_\alpha m_\beta \underbrace{(\chi_{W_\alpha}, \chi_{W_\beta})}_{= \delta_{\alpha\beta}}$$
$$= \sum_\alpha m_\alpha^2 ,$$

and this readily implies (iv). $\square$

The regular representation and group algebra

The group G acts on itself by left multiplication, and from this action we can construct a representation $\mathbb{C}^G$ with a basis $(e_g)_{g\in G}$ by

$$h.e_g = e_{hg} \qquad \forall h\in G,\ g\in G.$$

We denote this representation by $\mathbb{C}[G]$, and call it the (left) regular representation of G.

In fact, $\mathbb{C}[G]$ becomes an algebra (over $\mathbb{C}$) with a bilinear product defined on the basis elements by

$$e_h e_g = e_{hg}.$$

It is easy to calculate the character of $\mathbb{C}[G]$, since in the basis $(e_g)_{g\in G}$ the action of $h\in G$ has no diagonal entries unless $h=e\in G$:

$$\chi_{\mathbb{C}[G]}(h) = \begin{cases} \# G & \text{if } h=e \\ 0 & \text{if } h\neq e. \end{cases}$$

The multiplicity m_α of W_α in the regular representation is

$$m_\alpha = (\chi_{W_\alpha}, \chi_{\mathbb{C}[G]}) = \frac{1}{\# G}\sum_{g\in G} \overline{\chi_{W_\alpha}(g)}\,\chi_{\mathbb{C}[G]}(g)$$

$$= \frac{1}{\# G}\overline{\chi_{W_\alpha}(e)}\cdot \# G = \dim(W_\alpha).$$

Thus any irreducible representation of G appears in the regular representation with multiplicity equal to its dimension:

$$\boxed{\mathbb{C}[G] \cong \bigoplus_\alpha \dim(W_\alpha)\cdot W_\alpha}$$

Equating in particular the dimensions of the two sides we find that $\boxed{\# G = \sum_\alpha \dim(W_\alpha)^2}$.

Example Consider the symmetric group on 4 letters, S_4. It has five conjugacy classes:

- neutral element
- traspositions
- three-cycles
- four-cycles
- products of two disjoint transpositions.

What can we say about representations of S_4?
We know there are at most five different irreducible representations. The trivial and alternating representations are both one-dimensional irreducible. The sum of squares formula above says that

$$\sum_{\alpha} \dim(W_\alpha)^2 = \# S_4 = 4! = 24.$$

Let's subtract the known contributions of trivial and alternating representations:

$$\sum_{\alpha \neq \text{triv}, \text{alt}} \dim(W_\alpha)^2 = 24 - 1^2 - 1^2 = 22$$

and observe that this sum has at most three terms in it. But 22 is not a square, and not a sum of two squares, so we know that there are still 3 other irreducibles. Moreover, the only way of expressing 22 as a sum of 3 squares is $22 = 3^2 + 3^2 + 2^2$, so we know that S_4 has exactly five irreducibles in total, and their dimensions are

$$\underbrace{1,\ 1,}_{\substack{\text{trivial} \\ \text{and alternating}}}\ \underbrace{2,\ 3,\ 3}_{\substack{\text{other} \\ \text{irreducibles}}}.$$

In the exercises you will use the group algebra to show that

<u>Theorem:</u> The number of different irreducible representations of G equals the number of conjugacy classes of G, and their characters form an orthonormal basis for the space of class functions on G.

This suggests that the information about the irreducible representations is concisely summarized in the "<u>character table</u>" of G, which tabulates the values of the characters of irreducible representations on the conjugacy classes, as we did for S_3 in an earlier example:

S_3 character table

	e	transpos.	3-cycle
trivial	1	1	1
altern.	1	-1	1
2-dim	2	0	-1

The rows of the character table are orthonormal w.r.t. the inner product on class functions

$$(\varphi,\phi) = \frac{1}{\# G}\sum_{g\in G}\overline{\varphi(g)}\,\phi(g) = \frac{1}{\# G}\sum_{\text{conj. cl. } C}(\# C)\cdot\overline{\varphi(C)}\,\phi(C).$$

The colums are also orthogonal with respect to the appropriate inner product: for any two conjugacy classes C and D of G we have

$$\sum_{\substack{\alpha\\ \text{irred. rep.}}}\overline{\chi_\alpha(C)}\,\chi_\alpha(D) = \begin{cases}\#G/\#C & \text{if } C=D\\ 0 & \text{if } C\neq D\end{cases}$$

CONTINUOUS SYMMETRIES AND LIE GROUPS

So far we have focused on finite groups appropriate for discrete symmetries.

We now turn to continuous symmetries, and how they are analyzed in terms of infinitesimal transformations. For this, in addition to the group structure, one needs some topology ("continuous") and differentiability ("infinitesimal"). The notion of a Lie group incorporates just that.

Remark: The focus of this course is not geometry, and we do not assume background in differential geometry. The most important examples of Lie groups are matrix groups, with which one can work using ordinary multivariate calculus. To guide towards the general viewpoint, we nevertheless mention general Lie theoretic facts along with their concrete versions for matrix groups.

Def: A Lie group is a group G which is also a smooth manifold, such that the group operations

- multiplication $G \times G \longrightarrow G$, $(g,h) \longmapsto gh$
- inverse $G \longrightarrow G$, $g \longmapsto g^{-1}$

are smooth maps.

"space, which is locally homeomorphic to $\mathbb{R}^m$"

Smooth manifold M (of dimension $m \in \mathbb{N}$)?

- second countable Hausdorff topological space
- atlas: $(U_i, \phi_i)_{i \in I}$, collection of coordinate charts (U_i, ϕ_i)
  - $U_i \subset M$ are open sets, $\bigcup_{i \in I} U_i = M$ ("cover")
  - $\phi_i : U_i \to \mathbb{R}^m$ homeomorphisms to open subsets $\phi_i(U_i) \subset \mathbb{R}^n$
  - transition maps $\phi_i \circ \phi_j^{-1}$ are infinitely differentiable (whenever $U_i \cap U_j \neq \emptyset$ so that the map can be defined) (C^∞ between subsets of $\mathbb{R}^m$)

Smooth map between manifolds M and N?

$f : M \to N$ s.t. whenever

$\phi_i : U_i \to \mathbb{R}^m$ and $\psi_j : V_j \to \mathbb{R}^n$, $U_i \subset M$, $V_j \subset N$

are charts for M and N, resp., then

$\psi_j \circ f \circ \phi_i^{-1}$ is infinitely differentiable (between subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$)

Fundamental example

$GL_n(\mathbb{R})$, the group of invertible $n\times n$ real matrices, is a Lie group.

Check: $GL_n(\mathbb{R}) = \{ M \in \mathbb{R}^{n\times n} \mid \det(M) \neq 0 \}$

Since $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ is continuous, the subset $GL_n(\mathbb{R}) \subset \mathbb{R}^{n\times n}$ is open (as the preimage $\det^{-1}(\mathbb{R}\setminus\{0\})$ of the open set $\mathbb{R}\setminus\{0\}$)

In the neighborhood of any invertible matrix, the n^2 matrix entries define a local homeomorphism to an open subset of $\mathbb{R}^m$, $m = n^2$.

Transition maps are identities, thus C^∞.

Group multiplication is polynomial in matrix entries, thus C^∞. Inverse matrix is rational function in matrix entries, thus C^∞.

Def: A <u>Lie subgroup</u> G' of a Lie group G is a subgroup $G' \subset G$ which is a closed subset

Fact: A Lie subgroup has the structure of a smooth manifold, and becomes a Lie group as well.

Def: A <u>matrix group</u> is a closed subgroup of the Lie group $GL_n(\mathbb{R})$.

As we said, most important examples of Lie groups are matrix groups.

Example $SL_n(\mathbb{R})$ "special linear group"

$$SL_n(\mathbb{R}) = \{ M \in \mathbb{R}^{n\times n} \mid \det(M) = 1 \}$$

$\left(\begin{array}{l} \det(M\tilde{M}) \\ = \det(M)\det(\tilde{M}) \\ = 1 \cdot 1 = 1 \end{array} \right)$

Clearly $SL_n(\mathbb{R}) \subset GL_n(\mathbb{R})$ is a subgroup, and as the preimage of the closed set $\{1\}$ under the continuous map det, it is a closed subset.

Example $GL_n(\mathbb{C})$ "complex general linear group"

$$\overset{"}{=} \{ M \in \mathbb{C}^{n\times n} \mid \det(M) \neq 0 \}$$

Note that we can interpret $GL_n(\mathbb{C}) \subset GL_{2n}(\mathbb{R})$:

Indeed, an n-dimensional $\mathbb{C}$-vect. sp. with basis $e_1, e_2, \dots, e_n$ is a $2n$-dim. $\mathbb{R}$-vect. sp. with basis $e_1, ie_1, e_2, ie_2, \dots, e_n, ie_n$, and $\mathbb{C}$-linear maps are also $\mathbb{R}$-linear, and invertible maps remain invertible.

In practise, this amounts to replacing each entry $z = x + iy$ of the complex matrix by a 2×2 real block $\begin{bmatrix} x & -y \\ y & x \end{bmatrix}$.

Clearly thus $GL_n(\mathbb{C}) \subset GL_{2n}(\mathbb{R})$ is a subgroup, and it is a closed subset (the limit point of a convergent sequence in $GL_n(\mathbb{C}) \subset GL_{2n}(\mathbb{R})$ remains in $GL_n(\mathbb{C})$).

Example O_n and SO_n (orthogonal group and special orthogonal group)

$$O_n = \{ M \in \mathbb{R}^{n \times n} \mid M^T M = \mathbb{1}_{n\times n} \}$$

Note: $M^T M = \mathbb{1} \Longrightarrow M$ is invertible with $M^{-1} = M^T$.

If $M_1, M_2 \in O_n$, then we have $M_1 M_2 \in O_n$, since

$$(M_1 M_2)^T M_1 M_2 = M_2^T \underbrace{M_1^T M_1}_{= \mathbb{1}} M_2 = M_2^T M_2 = \mathbb{1}.$$

Therefore $O_n \subset GL_n(\mathbb{R})$ is a subgroup.

The condition $M^T M$ for a matrix $M = (m_{ij})_{i,j=1}^n$ amounts to the n^2 equations

$$\sum_{k=1}^n m_{ki} m_{kj} = \delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}.$$

The solutions to the (i,j)-th equation is the closed set $f_{ij}^{-1}(\{\delta_{ij}\})$, where $f_{ij} : GL_n(\mathbb{R}) \longrightarrow \mathbb{R}$ is the left hand side of the eqn: $f_{ij}(M) = \sum_k m_{ki} m_{kj}$.

The solution to all equations is the intersection

$$\bigcap_{i,j=1,\dots,n} f_{ij}^{-1}(\{\delta_{ij}\}) \subset GL_n(\mathbb{R}),$$

which is also closed.

The proof that $SO_n = \{ M \in O_n \mid \det(M) = 1 \}$ is a closed subgroup of $GL_n(\mathbb{R})$ is similar.

Example U_n and SU_n (unitary group and special unitary group)

$$U_n = \{ M \in \mathbb{C}^{n\times n} \mid M^\dagger M = \mathbb{1}_{n\times n} \}$$

$$SU_n = \{ M \in \mathbb{C}^{n\times n} \mid M^\dagger M = \mathbb{1}_{n\times n} \text{ and } \det(M) = 1 \}$$

($M^\dagger$ is the conjugate transpose of M)

Similarly, U_n and SU_n are closed in $GL_{2n}(\mathbb{R})$ because the defining equations make them intersections of preimages of closed sets under continuous maps.

Since Lie groups in general, and matrix groups in particular are topological spaces, we may ask whether they are e.g.

- (path-) connected
  - for any $g_1, g_2 \in G$, does there exist a continuous path $\gamma : [0,1] \longrightarrow G$ such that $\gamma(0) = g_1$ and $\gamma(1) = g_2$?
- simply connected
  - for any loop, i.e. a continuous path $\gamma : [0,1] \longrightarrow G$ with $\gamma(0) = \gamma(1) = g \in G$ can γ be continuously deformed ($[0,1] \ni s \longmapsto \gamma^{(s)}$ loop, $\gamma^{(0)} = \gamma$, $\gamma^{(1)} =$ constant loop) to the constant loop $\bar{\gamma}$, $\bar{\gamma}(t) \equiv g \quad \forall t \in [0,1]$?
- compact
  - does every sequence $(g_n)_{n \in \mathbb{N}}$ of elements $g_n \in G$ contain a subsequence $(g_{n_k})_{k \in \mathbb{N}}$, $n_k \xrightarrow[k \to \infty]{} \infty$, which converges in G, $g_{n_k} \xrightarrow[k \to \infty]{} g \in G$?

Connectedness, simply-connectedness and compactness turn out to be very important in the study of Lie groups.

Compactness examples

- $GL_n(\mathbb{R})$, $SL_n(\mathbb{R})$, $GL_n(\mathbb{C})$ are not compact for $n \geq 2$.
  - ▶ Consider e.g. the sequence $(g_k)_{k=1}^{\infty}$, where g_k is the diagonal matrix with entries $k, k^{-1}, 1, 1, \dots, 1$. It has no convergent subsequences.

- O_n, SO_n, U_n, SU_n are compact
  - ▶ From the defining equations one sees that the rows of these matrices are orthonormal. In particular these are bounded subsets in $\mathbb{R}^{n\times n}$ or $\mathbb{C}^{n\times n}$. They are also closed subsets. By Heine-Borel theorem, they are thus compact.

Connectedness examples

- $GL_n(\mathbb{R})$ is not connected
  - ▶ Suppose, by contrapositive, that there exists a continuous path $\gamma : [0,1] \to GL_n(\mathbb{R})$ such that $\gamma(0) = \mathbb{1}$ and $\gamma(1) = g$ with $\det(g) < 0$ (e.g. g diagonal with entries $-1, +1, +1, +1, \dots$). But determinant is continuous, so the continuous function $t \mapsto \det(\gamma(t))$ must have a zero at some $t \in (0,1)$, contradicting $\gamma(t) \in GL_n(\mathbb{R})$.
- $GL_n(\mathbb{C})$ is connected
  - ▶ Let $g_1, g_2 \in GL_n(\mathbb{C})$. The polynomial of z defined by $z \mapsto \det(z \cdot g_2 + (1-z) g_1)$ has finitely many zeroes in the complex plane. Therefore there exists a path from 0 to 1 in $\mathbb{C}$, $[0,1] \ni t \mapsto z(t) \in \mathbb{C}$ avoiding those zeroes. Then $\gamma(t) = z(t) \cdot g_2 + (1 - z(t)) g_1$ connects g_1 to g_2 in $GL_n(\mathbb{C})$.

The topology of SU_2

Theorem: The Lie group $SU_2 = \{ M \in \mathbb{C}^{2\times 2} \mid M^\dagger M = \mathbb{1} , \det(M) = 1 \}$ is homeomorphic to the three-sphere $S^3 \subset \mathbb{R}^4$. In particular SU_2 is compact, connected, and simply connected.

Proof: Write a 2×2 complex matrix as $M = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, $\alpha, \beta, \gamma, \delta \in \mathbb{C}$. Consider the requirements for $M \in SU_2$, they can be written in terms of entries $\alpha, \beta, \gamma, \delta$ as follows.

Calculate

$$M^\dagger M = \begin{bmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} |\alpha|^2 + |\gamma|^2 & \bar{\alpha}\beta + \bar{\gamma}\delta \\ \alpha\bar{\beta} + \gamma\bar{\delta} & |\beta|^2 + |\delta|^2 \end{bmatrix}.$$

The condition

$$M^\dagger M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

amounts to four complex equations for $\alpha, \beta, \gamma, \delta$.

The equality of $(1,1)$ entries reads $|\alpha|^2 + |\gamma|^2 = 1$, meaning that $\begin{bmatrix} \alpha \\ \gamma \end{bmatrix}$ should be a unit vector. Similarly the equality of $(2,2)$ entries requires $\begin{bmatrix} \beta \\ \delta \end{bmatrix}$ to be unit vec. The equalities of $(1,2)$ and $(2,1)$ entries are equivalent (complex conjugate equations) and each says that $\bar{\alpha}\beta + \bar{\gamma}\delta = 0$, i.e. that the vectors $\begin{bmatrix} \alpha \\ \gamma \end{bmatrix}$ and $\begin{bmatrix} \beta \\ \delta \end{bmatrix}$ are orthogonal.

Since the orthogonal complement of $\begin{bmatrix} \alpha \\ \gamma \end{bmatrix}$ is spanned by $\begin{bmatrix} -\bar{\gamma} \\ \bar{\alpha} \end{bmatrix}$, we must have $\begin{bmatrix} \beta \\ \delta \end{bmatrix} = s \begin{bmatrix} -\bar{\gamma} \\ \bar{\alpha} \end{bmatrix}$ for some $s \in \mathbb{C}$.

Assuming this, we calculate $\det(M) = \det \begin{bmatrix} \alpha & -s\bar{\gamma} \\ \gamma & s\bar{\alpha} \end{bmatrix} = s(|\alpha|^2 + |\gamma|^2) = s$ so the requirement $\det(M) = 1$ fixes $s = 1$.

We find that any $M \in SU_2$ is of the form $M = \begin{bmatrix} \alpha & -\bar{\gamma} \\ \gamma & \bar{\alpha} \end{bmatrix}$ for some $\alpha, \gamma \in \mathbb{C}$ with $|\alpha|^2 + |\gamma|^2 = 1$.

If we write $\alpha = x_1 + i x_2$ and $\gamma = x_3 + i x_4$ with $x_1, x_2, x_3, x_4 \in \mathbb{R}$ the condition $|\alpha|^2 + |\gamma|^2 = 1$ reads $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$, i.e. $[x_1\ x_2\ x_3\ x_4]^T \in S^3 \subset \mathbb{R}^4$.

Conversely, for $[x_1\ x_2\ x_3\ x_4]^T \in S^3 \subset \mathbb{R}^4$ the matrix

$$M = \begin{bmatrix} x_1 + i x_2 & -x_3 + i x_4 \\ x_3 + i x_4 & x_1 - i x_2 \end{bmatrix} \quad \text{is in } SU_2.$$

This formula thus defines a bijection $S^3 \to SU_2$, and continuity (and in fact smoothness) of this map is clear from the expression, so S^3 is homeomorphic (in fact diffeomorphic) to SU_2.

It is known that the unit sphere S^3 in $\mathbb{R}^4$ is compact, connected and simply connected, so we get the same topological properties for SU_2. $\square$

Intrinsic definition of orthogonal groups

We have defined O_n as the set of $n \times n$ real matrices such that $M^T M = \mathbb{1}$, but this is not the intrinsic definition which would realize SO_n as the symmetry group of some structure. We now turn to the intrinsic, coordinate independent definition.

Let V be a real vector space of dimension n, equipped with an inner product, i.e. a positive definite symmetric bilinear form $\beta : V \times V \to \mathbb{R}$. The group $O(V)$ is the symmetry group of the inner product space (V, β) — it consists of all transformations that preserve this structure: preserving the vector space structure means considering linear transformations $T: V \to V$, and preserving the inner product means that $\beta(v,w) = \beta(Tv, Tw)$ for all $v,w \in V$. Thus:

$$O(V) = \{ T : V \to V \text{ linear s.t. } \beta(v,w) = \beta(Tv, Tw) \ \forall v,w \in V \}$$

Another virtue of this definition is that it is coordinate independent. Occasionally, however, it is practical to work in specific coordinates, and this can be achieved as follows. Choose a basis $v_1, \dots, v_n \in V$ and set $B_{ij} = \beta(v_i, v_j)$. Then vectors $v = \sum_{i=1}^n x_i v_i$ $w = \sum_{j=1}^n y_j v_j$ have inner product

$$\beta(v,w) = \sum_{i,j=1}^n x_i B_{ij} y_j = x^T B y .$$

When the linear map $T : V \to V$ is represented as matrix $T v_i = \sum_{k=1}^n M_{ki} v_k$, we have

$$\beta(Tv, Tw) = \sum_{i,j,k,l=1}^n x_i M_{ki} B_{kl} M_{lj} y_j = x^T M^T B M y .$$

The invariance of β under T holds for all vectors if and only if $M^T B M = B$.

The standard basis of $\mathbb{R}^n$ and usual inner product correspond to $B = \mathbb{1}$, in which case the matrix M of T should satisfy $M^T M = \mathbb{1}$.

Groups preserving bilinear forms

Again, let V be a $\mathbb{R}$-vector space, and let $\beta: V \times V \to \mathbb{R}$ be a bilinear form. We often consider the symmetry groups of such structures, i.e. the subgroups of $\mathrm{Aut}(V)$ consisting of $T: V \to V$ such that $\beta(Tv, Tw) = \beta(v,w)$ $\forall v,w \in V$.

- If β is symmetric and positive definite, then the symmetry group is $O(V)$, "orthogonal group".
- If β is symmetric and has k positive and l negative eigenvalues, the symmetry group is still called "orthogonal group (with signature (k,l))".

<u>Ex</u>: In physics, special relativity concerns with transformations preserving the metric of Minkowski space-time, which is a bilinear form of signature $(k,l) = (1,3)$. The standard choice of β is $\beta([t\ x\ y\ z]^T, [t'\ x'\ y'\ z']) = tt' - xx' - yy' - zz'$. The corresponding group $O_{1,3}$ is called the "Lorentz group".

- If β is skew-symmetric and non-degenerate, i.e. a "symplectic form", then the symmetry group is denoted $Sp_n(\mathbb{R})$ and called "symplectic group" (Note: n must be even for such β to exist)

Finally, the following familiar case is a slight modification:

- Let V be a $\mathbb{C}$-vector space instead, and let β be a positive definite Hermitian form (linear in second, and conjugate linear in first argument). Then the symmetry group is "unitary group".

INFINITESIMAL TRANSFORMATIONS AND THE LIE ALGEBRA OF A LIE GROUP

The study of continuous symmetries may appear complicated, as the notion of a Lie group involves simultaneously algebra, geometry and topology. The basic reason why it is nevertheless tractable, is that by studying only infinitesimal symmetry transformations, we manage to linearize the problem, and one can systematically go back to the full Lie group from this linearization. The Lie algebra of the Lie group is precisely the infinitesimal, linearized version of the symmetry.

The first indication that an infinitesimal neighborhood of the neutral element contains all relevant information is:

Exercise Let G be a connected Lie group, and $U \subset G$ an open neighborhood of the neutral element $e \in G$. Show that the subgroup generated by elements in U is the entire group G.

We will see that under certain topological assumptions on G, "elements infinitesimally close to the neutral element" (the Lie algebra $\mathfrak{g}$ of G) contain all relevant information. As a sort of summary:

First principle

If G is connected, then any homom. $\varphi : G \to H$ is determined by its differential $\lambda : \mathfrak{g} \to \mathfrak{h}$, which is a Lie algebra homom.

Second principle

If G is connected and simply connected, then for any Lie alg. homom. $\lambda : \mathfrak{g} \to \mathfrak{h}$ there exists a unique homom. $\varphi : G \to H$.

As usual, infinitesimals are mathematically described by derivatives. For a smooth manifold M and a point $p \in M$ on it, the directional derivatives at p form a vector space T_pM, called the tangent space at p.

Consider smooth functions $f : M \to \mathbb{R}$.

Let $\gamma : (-\varepsilon, \varepsilon) \to M$ be a smooth path through $p = \gamma(0) \in M$. The directional derivative of f in the direction of γ is

$$\frac{d}{dt} f(\gamma(t))\Big|_{t=0} .$$

We denote by $\dot{\gamma}(0)$ this operation $C^\infty(M) \longrightarrow \mathbb{R}$. The tangent space T_pM is the span of such directional derivatives for all paths γ through p.

If M and N are two smooth manifolds, and $\varphi : M \to N$ a smooth map, then the derivative of φ at $p \in M$

$$d\varphi|_p : T_pM \longrightarrow T_{\varphi(p)}N$$

is defined by setting, for $f : N \to \mathbb{R}$ and $X \in T_pM$

$$\big((d\varphi|_p)(X)\big)(f) = X(f \circ \varphi),$$

which corresponds to mapping paths γ on M through $p \in M$ by φ to paths $\varphi \circ \gamma$ on N through $f(p) \in N$.

On a Lie group G, one can identify tangent spaces T_gG at different points $g \in G$:

left multiplication by g: $\quad L_g : G \to G, \; h \mapsto gh$ (smooth map)

derivative at neutral element: $(dL_g)|_e : T_eG \to T_gG$

note: $L_{g_1} \circ L_{g_2} = L_{g_1 g_2}$ and $(dL_{g_1})|_{g_2} \circ (dL_{g_2})|_e = (dL_{g_1 g_2})|_e$

Denote $\mathfrak{g} = T_eG$ ("Lie algebra of G") and identify T_gG with $\mathfrak{g}$ via $(dL_g)|_e$ ("left invariant vector fields")

The Lie algebra, one-parameter subgroups, and exponential map

G Lie group, $\mathfrak{g} = T_eG$ its Lie algebra

Def: A smooth path $\gamma : \mathbb{R} \to G$ is called a *one-parameter subgroup* in G if $\gamma(t+s) = \gamma(t)\gamma(s)$ for all $t, s \in \mathbb{R}$. ($\gamma : \mathbb{R} \to G$ homomorphism of Lie groups)

Let $X = \dot{\gamma}(0)$ denote the derivative of γ, so $X \in \mathfrak{g} = T_eG$.

Note that for any $t \in \mathbb{R}$

$$\begin{aligned}\frac{d}{dt}\gamma(t) &= \frac{d}{ds}\big(\gamma(t+s)\big)\Big|_{s=0} = \frac{d}{ds}\big(\gamma(t)\gamma(s)\big)\Big|_{s=0} \\ &= \frac{d}{ds}\Big(L_{\gamma(t)}(\gamma(s))\Big)\Big|_{s=0} = (dL_{\gamma(t)})\big|_e(\dot{\gamma}(0)) \\ &= (dL_{\gamma(t)})\big|_e(X) \quad =: \tilde{X}.\end{aligned}$$

via the identification "left invariant vector fields on G"

Conversely, given $X \in \mathfrak{g}$, the differential equation

$$\frac{d}{dt}\gamma(t) = (dL_{\gamma(t)})(X) =: \tilde{X}, \qquad \gamma(0) = e \in G,$$

has a solution (by existence of solutions to ODEs) and it satisfies $\gamma(t+s) = \gamma(t)\gamma(s)$ (by uniqueness of solutions to ODEs: both sides satisfy the same eq. by virtue of $L_{g_1} \circ L_{g_2} = L_{g_1 g_2}$).

Therefore, the Lie algebra $\mathfrak{g}$ can be identified with one-parameter subgroups $\mathfrak{g} \ni X \longleftrightarrow \gamma^X$.

$$\left(\frac{d}{dt}\gamma^X(t) = \tilde{X}\right)$$

Also, by uniqueness of ODE solutions,

$$\gamma^{\lambda X}(t) = \gamma^X(\lambda t) \quad \text{for any } \lambda \in \mathbb{R}.$$

We denote $\gamma^X(t) = \exp(tX)$

Exponential map: $\exp : \mathfrak{g} \longrightarrow G$

By construction $\exp((t+s)X) = \exp(tX)\exp(sX)$.

Matrix exponential

For $X \in \mathbb{R}^{n\times n}$, recall that

$$e^X = \sum_{n=0}^{\infty} \frac{1}{n!} X^n \quad \text{is a convergent series}$$

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

Thus $\mathbb{R} \ni t \longmapsto e^{tX} \in GL_n(\mathbb{R})$ is a one-parameter subgroup. Indeed also $\gamma(t) = e^{tX}$ is the unique solution of the ODE

$$\frac{d}{dt}\gamma(t) = \gamma(t) X \qquad \gamma(0) = \mathbb{1}$$

and the RHS can be understood as mapping the tangent vector X at $\mathbb{1} \in GL_n(\mathbb{R})$ by left multiplication to a vector at $\gamma(t) \in GL_n(\mathbb{R})$.

This identifies the Lie algebra $\mathfrak{gl}_n(\mathbb{R})$ of $GL_n(\mathbb{R})$ with the set of all $n\times n$ real matrices

$$\mathfrak{gl}_n(\mathbb{R}) = \mathbb{R}^{n\times n}$$

The exponential map is the matrix exponential

$$\exp : \mathfrak{gl}_n(\mathbb{R}) \longrightarrow GL_n(\mathbb{R})$$
$$X \longmapsto e^X .$$

For all matrix Lie groups, i.e. closed subgroups $G \subset GL_n(\mathbb{R})$, the exponential map is still the matrix exponential, defined on the tangent space $\mathfrak{g} = T_{\mathbb{1}} G \subset T_{\mathbb{1}} GL_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R})$ of the matrix group at $\mathbb{1}$.

We can thus characterize the Lie algebras of matrix groups $G \subset GL_n(\mathbb{R})$ as

$$\mathfrak{g} = \{ X \in \mathbb{R}^{n\times n} \mid e^{tX} \in G \ \ \forall t \in \mathbb{R} \} .$$

Exercise Show that $\det(e^{tX}) = e^{t\cdot tr(X)}$, for all $t \in \mathbb{R}$, $X \in \mathbb{R}^{n\times n}$.

Example Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ of $SL_n(\mathbb{R})$.

Recall: $SL_n(\mathbb{R}) = \{ M \in \mathbb{R}^{n\times n} \mid \det(M) = 1 \}$.

Now for $X \in \mathbb{R}^{n\times n}$, we have $\det(e^{tX}) = e^{t\cdot tr(X)}$, and therefore $e^{tX} \in SL_n(\mathbb{R})\ \forall t \iff tr(X) = 0$.

We conclude

$$\mathfrak{sl}_n(\mathbb{R}) = T_{\mathbb{1}} SL_n(\mathbb{R}) = \{ X \in \mathbb{R}^{n\times n} \mid tr(X) = 0 \}$$

= the set of all traceless $n\times n$ real matrices.

(The exponential of complex matrices works the same way!

Example: The Lie algebra of $GL_n(\mathbb{C})$ is

$\mathfrak{gl}_n(\mathbb{C}) = \mathbb{C}^{n\times n}$ = the set of all $n\times n$ complex matrices.

Example The Lie algebra of $SL_n(\mathbb{C}) = \{ M \in \mathbb{C}^{n\times n} \mid \det(M) = 1 \}$ is $\mathfrak{sl}_n(\mathbb{C}) = \{ X \in \mathbb{C}^{n\times n} \mid tr(X) = 0 \}$

= the set of traceless $n\times n$ complex matrices.)

Remark: The dimension of G is the dimension $\dim(T_e G)$.

Note: We get the dimensions of these Lie groups, since

$\dim(\mathfrak{gl}_n(\mathbb{R})) = n^2$ (as we knew by construction)

$\dim(\mathfrak{sl}_n(\mathbb{R})) = n^2 - 1$

(one linear condition $tr(X) = 0$ reduces the dimension by 1 — e.g. all other entries but the (n,n)-entry can be freely chosen).

Example Lie algebra of orthogonal group.

Recall: $O_n = \{ M \in \mathbb{R}^{n\times n} \mid M^T M = \mathbb{1} \}$

$SO_n = \{ M \in \mathbb{R}^{n\times n} \mid M^T M = \mathbb{1} \text{ and } \det(M) = 1 \}$

Claim: For $X \in \mathbb{R}^{n\times n}$ we have:

$$\forall t \in \mathbb{R}\ e^{tX} \in O_n \iff \forall t \in \mathbb{R}\ e^{tX} \in SO_n \iff X^T = -X$$

Pf: We prove "first" $\Rightarrow$ "third" $\Rightarrow$ "second" $\Rightarrow$ "first".

"$2^{nd} \Rightarrow 1^{st}$" Obvious since $SO_n \subset O_n$.

"$1^{st} \Rightarrow 3^{rd}$" Assume $e^{tX} \in O_n$ $\forall t$, i.e.

$$\begin{aligned} \mathbb{1} &= (e^{tX})^T e^{tX} \\ &= e^{tX^T} e^{tX} = (\mathbb{1} + tX^T + O(t^2))(\mathbb{1} + tX + O(t^2)) \\ &= \mathbb{1} + t\cdot(X^T + X) + O(t^2) \end{aligned}$$

The equality of coefficients of t requires $X^T = -X$.

"$3^{rd} \Rightarrow 2^{nd}$" Assume $X^T = -X$. Then for any t,

$$(e^{tX})^T e^{tX} = e^{tX^T} e^{tX} = e^{-tX} e^{tX} = \mathbb{1}.$$

Moreover, if $X^T = -X$, then all diagonal entries of X vanish and $\mathrm{tr}(X) = 0$.

Therefore $\det(e^{tX}) = e^{t\cdot \mathrm{tr}(X)} = e^{t\cdot 0} = 1$. $\square$

Therefore the Lie algebras of O_n and SO_n are

$$\underline{so}_n = \{ X \in \mathbb{R}^{n\times n} \mid X^T = -X \}$$

= the set of antisymmetric $n\times n$ real matrices.

Note: $\dim(\underline{so}_n) = \binom{n}{2} = \frac{n(n-1)}{2}$

An antisymmetric matrix $X \in \mathbb{R}^{n\times n}$ ($X^T = -X$) is determined by its strictly upper triangular entries X_{ij}, $1 \le i < j \le n$. The number of these equals $\binom{n}{2}$.

Example Exponentials $\underline{so}_3 \longrightarrow SO_3$

The Lie algebra $\underline{so}_3$ of SO_3 is three-dimensional (recall: $\dim(\underline{so}_n) = \frac{1}{2}n(n-1)$, so $\dim(\underline{so}_3) = \frac{1}{2}3\cdot 2 = 3$), consisting of antisymmetric 3×3 matrices.

As a basis we may take R_x, R_y, R_z given by

$$R_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{bmatrix}, \quad R_y = \begin{bmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad R_z = \begin{bmatrix} 0 & -1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Straightforward calculations give

$$R_x^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad R_x^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & +1 \\ 0 & -1 & 0 \end{bmatrix} = -R_x$$

and we see that the powers of R_x repeat modulo 4:

$$R_x^{4m+1} = -R_x^{4m+3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{bmatrix}$$
$$R_x^{4m+2} = -R_x^{4m+4} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{for } m \in \mathbb{Z}_{\geq 0}.$$

We can calculate the matrix exponential

$$e^{tR_x} = \mathbb{1} + \sum_{n=1}^{\infty} \frac{t^n}{n!} R_x^n = \mathbb{1} + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \cdots\right) R_x + \left(\frac{t^2}{2!} - \frac{t^4}{4!} + \frac{t^6}{6!} - \cdots\right) R_x^2$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(t) & -\sin(t) \\ 0 & \sin(t) & \cos(t) \end{bmatrix}.$$

This is the matrix of rotation around the x-axis by angle t in the positive direction.

The one parameter subgroup $t \mapsto e^{tR_x}$ is the subgroup in SO_3 consisting of rotations around x-axis.

Similarly

$$e^{tR_y} = \begin{bmatrix} \cos(t) & 0 & \sin(t) \\ 0 & 1 & 0 \\ -\sin(t) & 0 & \cos(t) \end{bmatrix}, \quad e^{tR_z} = \begin{bmatrix} \cos(t) & -\sin(t) & 0 \\ \sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

are rotations around y- and z-axes.

Example Lie algebra of unitary group

Recall: $U_n = \{ M \in \mathbb{C}^{n\times n} \mid M^\dagger M = \mathbb{1} \}$

$SU_n = \{ M \in \mathbb{C}^{n\times n} \mid M^\dagger M = \mathbb{1} \text{ and } \det(M) = 1 \}$

Claim For $X \in \mathbb{C}^{n\times n}$ we have

(i) $\forall t \in \mathbb{R}$ $e^{tX} \in U_n \iff X^\dagger = -X$

(ii) $\forall t \in \mathbb{R}$ $e^{tX} \in SU_n \iff X^\dagger = -X$ and $\mathrm{tr}(X) = 0$

Pf: Similar to O_n and SO_n. The difference is that a complex anti-Hermitian matrix X, $X^\dagger = -X$ is not automatically traceless. The diagonal entries must have vanishing real part, but their imaginary parts can be arbitrary.

Therefore the Lie algebras $\underline{u}_n$ of U_n and $\underline{su}_n$ of SU_n are

$$\underline{u}_n = \{ X \in \mathbb{C}^{n\times n} \mid X^\dagger = -X \}$$

= the set of anti-Hermitian $n\times n$ complex matrices

$$\underline{su}_n = \{ X \in \mathbb{C}^{n\times n} \mid X^\dagger = -X \text{ and } \mathrm{tr}(X) = 0 \}$$

= the set of traceless anti-Hermitian complex matrices

Note:
$$\dim(\underline{u}_n) = 2\cdot\binom{n}{2} + n = 2\,\frac{n(n-1)}{2} + n = n^2$$

An anti-Hermitian matrix X has arbitrary complex numbers as its strictly upper triangular entries ($2\cdot\binom{n}{2}$ real parameters) and arbitrary imaginary numbers as its diagonal entries (n real param.).

$$\dim(\underline{su}_n) = \dim(\underline{u}_n) - 1 = n^2 - 1$$

The condition $\mathrm{tr}(X) = 0$ determines one of the diagonal entries in terms of the other $n-1$.

The exponential map $\exp : \mathfrak{g} \to G$ provides the way to lift "infinitesimal elements" in $\mathfrak{g} = T_e G$ to elements in the Lie group G. A crucial property of it is:

<u>Fact</u>: There exists an open neighborhood $U \subset G$ of the neutral element $e \in G$, and an open neighborhood $V \subset \mathfrak{g}$ of zero $0 \in \mathfrak{g}$, such that

$$\exp : V \longrightarrow U \quad \text{is a bijection.}$$

<u>Def</u>: A <u>homomorphism of Lie groups</u> is a smooth map $\varphi : G \longrightarrow H$ from a Lie group G to a Lie group H, which is a group homomorphism.

<u>Def</u>: A (real or complex) <u>representation of a Lie group</u> G on a finite-dimensional real or complex vector space V is a Lie group homomorphism

$$\rho : G \longrightarrow \mathrm{Aut}(V).$$

Recall: If V is an $\mathbb{R}$-vect. sp. of $\dim_{\mathbb{R}}(V) = n$, then $\mathrm{Aut}(V) \cong GL_n(\mathbb{R})$ is a matrix Lie group.

If V is a $\mathbb{C}$-vect. sp. of $\dim_{\mathbb{C}}(V) = n$, then $\mathrm{Aut}(V) \cong GL_n(\mathbb{C}) \subset GL_{2n}(\mathbb{R})$ is a matrix Lie group.

THE FIRST PRINCIPLE OF INFINITESIMAL ELEMENTS

<u>Theorem</u> If G is a connected Lie group, then any Lie group homomorphism $\varphi : G \longrightarrow H$ is uniquely determined by $d\varphi|_e : \mathfrak{g} \longrightarrow \mathfrak{h}$.

This says that the "infinitesimal elements" $X \in \mathfrak{g} = T_e G$ contain a sufficient amount of information about homomorphisms, and thus in particular about representations of G.

Lemma If $\gamma : \mathbb{R} \to G$ is a one-parameter subgroup in G with $\dot{\gamma}(0) = X \in \mathfrak{g}$, and $\varphi : G \to H$ is a Lie group homomorphism, then $\tilde{\gamma} = \varphi \circ \gamma : \mathbb{R} \to H$ is a one-param. subgroup in H with $\dot{\tilde{\gamma}}(0) = (d\varphi|_e)(X) \in \mathfrak{h} = T_e H$.

Proof of lemma:
$$\tilde{\gamma}(t+s) = \varphi(\gamma(t+s)) = \varphi(\gamma(t)\gamma(s))$$
$$= \varphi(\gamma(t))\,\varphi(\gamma(s)) = \tilde{\gamma}(t)\,\tilde{\gamma}(s).$$
$$\frac{d}{dt}\tilde{\gamma}(t)\Big|_{t=0} = \frac{d}{dt}\big(\varphi(\gamma(t))\big)\Big|_{t=0} = (d\varphi|_e)(\dot{\gamma}(0)) = (d\varphi|_e)(X). \quad \square$$

Corollary: For a Lie group homomorphism $\varphi : G \to H$ we have
$$\varphi(\exp(X)) = \exp\big((d\varphi|_e)(X)\big) \quad \forall X \in \mathfrak{g}.$$

Proof of theorem: Let $U \subset G$ be a neighborhood of $e \in G$ s.t. $\exp : V \to U$ is bijective, with $V \subset \mathfrak{g}$.
Recall that the connected Lie group G is generated by elements in the neighborhood U of e.
Now, for any $g \in G$, write $g = g_1 \cdots g_n$, with $g_1, \dots, g_n \in U$, and write $g_j = \exp(X_j)$ for $X_j \in \mathfrak{g}$. Then $\varphi(g_j) = \varphi(\exp(X_j)) = \exp((d\varphi|_e)(X_j))$ is determined by $d\varphi|_e : \mathfrak{g} \to \mathfrak{h} = T_e H$.
But then $\varphi(g) = \varphi(g_1 \cdots g_n) = \varphi(g_1) \cdots \varphi(g_n)$ is also. $\square$

Matrix exponential, logarithm, and Baker-Campbell-Hausdorff formula

Let us see concretely for matrix Lie groups that the exponential map is locally bijective near zero.

For $X \in \mathbb{R}^{n\times n}$ (or more generally for $X \in \mathbb{C}^{n\times n}$) the series $\exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k$ is convergent in the operator norm $\|A\|_{op} = \sup_{v\in\mathbb{R}^n} \frac{\|Av\|}{\|v\|}$.

The series $\log(\mathbb{1}+X) = \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell-1}}{\ell} X^\ell$ is convergent when $\|X\|_{op} < 1$, and then $\exp(\log(\mathbb{1}+X)) = \mathbb{1}+X$.

Thus, in $G = GL_n(\mathbb{R})$ with $\mathfrak{g} = \mathfrak{gl}_n(\mathbb{R})$ we can set $U = \{M \in \mathbb{R}^{n\times n} \mid \|M-\mathbb{1}\|_{op} < 1\} \subset GL_n(\mathbb{R})$ and $V = \log(U) \subset \mathbb{R}^{n\times n} = \mathfrak{gl}_n(\mathbb{R})$, and we have that

$$\exp : \underset{\mathfrak{gl}_n(\mathbb{R})}{\underset{\cap}{V}} \longrightarrow \underset{GL_n(\mathbb{R})}{\underset{\cap}{U}} \quad \text{is bijective}$$

__Example:__ Let $X, Y \in \mathbb{R}^{n\times n}$, with $\|X\|_{op}, \|Y\|_{op} \leq M$ for some $M\in\mathbb{R}$.

Then
$$e^X e^Y = \left(\mathbb{1} + X + \tfrac{1}{2}X^2 + \mathcal{O}(\|X\|_{op}^3)\right)\left(\mathbb{1} + Y + \tfrac{1}{2}Y^2 + \mathcal{O}(\|Y\|_{op}^3)\right)$$
$$= \mathbb{1} + X + Y + \tfrac{1}{2}X^2 + XY + \tfrac{1}{2}Y^2 + \mathcal{O}(\max\{\|X\|_{op}, \|Y\|_{op}\}^3)$$

If $\|X\|_{op}, \|Y\|_{op}$ are sufficiently small, then $\|\mathbb{1} - e^X e^Y\|_{op} < 1$ and we have
$$\log(e^X e^Y) = X + Y + \tfrac{1}{2}X^2 + XY + \tfrac{1}{2}Y^2 - \tfrac{1}{2}(X+Y)^2 + \mathcal{O}((\cdots)^3)$$
$$= X + Y + \tfrac{1}{2}XY - \tfrac{1}{2}YX + \mathcal{O}((\cdots)^3)$$
$$= X + Y + \tfrac{1}{2}[X,Y] + \mathcal{O}((\cdots)^3)$$

Baker-Campbell-Hausdorff formula

__Exercise:__ Show that $\log(e^X e^Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] + \frac{1}{12}[Y,[Y,X]] + \dots$

Recall: For G a Lie group, the Lie algebra $\mathfrak{g}$ can be viewed in three different ways:

- as the tangent space $\mathfrak{g} = T_e G$ at neutral element
- as left-invariant vector fields on G by identification $T_g G \cong T_e G = \mathfrak{g}$ via $(dL_g)|_e$
- as one parameter subgroups, homomorphisms $\gamma : \mathbb{R} \to G$ via $\gamma \mapsto \dot{\gamma}(0) \in T_e G = \mathfrak{g}$

  $(\gamma(t) = \exp(tX)$ with $X = \dot{\gamma}(0) \in \mathfrak{g})$

LIE BRACKET

Fact: exp is bijective from a neighborhood V of 0 in $\mathfrak{g}$ to a neighborhood U of e in G.

Denote by $\log : U \to V$ the inverse.

Baker-Campbell-Hausdorff formula:

For $X, Y \in \mathfrak{g}$ in some neighborhood of 0, we have $\exp(X)\exp(Y) \in U$ so that $\log(\exp(X)\exp(Y))$ is defined. Denote

$\mu(X,Y) = \log(\exp(X)\exp(Y))$ i.e. $\exp(X)\exp(Y) = \exp(\mu(X,Y))$

BCH: $\mu(tX, tY) = t\cdot(X+Y) + t^2 \cdot \beta(X,Y) + \mathcal{O}(t^3) + \dots$

where β is a bilinear antisymmetric map

(We checked this for matrix Lie groups with $\beta(X,Y) = \frac{1}{2}XY - \frac{1}{2}YX = \frac{1}{2}[X,Y]$.)

Remark: Given that β is bilinear (see adjoint action, below) antisymmetry can be seen as follows:

Let $X \in \mathfrak{g}$. Then $\mu(tX, sX) = \log(\exp(tX + sX))$

$= \log(\exp((t+s)X)) = (t+s)X$

so $\beta(X,X) = 0$.

Then by polarization, for $X, Y \in \mathfrak{g}$:

$$\beta(X,Y) = \underbrace{\beta(X+Y, X+Y)}_{=0} - \underbrace{\beta(X,X)}_{=0} - \beta(Y,X) - \underbrace{\beta(Y,Y)}_{=0} = -\beta(Y,X)$$

For $\varphi: G \to H$ a homom. we have

We can use the BCH formula to define the Lie bracket : $[X,Y] = 2 \cdot \beta(X,Y)$.

We will give another approach, which would give an alternative definition, which is sometimes preferable. It is based on adjoint actions.

Adjoint actions :

Let G be a Lie group, and $\mathfrak{g} = T_e G$ its Lie algebra.

The group acts on itself by conjugation :

$$C_g(h) = g h g^{-1}$$

which is best written with left multiplication $L_g : h \mapsto gh$ and right multiplication $R_g : h \mapsto hg$ as

$$C_g = R_{g^{-1}} \circ L_g = L_g \circ R_{g^{-1}}$$

The conjugation by $g \in G$ is an automorphism of G i.e. a (Lie group) homomorphism $C_g : G \longrightarrow G$.

Moreover, the map $C : g \mapsto C_g$ is a homomorphism $C : G \longrightarrow \mathrm{Aut}(G)$ to the group of automorphisms of G.

Since for any $g \in G$, $C_g : G \longrightarrow G$ is a Lie group homomorphism, its derivative at $e \in G$ is a Lie algebra homomorphism

$$Ad_g := dC_g \quad : \quad \mathfrak{g} \longrightarrow \mathfrak{g} .$$

Using $C_g = L_g \circ R_{g^{-1}} = R_{g^{-1}} \circ L_g$, we get

$$Ad_g = (dL_g)|_{g^{-1}} \circ (dR_{g^{-1}})|_e = (dR_{g^{-1}})|_g \circ (dL_g)|_e .$$

Note that $g \longmapsto Ad_g$ defines a representation of G on $\mathfrak{g}$, since $C_{gh} = C_g \circ C_h$ and thus

$$Ad_{gh} = d(C_{gh})|_e = (dC_g)|_e \circ (dC_h)|_e = Ad_g \circ Ad_h .$$

This representation, $Ad : G \longrightarrow Aut(\mathfrak{g}) \cong GL(\mathfrak{g})$ the <u>adjoint representation</u> of G (or <u>adjoint action</u>).

We obtained Ad_g by looking at $h \longmapsto ghg^{-1} = C_g(h)$ for h infinitesimally close to e.

We can furthermore let g be infinitesimally close to e: consider the derivative of $Ad : G \longrightarrow Aut(\mathfrak{g}) \cong GL(\mathfrak{g})$ at e:

$$d(Ad)|_e \quad : \quad \mathfrak{g} \longrightarrow End(\mathfrak{g}) \cong \mathfrak{gl}(\mathfrak{g})$$

where we noticed that the Lie algebra of the Lie group $Aut(\mathfrak{g})$ is $End(\mathfrak{g}) = Hom(\mathfrak{g}, \mathfrak{g})$ (this is the same statement as: "the Lie algebra of $GL_n(\mathbb{R})$ is $\mathfrak{gl}_n(\mathbb{R}) = \mathbb{R}^{n \times n}$ ").

Denote $ad = d(Ad)|_e \quad : \quad \mathfrak{g} \longrightarrow End(\mathfrak{g})$.

Since $Ad : G \rightarrow Aut(\mathfrak{g})$ is a Lie group homom., ad is a Lie algebra homomorphism.

This representation of $\mathfrak{g}$ on $\mathfrak{g}$ is called the adjoin representation of the Lie algebra $\mathfrak{g}$.

Adjoint representations of matrix Lie groups and BCH formula

$G \subset GL_n(\mathbb{R})$ a matrix Lie group

$\mathfrak{g} \subset \mathfrak{gl}_n(\mathbb{R})$ its Lie algebra

conjugation by $g \in G$: $h \longmapsto g h g^{-1} = C_g(h)$

h infinitesimally close to e : $h = e^{\varepsilon Y} \longmapsto g e^{\varepsilon Y} g^{-1} = g(\mathbb{1} + \varepsilon Y + \dots) g^{-1}$

$$= \mathbb{1} + \varepsilon\, g Y g^{-1} + \dots$$

$$\mathrm{Ad}_g : \mathfrak{g} \longrightarrow \mathfrak{g}$$

$$Y \longmapsto g Y g^{-1} = \mathrm{Ad}_g(Y)$$

g infinitesimally close to e : $g = e^{\delta X}$

$$Y \longmapsto e^{\delta X} Y e^{-\delta X} = (\mathbb{1} + \delta X + \dots) Y (\mathbb{1} - \delta X - \dots)$$

$$= Y + \delta \cdot (XY - YX) + \dots$$

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

More generally, from the Baker–Campbell–Hausdorff formula

$$\exp(X)\exp(Y) = \exp(\mu(X,Y)) = \exp(X + Y + \beta(X,Y) + \dots)$$

one can show that in general

$$\mathrm{ad}_X(Y) = 2 \cdot \beta(X,Y).$$

We take either as a definition of the Lie bracket

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g} \qquad [X,Y] := \mathrm{ad}_X(Y) = 2\beta(X,Y).$$

<u>Theorem</u>: The following hold:

(i): $\mathrm{ad} : \mathfrak{g} \longrightarrow \mathrm{End}(\mathfrak{g})$ is a Lie algebra homomorphism

(ii): $\forall X, Y \in \mathfrak{g}$: $\mathrm{ad}_{[X,Y]} = \mathrm{ad}_X \circ \mathrm{ad}_Y - \mathrm{ad}_Y \circ \mathrm{ad}_X$

(iii): $\forall X, Y, Z \in \mathfrak{g}$: $[[X,Y],Z] = [X,[Y,Z]] - [Y,[X,Z]]$

(iv): $\forall X, Y, Z \in \mathfrak{g}$: $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0.$

"Jacobi identity"

Proof: (i) holds since $\mathrm{ad} = d(\mathrm{Ad})|_e$ and $\mathrm{Ad}: G \to \mathrm{Aut}(\mathfrak{g})$ is a Lie group homomorphism.

It is easy to check the equivalences of the conditions (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv). $\square$

The Lie algebra $\mathfrak{g}$ of a (real) Lie group G is thus a (real) vector space equipped with an antisymmetric bilinear operation ("the Lie bracket") which satisfies the Jacobi identity.

We give the following general definition.

Def: A Lie algebra over a field $\mathbb{K}$ ($\mathrm{char}(\mathbb{K}) \neq 2$) is a $\mathbb{K}$-vector space $\mathfrak{g}$ equipped with a $\mathbb{K}$-bilinear map $\mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$ denoted

$$\mathfrak{g} \times \mathfrak{g} \ni (X,Y) \longmapsto [X,Y] \in \mathfrak{g}$$

which satisfies

$$[X,Y] = -[Y,X] \quad \forall X,Y \in \mathfrak{g} \qquad \text{"antisymmetry"}$$

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \quad \forall X,Y,Z \in \mathfrak{g} \qquad \text{"Jacobi identity"}$$

In general we also define homomorphisms and representations of Lie algebras in the same way as before:

Def: A homomorphism of $\mathbb{K}$-Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ is a $\mathbb{K}$-linear map $\lambda: \mathfrak{g} \to \mathfrak{h}$ such that

$$\forall X,Y \in \mathfrak{g}: \quad \lambda([X,Y]_{\mathfrak{g}}) = [\lambda(X), \lambda(Y)]_{\mathfrak{h}}.$$

Recall: The derivative $(d\varphi)|_e : \mathfrak{g} \to \mathfrak{h}$ of a Lie group homomorphism $\varphi: G \to H$ is a homomorphism of $\mathbb{R}$-Lie algebras.

Def: A representation of a $\mathbb{K}$-Lie algebra $\mathfrak{g}$ on a $\mathbb{K}$-vector space V is a Lie algebra homom. $\rho: \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$.

The second principle of infinitesimal transformations

Let us begin by taking an alternative point of view to the first principle — that a Lie group homomorphism $\varphi : G \to H$ is determined by the derivative $\lambda = d\varphi|_e : \mathfrak{g} \to \mathfrak{h}$ at the neutral element, if G is connected.

Let $g \in G$, and let $\gamma : I \to G$ ($I = [a,b]$ interval) be a path from $\gamma(a) = e$ to $\gamma(b) = g$ (such smooth paths exist since G is connected).

Encode the path γ in its "steering":

write $\frac{d}{dt}\gamma(t) = (dL_{\gamma(t)})|_e (X(t))$, where $X(t) \in \mathfrak{g}$.

We get a path $I \to \mathfrak{g}$ $t \mapsto X(t)$ in the Lie algebra $\mathfrak{g}$ which "steers" the path γ.

One can recover the value $\varphi(g)$ of the homomorphism φ at g by looking at the path $\tilde{\gamma} = \varphi \circ \gamma$ in H. Write similarly $\frac{d}{dt}\tilde{\gamma}(t) = (dL_{\tilde{\gamma}(t)})|_e (\tilde{X}(t))$, where $\tilde{X}(t) \in \mathfrak{h}$.

We claim that the "steering" $\tilde{X}$ of $\tilde{\gamma}$ is determined by the "steering" X of γ via $\tilde{X}(t) = \lambda(X(t))$, where $\lambda = (d\varphi)|_e : \mathfrak{g} \to \mathfrak{h}$.

Check: φ homom. $\Rightarrow$ $\varphi \circ L_g = L_{\varphi(g)} \circ \varphi$.

Thus $d\varphi|_g = (dL_{\varphi(g)})|_e \circ d\varphi|_e = dL_{\varphi(g)}|_e \circ \lambda$.

Now
$$\begin{aligned}
\frac{d}{dt}\tilde{\gamma}(t) &= \frac{d}{dt}\varphi(\gamma(t)) \\
&= d\varphi|_{\gamma(t)}\left((dL_{\gamma(t)})|_e (X(t))\right) \\
&= (dL_{\varphi(\gamma(t))})|_e\left((d\varphi)|_e (X(t))\right) \\
&= (dL_{\tilde{\gamma}(t)})|_e\left(\lambda(X(t))\right) \qquad \square
\end{aligned}$$

Thus, given the path γ to g, we get $X(t)$ and thus $Y(t) = \lambda(X(t))$, and we can solve the ODE

$$\frac{d}{dt}\tilde{\gamma}(t) = (dL_{\tilde{\gamma}(t)})|_e \tilde{X}(t), \qquad \tilde{\gamma}(a) = e$$

on H to find the value $\varphi(g) = \varphi(\gamma(b)) = \tilde{\gamma}(b)$.

This recovers:

"First principle": For G connected, a homom. $\varphi: G \to H$ is determined by $\lambda = (d\varphi)|_e : \mathfrak{g} \to \mathfrak{h}$.

To get to the second principle, we will have to consider whether this way of recovering the values of φ gives a well-defined unique answer independent of the path γ from e to g.

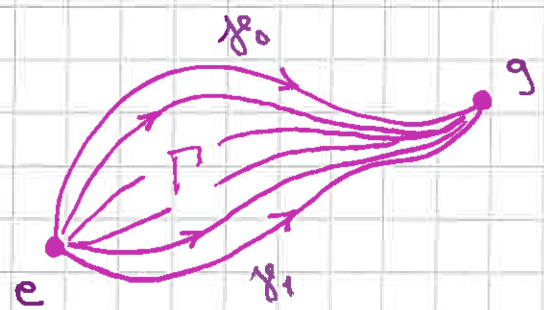
Suppose that we have two paths $\gamma_0, \gamma_1 : I \to G$ such that $\gamma_0(a) = e = \gamma_1(a)$, $\gamma_0(b) = g = \gamma_1(b)$.

Let us assume that one can be smoothly deformed to the other ("homotopy"), i.e. there exists

$$\Gamma : I \times [0,1] \longrightarrow G \quad \text{s.t.}$$

$\forall t \in I$: $\Gamma(t,0) = \gamma_0(t)$ $\qquad$ $\Gamma(t,1) = \gamma_1(t)$

$\forall s \in [0,1]$: $\Gamma(a,s) = e$ $\qquad$ $\Gamma(b,s) = g$



"deformation of γ_0 to γ_1"

(homotopy)

Suppose that we are given a Lie alg. homomorphism $\lambda: \mathfrak{g} \to \mathfrak{h}$.

Let $X(t,s) \in \underline{g}$ be the "steering" for the path $t \mapsto \Gamma(t,s)$, i.e.

$$\frac{\partial}{\partial t}\Gamma(t,s) = (dL_{\Gamma(t,s)})|_e (X(t,s)).$$

Define $\tilde{\Gamma} : I \times [0,1] \longrightarrow H$ by pushing the steering to $\underline{h}$ by $\lambda : \underline{g} \to \underline{h}$, i.e.

$$\tilde{\Gamma}(a,s) = e \in H \qquad \frac{\partial}{\partial t}\tilde{\Gamma}(t,s) = (dL_{\tilde{\Gamma}(t,s)})|_e (\tilde{X}(t,s))$$

where $\tilde{X}(t,s) = \lambda(X(t,s))$.

We will similarly encode the change of the deformation parameter s into $Y(t,s) \in \underline{g}$ by

$$\frac{\partial}{\partial s}\Gamma(t,s) = (dL_{\Gamma(t,s)})|_e (Y(t,s))$$

and let us agree to write also $\tilde{Y}(t,s) \in \underline{h}$ for

$$\frac{\partial}{\partial s}\tilde{\Gamma}(t,s) = (dL_{\tilde{\Gamma}(t,s)})|_e (\tilde{Y}(t,s)).$$

We claim that there is a "zero-curvature" relation between the "steerings" $X, Y : I \times [0,1] \longrightarrow \underline{g}$.

<u>Lemma</u> We have $\displaystyle \frac{\partial}{\partial s}X(t,s) = \frac{\partial}{\partial t}Y(t,s) + [X(t,s), Y(t,s)]$ $(\bigstar)$

<u>Proof</u>: Compute the second derivative $\frac{\partial}{\partial s}\frac{\partial}{\partial t}\Gamma$ by expanding to the second order

$$\begin{aligned}\Gamma(t+\varepsilon, s+\delta) &= \Gamma(t+\varepsilon, s)\exp(\delta \cdot Y(t+\varepsilon, s) + \dots) \\ &= \Gamma(t,s)\exp(\varepsilon X(t,s) + \dots)\exp(\delta \cdot Y(t,s) + \delta\varepsilon \tfrac{\partial}{\partial t}Y(t,s) + \dots) \\ &= \Gamma(t,s)\exp(\varepsilon X(t,s) + \delta \cdot Y(t,s) + \delta\varepsilon \tfrac{\partial}{\partial t}Y(t,s) + \tfrac{\varepsilon\delta}{2}[X(t,s), Y(t,s)] + \dots)\end{aligned}$$

and by varying the variables in the opposite order

$$\begin{aligned}\Gamma(t+\varepsilon, s+\delta) &= \Gamma(t, s+\delta)\exp(\varepsilon \cdot X(t, s+\delta) + \dots) \\ &= \Gamma(t,s)\exp(\delta Y(t,s) + \dots)\exp(\varepsilon X(t,s) + \varepsilon\delta \tfrac{\partial}{\partial s}X(t,s) + \dots) \\ &= \Gamma(t,s)\exp(\delta Y(t,s) + \varepsilon X(t,s) + \varepsilon\delta \tfrac{\partial}{\partial s}X(t,s) + \tfrac{\delta\varepsilon}{2}[Y(t,s), X(t,s)] + \dots)\end{aligned}$$

The results must be the same, i.e.

$$\frac{\partial}{\partial t} Y(t,s) + \frac{1}{2}\left[X(t,s), Y(t,s)\right] = \frac{\partial^2}{\partial t\,\partial s}\Gamma(t,s) = \frac{\partial}{\partial s} X(t,s) + \frac{1}{2}\left[Y(t,s), X(t,s)\right].$$

The asserted "zero-curvature equation" follows by antisymmetry of the bracket. $\square$

For the same reason we must also have (for $\tilde{\Gamma}$)

$$\frac{\partial}{\partial s}\tilde{X}(t,s) = \frac{\partial}{\partial t}\tilde{Y}(t,s) + \left[\tilde{X}(t,s), \tilde{Y}(t,s)\right]. \qquad (\tilde{\star})$$

We moreover have the initial condition

$$\tilde{Y}(a,s) = \frac{\partial}{\partial s}\Gamma(a,s) = \frac{\partial}{\partial s} e = 0$$

so the equation $(\tilde{\star})$ determines $\tilde{Y}(t,s)$ uniquely, given $\tilde{X}(t,s) = \lambda(X(t,s))$.

On the other hand, applying $\lambda : \underline{g} \to \underline{h}$ to $(\star)$ we get

$$\underbrace{\lambda\left(\frac{\partial}{\partial s} X(t,s)\right)}_{\substack{= \frac{\partial}{\partial s}\tilde{X}(t,s) \\ \text{by definition} \\ \text{of } X(t,s)}} = \lambda\left(\frac{\partial}{\partial t} Y(t,s)\right) + \underbrace{\lambda\left(\left[X(t,s), Y(t,s)\right]\right)}_{\substack{= \left[\lambda(X(t,s)), \lambda(Y(t,s))\right] \\ = \left[\tilde{X}(t,s), \lambda(Y(t,s))\right].}}$$

by homomorphism property of $\lambda : \underline{g} \to \underline{h}$

Thus $\lambda(Y(t,s))$ is the solution of $(\tilde{\star})$, i.e. we in fact have $\tilde{Y}(t,s) = \lambda(Y(t,s))$.

Now we still note that

$$Y(b,s) = \frac{\partial}{\partial s}\Gamma(b,s) = \frac{\partial}{\partial s} g = 0$$

and therefore $\tilde{Y}(b,s) = \lambda(Y(b,s)) = 0$, which implies

$$\frac{\partial}{\partial s}\tilde{\Gamma}(b,s) = 0, \qquad \text{i.e.} \qquad \tilde{\Gamma}(b,s) = \text{constant} \in H.$$

This is the uniqueness conclusion we wanted:

Proposition Let $\lambda : \underline{g} \to \underline{h}$ be a Lie algebra homomorphism, and $\gamma_0 : I \to G$ and $\gamma_1 : I \to G$ two homotopic paths ("deformable to each other"). Then the steerings along γ_0 and γ_1, pushed by $\lambda : \underline{g} \to \underline{h}$ to obtain paths $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$ on H result in the same end point $\tilde{\gamma}_0(b) = \tilde{\gamma}_1(b) \in H$.

This gives the second principle:

Theorem Let G be a connected, simply connected Lie group, and H a Lie group, and $\lambda : \underline{g} \to \underline{h}$ is a Lie algebra homomorphism between their Lie algebras. Then there exists a (unique) Lie group homomorphism $\varphi : G \to H$ such that $(d\varphi)|_e = \lambda : \underline{g} \to \underline{h}$.

Proof: Any two paths in G from e to $g \in G$ are homotopic, since G is simply connected. Therefore the push by λ of steering along any path to $g \in G$ gives the same endpoint in H, denoted by $\varphi(g) \in H$. This φ is by construction a homomorphism (steer along concatenated paths) and its derivative at e is λ (steer along infinitesimal paths). Uniqueness of such a $\varphi : G \to H$ follows from the First principle. $\square$

<u>Example</u> Recall that the Lie algebra $\underline{so}_3$ of SO_3 is

$$\underline{so}_3 = \{ X \in \mathbb{R}^{3\times 3} \mid X^T = -X \}.$$

It is 3-dimensional (more generally $\dim(\underline{so}_n) = \frac{1}{2} n(n-1)$) and a basis of it is

$$R_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{bmatrix}, \quad R_y = \begin{bmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad R_z = \begin{bmatrix} 0 & -1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

which are the "infinitesimal rotations around x, y, z axes".
Let us compute the brackets of $\underline{so}_3$ in this basis.
By antisymmetry, $[R_x, R_x] = 0 = [R_y, R_y] = [R_z, R_z]$
and $[R_x, R_y] = -[R_y, R_x]$, $[R_y, R_z] = -[R_z, R_y]$, $[R_z, R_x] = -[R_x, R_z]$.
It thus remains to compute three nontrivial brackets.
For example

$$[R_x, R_y] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & +1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = R_z.$$

Similarly, one calculates $[R_y, R_z] = R_x$ and $[R_z, R_x] = R_y$.

<u>Fact</u>: $\exp : \underline{so}_3 \longrightarrow SO_3$ is surjective.

<u>Idea of proof</u>: Euler's axis theorem states that any rotation $M \in SO_3$ leaves some axis fixed (i.e. $\exists v \neq 0$ in $\mathbb{R}^3$ s.t. $Mv = v$). By a conjugation, we may assume that the axis is the x-axis. Then we have

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SO_2, \quad \text{i.e.}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$ for some θ. This M is the rotation around x-axis by angle θ, and by a calculation, $M = \exp(\theta R_x)$. In the general case $M = A \exp(\theta R_x) A^{-1} = \exp(\theta (A R_x A^{-1}))$, for some $A \in SO_3$.

<u>Pauli spin matrices</u>: The following are frequently useful.

Set $\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, $\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

These Pauli spin matrices are Hermitian, $\sigma_j^\dagger = \sigma_j$, and they form a basis of the three-dimensional real vector space of traceless Hermitian 2×2 matrices.

Moreover, the products can be calculated, and give

$$\sigma_j^2 = \mathbb{1} \qquad \sigma_j \sigma_{j+1} = i\,\sigma_{j+2} \qquad \sigma_{j+1}\sigma_j = -i\,\sigma_{j+2}$$

(with indices interpreted modulo 3).

Together with the unit matrix $\mathbb{1}$, the Pauli spin matrices thus form a basis of the algebra of Hermitian matrices (note: the product of Hermitian matrices is Hermitian).

<u>Example</u>: Recall that the Lie algebra $\underline{su}_2$ of SU_2 is

$$\underline{su}_2 = \left\{ X \in \mathbb{C}^{2\times 2} \;\middle|\; X^\dagger = -X \,;\; tr(X) = 0 \right\}$$

It is a three-dimensional real Lie algebra (recall that more generally, $\dim(\underline{su}_n) = n^2 - 1$) consisting of all traceless anti-Hermitian 2×2 matrices.

A convenient basis for $\underline{su}_2$ is

$$S_x = -\frac{i}{2}\sigma_1 \,, \quad S_y = -\frac{i}{2}\sigma_2 \,, \quad S_z = -\frac{i}{2}\sigma_3 \,.$$

From the products of Pauli matrices we easily get

$$[S_x, S_y] = S_z \,, \quad [S_y, S_z] = S_x \,, \quad [S_z, S_x] = S_y$$

and all other brackets follow by antilinearity.

We see that $\underline{su}_2 \cong \underline{so}_3$, with an explicit Lie algebra isomorphism $\underline{su}_2 \to \underline{so}_3$ defined by

$$S_x \mapsto R_x \,, \quad S_y \mapsto R_y \,, \quad S_z \mapsto R_z \,.$$

Adjoint action of SU_2

$$SU_2 = \{ M \in \mathbb{C}^{2\times 2} \mid M^\dagger M = \mathbb{1} , \quad \det(M) = 1 \}$$
$$\mathfrak{su}_2 = \{ X \in \mathbb{C}^{2\times 2} \mid X^\dagger = -X , \quad \mathrm{tr}(X) = 0 \}$$

Adjoint action of SU_2 on $\mathfrak{su}_2$:

$M \in SU_2$, $X \in \mathfrak{su}_2$: $\quad \mathrm{Ad}_M(X) = M X M^{-1} = M X M^\dagger$.

Define a bilinear form $b : \mathfrak{su}_2 \times \mathfrak{su}_2 \longrightarrow \mathbb{R}$ by

$$b(X,Y) = -2 \cdot \mathrm{tr}(XY) .$$

Claim: The bilinear form $b : \mathfrak{su}_2 \times \mathfrak{su}_2 \longrightarrow \mathbb{R}$ is

(i) symmetric : $b(X,Y) = b(Y,X) \quad \forall X,Y \in \mathfrak{su}_2$

(ii) positive definite : $b(X,X) > 0 \quad \forall X \neq 0$

(iii) Ad-invariant : $b(\mathrm{Ad}_M(X), \mathrm{Ad}_M(Y)) = b(X,Y)$
$\forall X,Y \in \mathfrak{su}_2 \quad \forall M \in SU_2$.

Pf: (i) follows by cyclicity of trace ; $\mathrm{tr}(XY) = \mathrm{tr}(YX)$.

To see (ii), we calculate using the Pauli spin matrices that

$$b(S_x, S_x) = -2\,\mathrm{tr}\left(\left(-\tfrac{i}{2}\sigma_1\right)^2 \right) = -2 \cdot \tfrac{-1}{4}\, \mathrm{tr}(\sigma_1^2) = +\tfrac{1}{2}\,\mathrm{tr}(\mathbb{1}) = 1$$

and similarly $b(S_y, S_y) = 1$, $b(S_z, S_z) = 1$,

whereas

$$b(S_x, S_y) = -2 \cdot \left(\tfrac{-i}{2}\right)^2 \mathrm{tr}(\sigma_1 \sigma_2) = -2 \cdot \left(\tfrac{-i}{2}\right)^2 i \cdot \underbrace{\mathrm{tr}(\sigma_3)}_{=0} = 0$$

and similarly $b(S_y, S_z) = 0$, $b(S_z, S_x) = 0$.

This shows that the basis S_x, S_y, S_z is orthonormal for b, and in particular b is positive definite.

Ad-invariance also follows from cyclicity of trace :

$$\mathrm{tr}(\mathrm{Ad}_M(X)\,\mathrm{Ad}_M(Y)) = \mathrm{tr}(M X M^{-1} M Y M^{-1}) = \mathrm{tr}(M X Y M^{-1})$$
$$= \mathrm{tr}(X Y M^{-1} M) = \mathrm{tr}(XY) .$$

$\square$

In other words, the 3-dim. space $\underline{su}_2$ equipped with the inner product b can be identified with the Euclidean space $\mathbb{R}^3$, and the adjoint action $Ad_M : \underline{su}_2 \longrightarrow \underline{su}_2$ of any $M \in SU_2$ is an orthogonal transformation of that space.
Note furthermore that the determinant of Ad_M is $+1$ for all $M \in SU_2$: this is certainly true for $M = \mathbb{1}$ because $Ad_{\mathbb{1}} = id_{\underline{su}_2}$, and since $\det$ is continuous and $SU_2 \simeq S^3$ is connected, and $\det(Ad_M) \in \{\pm 1\}$ by orthogonality of Ad_M, we see that $\det(M) = +1$ for all $M \in SU_2$.

Thus we can interpret $Ad : SU_2 \rightarrow Aut(\underline{su}_2)$ as a homomorphism $\pi : SU_2 \longrightarrow SO_3$.

Let us finally calculate $ad : \underline{su}_2 \longrightarrow End(\underline{su}_2)$, i.e. the derivative $(d\pi)|_e : \underline{su}_2 \longrightarrow \underline{so}_3$.
On $\underline{su}_2$ we use the orthonormal basis S_x, S_y, S_z.
Recall that $ad_X(Y) = [X, Y]$ (one definition of Lie bracket).
Then from already calculated brackets we get

$$ad_{S_x} : \quad S_x \longmapsto 0, \quad S_y \longmapsto S_z, \quad S_z \longmapsto -S_y$$
$$ad_{S_y} : \quad S_x \longmapsto -S_z, \quad S_y \longmapsto 0, \quad S_z \longmapsto S_x$$
$$ad_{S_z} : \quad S_x \longmapsto S_y, \quad S_y \longmapsto -S_x, \quad S_z \longmapsto 0.$$

The matrices of these maps in the basis S_x, S_y, S_z give for $(d\varphi)|_e : \underline{su}_2 \longrightarrow \underline{so}_3$

$$(d\pi)|_e(S_x) = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}}_{= R_x}, \quad (d\pi)|_e(S_y) = \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}}_{= R_y}, \quad (d\pi)|_e(S_z) = \underbrace{\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{= R_z}.$$

Recall that SU_2 is homeomorphic (diffeomorphic) to the unit sphere in 4-dimensions, $S^3 \subset \mathbb{R}^4$, via

$$\mathbb{R}^4 \supset S^3 \ni (x_1, x_2, x_3, x_4) \longmapsto \begin{bmatrix} x_1 + i x_2 & -x_3 + i x_4 \\ x_3 + i x_4 & x_1 - i x_2 \end{bmatrix} \in SU_2 .$$

From this we concluded in particular that SU_2 is connected and simply connected (these are considered known topological properties of S^3).

Another view on the tangent space $\underline{su}_2 = T_{\mathbb{1}} SU_2$ comes from identifying it with the tangent space $T_p S^3$ of S^3 at $p = (1,0,0,0)$.

The vectors $\frac{\partial}{\partial x_2}\big|_p$, $\frac{\partial}{\partial x_3}\big|_p$, $\frac{\partial}{\partial x_4}\big|_p$ span $T_p S^3$, while $\frac{\partial}{\partial x_1}\big|_p$ is normal to the surface of the sphere S^3 in $\mathbb{R}^4$. By an easy calculation we find that under the identification $S^3 \longrightarrow SU_2$

$$\frac{\partial}{\partial x_2}\Big|_p \longmapsto \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = i\sigma_3 = -2 \cdot S_z \in \underline{su}_2$$

$$\frac{\partial}{\partial x_3}\Big|_p \longmapsto \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = -i\sigma_3 = 2 \cdot S_y \in \underline{su}_2$$

$$\frac{\partial}{\partial x_4}\Big|_p \longmapsto \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = i\sigma_1 = -2 \cdot S_x \in \underline{su}_2 .$$

The adjoint action of $M = \begin{bmatrix} x_1 + i x_2 & -x_3 + i x_4 \\ x_3 + i x_4 & x_1 - i x_2 \end{bmatrix} \in SU_2$ on $\underline{su}_2$ can be calculated explicitly in the basis S_x, S_y, S_z, with the result

$$Ad_M : \begin{aligned} S_x &\mapsto (x_1^2 - x_2^2 - x_3^2 + x_4^2) S_x - 2(x_1x_2 + x_3x_4) S_y - 2(x_1x_3 - x_2x_4) S_z \\ S_y &\mapsto 2(x_1x_2 - x_3x_4) S_x + (x_1^2 - x_2^2 + x_3^2 - x_4^2) S_y - 2(x_2x_3 + x_1x_4) S_z \\ S_z &\mapsto 2(x_1x_3 - x_2x_4) S_x + 2(x_1x_4 - x_2x_3) + (x_1^2 + x_2^2 - x_3^2 - x_4^2) S_z \end{aligned}$$

Theorem There exists a Lie group homomorphism $\pi : SU_2 \longrightarrow SO_3$, which is surjective and $\mathrm{Ker}(\pi) = \{\pm \mathbb{1}\} \cong \mathbb{Z}/2\mathbb{Z}$. In particular

$$SU_2 / \{\pm 1\} \cong SO_3 .$$

Proof: $\mathrm{Ad} : SU_2 \longrightarrow \mathrm{Aut}(\underline{su}_2)$ can be interpreted as $\pi : SU_2 \longrightarrow SO_3$ as seen above. The derivative $(d\pi)_{1e} : \underline{su}_2 \longrightarrow \underline{so}_3$ is a surjective linear map. Therefore the image of π contains some neighborhood of $\mathbb{1} \in SO_3$, and since $\mathrm{Im}(\pi)$ is a subgroup of the connected Lie group SO_3, we must have $\mathrm{Im}(\pi) = SO_3$.

(Alternatively, using the fact that $\exp_{SO_3} : \underline{so}_3 \longrightarrow SO_3$ is surjective, and $\varphi(\exp_{SU_2}(X)) = \exp_{SO_3}((d\varphi)_{1e}(X))$, we see that surjectivity of $(d\varphi)_{1e} : \underline{su}_2 \to \underline{so}_3$ implies the surjectivity of $\varphi : SU_2 \to SO_3$.)

We use the expression for $\mathrm{Ad}_M : \underline{su}_2 \longrightarrow \underline{su}_2$ with $M = \begin{bmatrix} x_1 + i x_2 & -x_3 + i x_4 \\ x_3 + i x_4 & x_1 - i x_2 \end{bmatrix}$ to find the kernel: for all diagonal entries to be 1, we must have $x_1^2 = 1$ and $x_2^2 = x_3^2 = x_4^2 = 0$, and this is also sufficient for $\mathrm{ad}_M = \mathrm{id}_{\underline{su}_2}$. The two options $x_1 = \pm 1$ correspond to $\pm \mathbb{1} \in \mathrm{Ker}(\varphi)$. $\square$

Remark SO_3 is not simply connected.

Indeed, take a path $\tilde{\gamma} : [0,1] \to SU_2$ in SU_2 from $+\mathbb{1}_{2\times 2}$ to $-\mathbb{1}_{2\times 2}$, and map it by $\pi : SU_2 \to SO_3$ to a path $\gamma = \pi \circ \tilde{\gamma}$ on SO_3.

This path γ is a loop on SO_3,

$\gamma(0) = \pi(\tilde{\gamma}(0)) = \pi(+\mathbb{1}_{2\times 2}) = \mathbb{1}_{3\times 3}$ and

$\gamma(1) = \pi(\tilde{\gamma}(1)) = \pi(-\mathbb{1}_{2\times 2}) = \mathbb{1}_{3\times 3}$, but this loop can not be deformed to the trivial loop (a lifting of such a deformation to SU_2 would move $-\mathbb{1}_{2\times 2}$ to $+\mathbb{1}_{2\times 2}$ continuously, while staying in $\mathrm{Ker}(\pi) = \{\pm \mathbb{1}_{2\times 2}\} \subset SU_2$ — which is impossible).

So how can we study the representations of SO_3 with the help of our two principles:

- "1st principle": If G is a connected Lie group then a homom. $\varphi : G \to H$ is determined by $\lambda = (d\varphi)|_e : \mathfrak{g} \to \mathfrak{h}$
- "2nd principle" If G is conn. and simply-conn. Lie group then for any Lie algebra homom. $\lambda : \mathfrak{g} \to \mathfrak{h}$ there exists a homom. $\varphi : G \to H$ s.t. $(d\varphi)|_e = \lambda$.

Look at SU_2, which is connected and simply connected, and which covers SO_3 by $\pi : SU_2 \to SO_3$. so that $SO_3 \cong SU_2 / \mathrm{Ker}(\pi)$.

Indeed, given $\lambda : \underline{so}_3 \to \mathfrak{h}$ a Lie alg. homom., we use $\underline{so}_3 \cong \underline{su}_2$ and the 2nd principle to get a homom. $\varphi : SU_2 \to H$. Then whether we can factor

$$\begin{tikzcd} SU_2 \arrow[r, "\varphi"] \arrow[d, "\pi"'] & H \\ SO_3 \arrow[ur, dashed, "\bar{\varphi}"'] & \end{tikzcd}$$

$\varphi \overset{?}{=} \bar{\varphi} \circ \pi$ is determined by whether $\mathrm{Ker}(\varphi) \subset \mathrm{Ker}(\pi)$.

This characterizes homomorphisms from SO_3 and in particular representations of SO_3

The idea to use $\pi : SU_2 \to SO_3$ to study representations of the connected but not simply connected Lie group SO_3 is a particular case of the following general idea.

For a connected Lie group G, the universal covering manifold $\tilde{G}$ of G is also a Lie group. By construction $\tilde{G}$ is connected and simply connected, and the covering map $\pi : \tilde{G} \to G$ is a homomorphism, whose kernel $Ker(\pi) \subset \tilde{G}$ is a discrete subgroup of $\tilde{G}$, which lies in the center of $\tilde{G}$ (elements of $Ker(\pi)$ commute with all elements of $\tilde{G}$). We get an isomorphism

$$G \cong \tilde{G} / Ker(\pi),$$

and the Lie algebras $\mathfrak{g}$ and $\tilde{\mathfrak{g}}$ of G and $\tilde{G}$ are isomorphic, $\mathfrak{g} \cong \tilde{\mathfrak{g}}$.

We use 1^{st} and 2^{nd} principles to see that Lie algebra homomorphisms $\lambda : \mathfrak{g} \to \mathfrak{h}$ are in 1-1 correspondence with Lie group homomorphisms $\varphi : \tilde{G} \to H$.

All Lie group homomorphisms $G \to H$ are obtained by considering only those $\varphi : \tilde{G} \to H$ which are trivial on $Ker(\pi) \subset \tilde{G}$, i.e. $\varphi(k) = e \quad \forall k \in Ker(\pi) \subset \tilde{G}$: in this case we can factor

$$\begin{tikzcd} \tilde{G} \arrow[r, "\varphi"] \arrow[dr, "\pi"'] & H \\ & \tilde{G} \arrow[u, dotted, "\bar{\varphi}"'] \end{tikzcd}$$

4. Representations of $\mathfrak{sl}_2(\mathbb{C})$

We start by analyzing an easy but fundamental case, namely the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. It is a three-dimensional complex Lie algebra.

The importance of focusing on this particular case stems for example from the following:

- The complex Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ is isomorphic to the complexification of the real Lie algebras $\mathfrak{so}_3$ and $\mathfrak{su}_2$, i.e., $\mathfrak{so}_3 \otimes \mathbb{C} = \mathfrak{sl}_2(\mathbb{C})$ and $\mathfrak{su}_2 \otimes \mathbb{C} = \mathfrak{sl}_2(\mathbb{C})$. As such, the complex representations of $\mathfrak{so}_3$ and $\mathfrak{su}_2$ are exactly the same as those of $\mathfrak{sl}_2(\mathbb{C})$. In particular, by understanding the representations of $\mathfrak{sl}_2(\mathbb{C})$, we will ultimately understand the representations of the very important Lie groups SO_3 and SU_2, whose Lie algebras are $\mathfrak{so}_3$ and $\mathfrak{su}_2$.
- The complex Lie algebra $\mathfrak{sl}_2(\mathbb{C})$, viewed as a six-dimensional real Lie algebra, is isomorphic to the Lie algebra of the Lorentz group, i.e. the group of linear transformations of the Minkowski space-time.
- The analysis of all semisimple Lie algebras $\mathfrak{g}$ and their representations will be achieved by finding subalgebras in $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$, and applying our knowledge of the representation theory of $\mathfrak{sl}_2(\mathbb{C})$. Despite the importance of $\mathfrak{sl}_2(\mathbb{C})$ for its own sake (witnessed, e.g., by the previous examples), this is really the fundamental reason for studying it!

4.1. The Lie algebra $\mathfrak{sl}_2(\mathbb{C})$

Recall that $\mathfrak{sl}_2(\mathbb{C})$ is the set

$$\mathfrak{sl}_2(\mathbb{C}) = \left\{ M \in \mathbb{C}^{2\times 2} \;\middle|\; \mathrm{tr}(M) = 0 \right\}$$

of traceless (complex) two-by-two matrices, equipped with the Lie bracket $[M_1, M_2] = M_1 M_2 - M_2 M_1$. As a (complex) vector space, it is three dimensional (cf. Exercise [???]), and we will use the basis

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{II.7}$$

for it. The brackets of these basis elements are

$$[H, E] = 2E, \qquad [H, F] = -2F, \qquad [E, F] = H. \tag{II.8}$$

The chosen basis elements are quite simple matrices, but more importantly this basis choice is a fundamental instance of a canonical basis that can be chosen for any semisimple Lie algebra. This should become clear gradually, and at least by the time we treat the general structure of semisimple Lie algebras.

We can immediately give two examples of representations of $\mathfrak{sl}_2(\mathbb{C})$.

Example II.20. The space $V = \mathbb{C}^2$ is naturally a representation of $\mathfrak{sl}_2(\mathbb{C})$: any element $X \in \mathfrak{sl}_2(\mathbb{C})$ is a 2×2-matrix, which we let act on any vector $v \in V = \mathbb{C}^2$ by matrix multiplication Xv. This two-dimensional representation is called the standard representation of $\mathfrak{sl}_2(\mathbb{C})$.

Example II.21. The adjoint representation of $\mathfrak{sl}_2(\mathbb{C})$ is the vector space $V = \mathfrak{sl}_2(\mathbb{C})$ equipped with the adjoint action: for $X \in \mathfrak{sl}_2(\mathbb{C})$ and $Y \in V = \mathfrak{sl}_2(\mathbb{C})$, we set

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

This is a three-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$.

Concretely, in the basis E, H, F of $\mathfrak{sl}_2(\mathbb{C})$, the adjoint representation $\rho\colon \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}\,(\mathfrak{sl}_2(\mathbb{C}))$ becomes, in view of (II.8),

$$\rho(E) = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \rho(H) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad \rho(E) = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}.$$

4.2. The irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$

Let V be a finite dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$. We will use:

Fact II.10. The action of H on V is diagonalizable.

This fact follows from the preservation of Jordan form (see [**FH91**]), but it is also not particularly difficult to verify directly either.

By Fact II.18, we have an eigenspace decomposition

$$V = \bigoplus_{\mu} V_\mu, \tag{II.9}$$

where μ runs over the eigenvalues of H on V, a priori some finite collection of complex numbers, and V_μ are the corresponding eigenspaces for H

$$V_\mu = \left\{ v \in V \;\middle|\; Hv = \mu v \right\},$$

The decomposition (II.9) completely describes the action of H on V, and the remaining task is to describe the action of E and F — in particular, to see what E and F do to the H-eigenspaces V_μ. Suppose that $v \in V_\mu$. Consider the vector $Ev \in V$. We can figure out the action of H on it by an easy but important calculation which uses the commutator of H and E given by the bracket (II.8).

Fundamental calculation (first time):

$$\begin{aligned} H(Ev) &= E(Hv) + [H, E]v \\ &= E(\mu v) + 2Ev \\ &= (\mu + 2)\, Ev. \end{aligned}$$

This calculation shows that if v is an eigenvector of H with eigenvalue μ, then Ev is an eigenvector of H with eigenvalue $\mu + 2$ (although not necessarily a non-zero vector). In other words, for any μ we have

$$E\colon V_\mu \to V_{\mu+2}.$$

By an entirely similar calculation we see that $F\colon V_\mu \to V_{\mu-2}$.

If we assume that V is an irreducible representation, then it follows that the eigenvalues μ of H differ from each other by integer multiples of two. Indeed, if $\mu' \in \mathbb{C}$ is one eigenvalue of H, then the subspace

$$\bigoplus_{n \in \mathbb{Z}} V_{\mu' + 2n}$$

is invariant not only for H but also for E and F, and therefore actually invariant for the entire $\mathfrak{sl}_2(\mathbb{C})$. Thus the subspace is a subrepresentation, and by irreducibility it must be the entire V. In fact we can conclude a little more. For irreducible V

the H-eigenvalues μ must form an uninterrupted string of complex numbers, of the form

$$\zeta, \zeta+2, \zeta+4, \ldots, \zeta+2(k-1), \zeta+2k,$$

since otherwise the direct sum of only a subset of eigenspaces would be invariant for H, E, and F, and would thus be a proper subrepresentation of V.

So, assume from now on that V is a finite dimensional irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$. Denote by $\lambda = \zeta + 2k$ the last number in the above string of H-eigenvalues — a priori we have $\lambda \in \mathbb{C}$, but we will soon see that λ must be a non-negative integer. Choose a non-zero vector $v_0 \in V_\lambda$. Note that $V_{\lambda+2} = \{0\}$, so necessarily we have $Ev_0 = 0$. We will need to understand the action of F on v_0, and concerning that, we have the following:

Claim II.11. Denote $v_m = F^m v_0$, for $m \in \mathbb{Z}_{\geq 0}$. Then the vectors $v_0, v_1, v_2, \ldots$ span V.

Proof. Let $W \subset V$ be the subspace spanned by the above vectors, $W = \mathrm{span}\,\{F^m v_0 \,|\, m \in \mathbb{Z}_{\geq 0}\}$. By irreducibility of V, it suffices to show that W is invariant under H, E, and F. By definition W is invariant under F. Since $F^m v_0 \in V_{\lambda - 2m}$, it is also invariant under H. It suffices to check that $EW \subset W$. We calculate

$$\begin{aligned} E(F^m v_0) &= [E,F](F^{m-1}v_0) - F\big(E(F^{m-1}v_0)\big) \\ &= H(F^{m-1}v_0) - F\big(E(F^{m-1}v_0)\big). \end{aligned} \tag{II.10}$$

We know that the first term, $H(F^{m-1}v_0) = \big(\lambda - 2(m-1)\big)\, F^{m-1}v_0$, is in W. If we already knew that $E(F^{m-1}v_0)$ is in W, we could thus conclude that also the second term is in W, and thus that $E(F^m v_0) \in W$. This is proved by induction on m. Equation (II.10) serves as the induction step, and to complete the proof, we note that in the case $m = 0$ we have $E(F^0 v_0) = Ev_0 = 0 \in W$ by an earlier observation. In fact by this induction we can prove not only that $E(F^m v_0) \in W$, but we moreover obtain the explicit formula

$$E(F^m v_0) \;=\; (\lambda - m + 1)\, m\; F^{m-1} v_0. \tag{II.11}$$

□

The calculation above has some interesting consequences.

Observation II.12. All eigenspaces V_μ of H are one-dimensional.

Proof. Indeed, $\mu = \lambda - 2m$ for some $m \in \mathbb{Z}_{\geq 0}$ and $V_{\lambda-2m} = \mathrm{span}\,\{F^m v_0\}$. □

Observation II.13. The representation V is determined by the number λ.

Proof. Indeed, if d is the smallest power of F that annihilates v_0, then we see that the vectors $F^m v_0$ for $m = 0, 1, 2, \ldots, d-1$ form a basis of V. We have described explicitly the action of H, E, and F on each basis vector, and the matrix elements of H, E, and F only involved λ as a parameter. □

Observation II.14. The dimension of V is $\lambda + 1$, and in particular λ is a non-negative integer, $\lambda = \dim(V) - 1 \in \mathbb{Z}_{\geq 0}$.

Proof. Let again d be the smallest power of F that annihilates v_0. Note that $d = \dim(V)$. The calculation (II.11) is perfectly valid also for $m = d$, so we get

$$0 \;=\; E(F^d v_0) \;=\; (\lambda - d + 1)\, d\; F^{d-1} v_0.$$

But since $F^{d-1}v_0 \neq 0$, the prefactor on the right-hand-side must vanish, $(\lambda - d + 1)\, d = 0$. Also $d > 0$, so we must have $\lambda - d + 1 = 0$, that is $d = \lambda + 1$. □

The final obsevation below follows directly from the earlier ones.

Observation II.15. The eigenvalues of H on V are

$$\lambda, \lambda - 2, \lambda - 4, \ldots, -\lambda + 4, -\lambda + 2, -\lambda$$

and the multiplicity of each eigenvalue is one. In particular, the H-eigenvalues are all integers, they all have the same parity, and they are symmetric about the origin (i.e. if μ is an eigenvalue, then so is $-\mu$).

We conclude by the following complete description of all irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$.

Theorem II.22. *For each $\lambda \in \mathbb{Z}_{\geq 0}$ there exists an irreducible $\lambda + 1$-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$ with basis $v_0, v_1, \ldots, v_\lambda$ and the actions of H, E, and F on this basis given by*

$$F v_m = \begin{cases} v_{m+1} & \text{for } 0 \leq m < \lambda \\ 0 & \text{for } m = \lambda \end{cases}$$

$$E v_m = \begin{cases} 0 & \text{for } m = 0 \\ (\lambda - m + 1) m\, v_{m-1} & \text{for } 0 < m \leq \lambda \end{cases}$$

$$H v_m = (\lambda - 2m)\, v_m \qquad \text{for all } m.$$

Denote this representations by $L(\lambda)$, Any irreducible finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$ is isomorphic to $L(\lambda)$, for some $\lambda \in \mathbb{Z}_{\geq 0}$.

Proof. We have almost proven this already: $L(\lambda)$ is the representation we have analyzed in this section. We have shown that any finite dimensional irreducible representation of dimension $d \in \mathbb{Z}_{>0}$ must be $L(\lambda)$ for $\lambda = d - 1$. However, we have not yet strictly speaking shown that such a representation ideed exists. To show the existence, it remains to check that the formulas given above for the linear operators H, E, and F on the vector space $L(\lambda)$ with basis $v_0, v_1, \ldots, v_\lambda$ actually do define a representation of $\mathfrak{sl}_2(\mathbb{C})$. The only thing to check is that for any $Z, W \in \mathfrak{sl}_2(\mathbb{C})$ the action of the bracket $[Z, W]$ on $L(\lambda)$ equals the commutator of the actions of Z and W. By looking at the calculations done in this section again, you will notice that we have in fact done everything that is needed in such a check. □

Let us make some final observations which are useful in analyzing representations of $\mathfrak{sl}_2(\mathbb{C})$ that we might encounter. We will use the following fact.

Fact II.16. Any (finite dimensional) representation of $\mathfrak{sl}_2$ is a direct sum of irreducible representations.

This is the general property of complete reducibility for semisimple Lie algebras, see [???]. It could also be verified more directly in the present case, see Exercise [???].

Observation II.17. We have:

- Any representation of $\mathfrak{sl}_2(\mathbb{C})$, in which the H-eigenvalues have the same parity and occur with multiplicity one, is necessarily irreducible.
- The number of irreducible subrepresentations of a (finite dimensional) representation of $\mathfrak{sl}_2(\mathbb{C})$ is the sum of multiplicities of 0 and 1 as H-eigenvalues.

4.3. Examples of representations of $\mathfrak{sl}_2(\mathbb{C})$

4.3.1. *The standard representation*

In Example II.20, we noted that the space $V = \mathbb{C}^2$ is a representation of $\mathfrak{sl}_2(\mathbb{C})$, when the elements of $\mathfrak{sl}_2(\mathbb{C})$ are understood as 2×2-matrices such as in (II.7), and the action of such a matrix on a vector in $\mathbb{C}^2$ is by the usual matrix-vector multiplication. This representation is called the standard representation of $\mathfrak{sl}_2(\mathbb{C})$. If $x = [1\ 0]^\top$ and $y = [0\ 1]^\top$ are the standard basis, then we have $Hx = x$ and $Hy = -y$, so that the H-eigenvalues are $+1$ and -1, and the corresponding eigenspaces are $\mathbb{C}x$ and $\mathbb{C}y$. From Observation II.17 it follows that the standard representation V is irreducible, so by dimensionality in fact $V \cong L(1)$.

4.3.2. *The tensor square of the standard representation*

As above, denote by $V = \mathbb{C}^2 = L(1)$ the standard representation. Consider the representation $V \otimes V$. The H-eigenvalues on $V \otimes V$ are $+2$ with multiplicity one (eigenvector $x \otimes x$), 0 with multiplicity two (eigenvectors $x \otimes y$ and $y \otimes x$), and -2 with multiplicity one (eigenvector $y \otimes y$). Note that because of the multiplicities, Observation II.17 shows that $V \otimes V$ is not irreducible, but instead decomposes into a direct sum of two irreducible subrepresentations.

Note that $V \otimes V = \mathrm{Sym}^2 V \oplus \bigwedge^2 V$ as a vector space, and also as a representation of $\mathfrak{sl}_2(\mathbb{C})$. The two irreducible subrepresentations of the tensor square $V \otimes V$ of the standard representation are the symmetric square[7] $\mathrm{Sym}^2 V \cong L(2)$, and the alternating square[8] $\bigwedge^2 V \cong L(0)$. Here, $\bigwedge^2 V \cong L(0)$ in fact coincides with the trivial representation.

4.3.3. *The adjoint representation*

In Example II.21, we noted that The vector space $\mathfrak{sl}_2(\mathbb{C})$ is a representation of the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ by the adjoint action. Note that $\mathrm{ad}_H(E) = 2\,E$, $\mathrm{ad}_H(H) = 0$, and $\mathrm{ad}_H(F) = -2\,F$, so that the H-eigenvalues are $+2$, 0, and -2, each with multiplicity one. The corresponding H-eigenspaces are $\mathbb{C}E$, $\mathbb{C}H$, and $\mathbb{C}F$. From Observation II.17 it follows that the adjoint representation is irreducible, in fact isomorphic to $L(2)$, by dimensionality again.

[7]Note that $\dim(\mathrm{Sym}^2 V) = 3$, basis x^2, xy, y^2.

[8]Note that $\dim(\bigwedge^2 V) = 1$, basis $x \wedge y$.

5. Lifting representations from Lie algebra to Lie group

We now illustrate how, in practice, the understanding of representations of a complex Lie algebra (such as $\mathfrak{sl}_2(\mathbb{C})$ in the previous section) allows us to study continuous symmetries that are described by a real Lie group (such as SU_2 or SO_3).

First of all, we want to note that as long as one is interested in complex representations, we are allowed to replace a real Lie algebra by its complexification.

Lemma II.23. *Let $\mathfrak{g}$ be a real Lie algebra, and $\mathfrak{g}_\mathbb{C} = \mathfrak{g} \otimes_\mathbb{R} \mathbb{C} = \mathfrak{g} \oplus \mathrm{i}\,\mathfrak{g}$ its complexification. Then any complex representation of $\mathfrak{g}$ has a unique structure of representation of $\mathfrak{g}_\mathbb{C}$ (which restricts back to $\mathfrak{g}$ to the original one), and $\mathrm{Hom}_\mathfrak{g}(V, W) = \mathrm{Hom}_{\mathfrak{g}_\mathbb{C}}(V, W)$. In other words, the categories of complex representations of $\mathfrak{g}$ and $\mathfrak{g}_\mathbb{C}$ are equivalent.*

Proof. Let $\rho\colon \mathfrak{g} \to \mathrm{End}\,(V)$ be a representation of $\mathfrak{g}$ on a complex vector space V. The only $\mathbb{C}$-linear way to extend it to $\mathfrak{g}_\mathbb{C}$ is to define $\rho_\mathbb{C}\colon \mathfrak{g}_\mathbb{C} \to \mathrm{End}\,(V)$ by setting $\rho_\mathbb{C}(X + \mathrm{i}\,Y) = \rho(X) + \mathrm{i}\,\rho(Y)$. We leave it to the reader to check that this extension maps brackets in $\mathfrak{g}_\mathbb{C}$ to commutators in $\mathrm{End}\,(V)$, and thus defines a representation of $\mathfrak{g}_\mathbb{C}$. Note that the converse direction is clear — any representation of $\mathfrak{g}_\mathbb{C}$ restricts to a representation of $\mathfrak{g} \subset \mathfrak{g}_\mathbb{C}$.

As for morphisms of representations, if $f_\mathbb{C}\colon V \to W$ is a morphism of $\mathfrak{g}_\mathbb{C}$-representations, then a fortiori it is a morphism of $\mathfrak{g}$-representations. We only need to show the other direction, that if $f\colon V \to W$ is a morphism of $\mathfrak{g}$-representations, then it is also a morphism of $\mathfrak{g}_\mathbb{C}$-representations. But this is clear by $\mathbb{C}$-linearity of f and the way the representations $\rho^V_\mathbb{C}$ and $\rho^W_\mathbb{C}$ extend ρ^V and ρ^W. □

Example II.24. Recall that the three-dimensional real Lie algebras $\mathfrak{su}_2$ and $\mathfrak{so}_3$ are isomorphic. We next observe that the complexification of either one is the three-dimensional complex Lie algebra $\mathfrak{sl}_2(\mathbb{C})$.

Consider for example $\mathfrak{so}_3$ with basis R^x, R^y, R^z such that $[R^x, R^y] = R^z$, $[R^y, R^z] = R^x$, and $[R^z, R^x] = R^y$, see Example II.17. The complexification $\mathfrak{so}_3(\mathbb{C}) = \mathfrak{so}_3 \otimes_\mathbb{R} \mathbb{C}$ has a corresponding basis (now over $\mathbb{C}$), which we for clarity denote here by $R^x_c = R^x \otimes 1, R^y_c = R^y \otimes 1, R^z_c = R^z \otimes 1$. The Lie brackets of these basis elements in $\mathfrak{so}_3(\mathbb{C})$ are just

$$[R^x_c, R^y_c]_{\mathfrak{so}_3(\mathbb{C})} = R^z_c, \qquad [R^y_c, R^z_c]_{\mathfrak{so}_3(\mathbb{C})} = R^x_c, \qquad [R^z_c, R^x_c]_{\mathfrak{so}_3(\mathbb{C})} = R^y_c.$$

We now change to another basis. Denote $R^0 = -2\mathrm{i}\,R^z_c$ and $R^+ = R^x_c + \mathrm{i}\,R^y_c$ and $R^- = R^x_c - \mathrm{i}\,R^y_c$ — clearly R^0, R^+, R^- also forms a basis of $\mathfrak{so}_3(\mathbb{C})$. The brackets of these new basis elements are easily calculated using the $\mathbb{C}$-bilinearity of $[\cdot,\cdot]_{\mathfrak{so}_3(\mathbb{C})}$ and the brackets of R^x_c, R^y_c, R^z_c — we get

$$[R^0, R^+]_{\mathfrak{so}_3(\mathbb{C})} = 2R^+, \qquad [R^0, R^-]_{\mathfrak{so}_3(\mathbb{C})} = 2R^-, \qquad [R^+, R^-]_{\mathfrak{so}_3(\mathbb{C})} = R^+.$$

Comparing with the brackets of H, E, F in $\mathfrak{sl}_2(\mathbb{C})$ given in Equation (II.8), we immediately see that the map $\mathfrak{so}_3(\mathbb{C}) \to \mathfrak{sl}_2(\mathbb{C})$ defined by linear extension of $R^0 \mapsto H$, $R^+ \mapsto E$, $R^- \mapsto F$ is a Lie algebra isomorphism, $\mathfrak{so}_3(\mathbb{C}) \cong \mathfrak{sl}_2(\mathbb{C})$. Similarly we have $\mathfrak{su}_2(\mathbb{C}) \cong \mathfrak{sl}_2(\mathbb{C})$.

In particular we have equivalences

$$\{\text{complex rep'ns of } \mathfrak{su}_2\} \leftrightarrow \{\text{complex rep'ns of } \mathfrak{sl}_2(\mathbb{C})\} \leftrightarrow \{\text{complex rep'ns of } \mathfrak{so}_3\}.$$

Recall that we found that the finite dimensional irreducible representations of the three-dimensional complex Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ are $L(\lambda)$, with $\lambda \in \mathbb{Z}_{\geq 0}$. By Lemma II.23, then, these are also the finite dimensional irreducible complex representations of the real Lie algebras $\mathfrak{su}_2$ and $\mathfrak{so}_3$.

The fact that allows to get from representations of Lie algebras to representations of Lie groups is the following consequence of our two principles for Lie groups.

- Let G be a Lie group and $\mathfrak{g}$ its Lie algebra.
  (i) Every representation $\varrho\colon \mathrm{G} \to \mathrm{Aut}(V)$ of the Lie group G defines a representation $\rho = (\mathrm{d}\varrho)|_e \colon \mathfrak{g} \to \mathrm{End}(V)$ of the Lie algebra $\mathfrak{g}$, and any intertwining map of representations of G is an intertwining map of representations of $\mathfrak{g}$.
  (ii) If G is simply connected, then $\varrho \mapsto \rho = (\mathrm{d}\varrho)|_e$ gives an equivalence of categories of representations of G and representations of $\mathfrak{g}$. In particular, every representation of the Lie algebra $\mathfrak{g}$ is the derivative at e of some representation of the Lie group G.

Example II.25. Recall that SU_2 is simply connected by Theorem II.12. As a special case of the theorem above we get the equivalence

$$\{\text{representations of } \mathrm{SU}_2\} \;\leftrightarrow\; \{\text{representations of } \mathfrak{su}_2\}\,.$$

In particular, the irreducible complex representations of SU_2 are $L(\lambda)$, $\lambda \in \mathbb{Z}$.

The easiest way to give the explicit SU_2 action on $L(\lambda)$ is perhaps to realize that $L(\lambda) = \mathrm{Sym}^\lambda \mathbb{C}^2$ is a symmetric tensor product of the standard representation $\mathbb{C}^2$. The action of SU_2 on the standard representation $\mathbb{C}^2$ is the obvious matrix-vector multiplication, and the action on the symmetric tensor power can be read off from here. The example of the three-dimensional irreducible $L(2)$, for example, in the basis x^2, xy, y^2, gives that

$$\begin{bmatrix} \xi_1 + \mathrm{i}\xi_2 & -\xi_3 + \mathrm{i}\xi_4 \\ \xi_3 + \mathrm{i}\xi_4 & \xi_1 - \mathrm{i}\xi_2 \end{bmatrix} \in \mathrm{SU}_2$$

is represented by the matrix

$$\begin{bmatrix} \xi_1^2 + 2\mathrm{i}\xi_2\xi_1 - \xi_2^2 & -\xi_1\xi_3 - \mathrm{i}\xi_2\xi_3 + \mathrm{i}\xi_1\xi_4 - \xi_2\xi_4 & \xi_3^2 - 2\mathrm{i}\xi_4\xi_3 - \xi_4^2 \\ 2\xi_1\xi_3 + 2\mathrm{i}\xi_2\xi_3 + 2\mathrm{i}\xi_1\xi_4 - 2\xi_2\xi_4 & \xi_1^2 + \xi_2^2 - \xi_3^2 - \xi_4^2 & -2\xi_1\xi_3 + 2\mathrm{i}\xi_2\xi_3 + 2\mathrm{i}\xi_1\xi_4 + 2\xi_2\xi_4 \\ \xi_3^2 + 2\mathrm{i}\xi_4\xi_3 - \xi_4^2 & \xi_1\xi_3 - \mathrm{i}\xi_2\xi_3 + \mathrm{i}\xi_1\xi_4 + \xi_2\xi_4 & \xi_1^2 - 2\mathrm{i}\xi_2\xi_1 - \xi_2^2 \end{bmatrix}.$$

Although the statement of the previous fact appears to only concern simply connected Lie groups, it can in fact be used for any connected Lie groups G. We only need to pass through the universal cover $\tilde{\mathrm{G}}$.

Example II.26. The group SO_3 of rotations of the Euclidean space $\mathbb{R}^3$ is connected but not simply connected: by Theorem **??** its universal cover is SU_2, and the kernel of the covering map $\phi\colon \mathrm{SU}_2 \to \mathrm{SO}_3$ is the two element subgroup $\Gamma = \{\pm \mathbb{I}_2\}$ of the center of SU_2. We have $\mathrm{SO}_3 = \mathrm{SU}_2/\Gamma$.

By Example II.25, the irreducible complex representations of SU_2 are the same as the irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$, i.e., $L(\lambda)$ for $\lambda \in \mathbb{Z}_{\geq 0}$. To get the irreducible representations of SO_3, the remaining question is: which ones among $L(\lambda)$ are trivial on Γ?

The solution is easy once we notice that

$$-\mathbb{I}_2 = \exp(2\pi S^z) \in \mathrm{SU}_2, \qquad \text{where } S^z = -\frac{\mathrm{i}}{2}\sigma_3 = \begin{bmatrix} -\frac{\mathrm{i}}{2} & 0 \\ 0 & \frac{\mathrm{i}}{2} \end{bmatrix} \in \mathfrak{su}_2.$$

To lift a representation $\rho\colon \mathfrak{su}_2 \to \mathrm{End}(V)$ to a representation of $\varrho\colon \mathrm{SU}_2 \to \mathrm{Aut}(V)$, we must set $\varrho(\exp(X)) = \exp(\rho(X))$. In particular we have $\varrho(-\mathbb{I}_2) = \exp\big(2\pi\rho(S^z)\big)$. On $L(\lambda)$, the operator $\rho(S^z) = \frac{1}{2}\mathrm{i}\,\rho(H)$ is diagonalizable with eigenvalues $\mathrm{i}\,\frac{\lambda}{2}$, $\mathrm{i}\,(\frac{\lambda}{2}-1)$, $\ldots$, $-\mathrm{i}\,\frac{\lambda}{2}$. If λ is an even integer, then these are integer multiples of i and $\varrho(-\mathbb{I}_2) = \exp\big(2\pi\rho(S^z)\big)$ is the identity operator on the representation, so the representation is trivial on $\Gamma = \{\pm\mathbb{I}_2\}$. If λ is an odd integer, then the eigenvalues of S^z are half-integer multiples of i, and $\varrho(-\mathbb{I}_2) = \exp\big(2\pi\rho(S^z)\big)$ is minus identity, so the representation is non-trivial on $\Gamma = \{\pm\mathbb{I}_2\}$.

We conclude that the irreducible complex representations of SO_3 are $L(\lambda)$ with $\lambda \in 2\,\mathbb{Z}_{\geq 0}$.

6. Representations of $\mathfrak{sl}_3(\mathbb{C})$

We already showed how to find and construct all irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$, and how to apply the results to representations of Lie groups, whose Lie algebras have $\mathfrak{sl}_2(\mathbb{C})$ as their complexification, e.g., SU_2 and SO_3.

We will proceed to treat more complicated (semisimple) Lie algebras. We start in this section by considering $\mathfrak{sl}_3(\mathbb{C})$. The representations of $\mathfrak{sl}_3(\mathbb{C})$ are needed for example in quantum chromodynamics (QCD), the theory of strong interactions that govern the atomic nucleai. Besides their direct relevance, the analysis of the structure and representations of $\mathfrak{sl}_3(\mathbb{C})$ will serve as a wonderful example of what happens with semisimple Lie algebras in full generality.

We will follow a similar strategy as in the case of $\mathfrak{sl}_2(\mathbb{C})$ to analyze the structure of $\mathfrak{sl}_3(\mathbb{C})$ and its representations. We only require some new ideas, or rather reinterpretations of a few concepts and arguments. These ideas turn out to be powerful — with them, we will be able to handle any semisimple Lie algebra.

6.1. The Lie algebra $\mathfrak{sl}_3(\mathbb{C})$

Recall that $\mathfrak{sl}_3(\mathbb{C})$ is the set

$$\mathfrak{sl}_3(\mathbb{C}) = \big\{ M \in \mathbb{C}^{3\times 3} \;\big|\; \mathrm{tr}(M) = 0 \big\}$$

of traceless (complex) three-by-three matrices, equipped with the Lie bracket $[M_1, M_2] = M_1 M_2 - M_2 M_1$. As a (complex) vector space, it is eight dimensional

$$\dim\big(\mathfrak{sl}_3(\mathbb{C})\big) = 8.$$

Indeed, the nine entries $X_{i,j}$, $1 \leq i, j \leq 3$, of a matrix $X \in \mathfrak{sl}_3(\mathbb{C})$ can be chosen arbitrarily subject to just one linear condition, $\mathrm{tr}(X) = X_{1,1} + X_{2,2} + X_{3,3} = 0$.

Remark II.27. For calculations below, we recall the definition and properties of the elementary matrices E^{kl}. For a general dimension $n \in \mathbb{N}$ and for $1 \leq k, l \leq n$, the *elementary matrix* $E^{kl} \in \mathbb{K}^{n\times n}$ is the matrix whose (k,l)-entry is one, and all other entries are zeroes, $E^{kl}_{ij} = \delta_{k,i}\,\delta_{l,j}$. The products of such matrices are

$$E^{kl} E^{k'l'} = \delta_{l,k'}\, E^{kl'},$$

as is verified by the following direct calculation

$$\begin{aligned}\big(E^{kl}E^{k'l'}\big)_{ij} &= \sum_m E^{kl}_{im} E^{k'l'}_{mj} = \sum_m \delta_{k,i}\,\delta_{l,m}\,\delta_{k',m}\,\delta_{l',j} = \delta_{l,k'}\,\delta_{k,i}\,\delta_{l',j} \\ &= \delta_{l,k'}\, E^{kl'}_{ij}.\end{aligned}$$

The n^2 elementary matrices E^{kl} form a basis of $\mathfrak{gl}_n(\mathbb{K})$, and the brakets in $\mathfrak{gl}_n(\mathbb{K})$ (and thus also in any Lie subalgebra $\mathfrak{g} \subset \mathfrak{gl}_n(\mathbb{K})$) read

$$\begin{aligned}[E^{kl}, E^{k'l'}] &= E^{kl}E^{k'l'} - E^{k'l'}E^{kl} \\ &= \delta_{l,k'}\, E^{kl'} - \delta_{l',k}\, E^{k'l}. \end{aligned} \tag{II.12}$$

In our analysis of $\mathfrak{sl}_3(\mathbb{C})$, we will follow steps modelled on those that we took in the analysis of $\mathfrak{sl}_2(\mathbb{C})$ in the previous lecture. For $\mathfrak{sl}_2(\mathbb{C})$, our analysis relied first of all on a good choice of basis H, E, F — we split any representation (including the adjoint representation on $\mathfrak{sl}_2(\mathbb{C})$ itself) to eigenspaces of H, and figured out how E and F acted on the eigenspaces. The task now is to find the appropriate generalizations.

The good idea turns out to be not to pick just one element to diagonalize, but rather to take an entire subspace $\mathfrak{h} \subset \mathfrak{sl}_3(\mathbb{C})$ to be diagonalized simultaneously. Such a simultaneous diagonalization in any representation succeeds if all the needed operators commute with each other, which is guaranteed if $\mathfrak{h}$ is an abelian subalgebra of $\mathfrak{sl}_3(\mathbb{C})$. We choose $\mathfrak{h}$ to consist of all diagonal matrices in $\mathfrak{sl}_3(\mathbb{C})$, i.e.,

$$\mathfrak{h} = \left\{ \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} \;\middle|\; a_1, a_2, a_3 \in \mathbb{C},\ a_1 + a_2 + a_3 = 0 \right\}. \tag{II.13}$$

All diagonal matrices indeed commute with each other, so $[\mathfrak{h}, \mathfrak{h}] = 0$, and the simultaneous diagonalization of the action of all $H \in \mathfrak{h}$ is possible.

Since we are not considering the diagonalization of a single linear operator, but an entire space of operators, the concept of eigenvalue needs to be appropriately generalized. If V is a representation, and $v \in V$ is a simultaneous eigenvector for the action of all $H \in \mathfrak{h}$, then we have

$$Hv = \mu(H)v \qquad \forall H \in \mathfrak{h}, \tag{II.14}$$

where $\mu(H)$ denotes the eigenvalue of the action of $H \in \mathfrak{h}$. Obviously $\mu(H)$ depends linearly on H, and so defines a linear functional $\mu \colon \mathfrak{h} \to \mathbb{C}$, i.e., an element $\mu \in \mathfrak{h}^*$ of the dual of $\mathfrak{h}$. This is the appropriate generalization of eigenvalues and eigenvectors. We call $\mu \in \mathfrak{h}^*$ a *weight* and $v \in V$ satisfying (II.14) a *weight vector* (of weight μ). Analogously to the decomposition (II.9), any finite-dimensional representation V of $\mathfrak{sl}_3(\mathbb{C})$ has a decomposition

$$V = \bigoplus_{\mu} V_\mu, \tag{II.15}$$

where μ runs over weights V, a priori some finite collection of linear functionals $\mu \in \mathfrak{h}^*$, and V_μ are the corresponding *weight spaces* for $\mathfrak{h}$

$$V_\mu = \left\{ v \in V \;\middle|\; \forall H \in \mathfrak{h} : Hv = \mu(H)v \right\}. \tag{II.16}$$

We have $\dim(\mathfrak{h}) = 2$, and to be concrete we can take a basis $H^{1,2} = E^{1,1} - E^{2,2}$, $H^{2,3} = E^{2,2} - E^{3,3}$ for $\mathfrak{h}$. It is convenient to write the dual elements as linear combinations of η^i, $i = 1, 2, 3$, defined on all diagonal 3×3-matrices by

$$\eta^i\Big(\sum_{j=1}^{3} a_j E^{j,j}\Big) = a_i.$$

As a basis of the dual, we can then take for example $\eta^1 - \eta^2$ and $\eta^2 - \eta^3$, but we remark that all η^i, $i = 1, 2, 3$, make sense as elements of $\mathfrak{h}^*$.[9]

Example II.28. The space $V = \mathbb{C}^3$ is naturally a representation of $\mathfrak{sl}_3(\mathbb{C})$: any element $X \in \mathfrak{sl}_3(\mathbb{C})$ is a 3×3-matrix, which we let act on any vector $v \in V = \mathbb{C}^3$ by matrix multiplication Xv. This three-dimensional representation is called the standard representation of $\mathfrak{sl}_3(\mathbb{C})$.

The standard basis vectors $e_1, e_2, e_3 \in \mathbb{C}^3$ are weight vectors, with respective weights η^1, η^2, η^3. The weight space decomposition of the standard representation $\mathbb{C}^3$ of $\mathfrak{sl}_3(\mathbb{C})$ is thus

$$\mathbb{C}^3 = \mathbb{C}e_1 \oplus \mathbb{C}e_2 \oplus \mathbb{C}e_3 = (\mathbb{C}^3)_{\eta^1} \oplus (\mathbb{C}^3)_{\eta^2} \oplus (\mathbb{C}^3)_{\eta^3}.$$

[9] Acting on $\mathfrak{h}$, the elements η^1, η^2, η^3 are not linearly independent, of course, since $\eta^1(H) + \eta^2(H) + \eta^3(H) = 0$ holds for any traceless diagonal matrix H.

Example II.29. Recall that if V is a representation of a Lie algebra $\mathfrak{g}$, then the dual V^* becomes a representation by defining, for any $X \in \mathfrak{g}$ and $\varphi \in V^*$, the dual element $X.\varphi$ as $v \mapsto -\varphi(X.v)$ for all $v \in V$.

The dual V^* of the standard representation $V = \mathbb{C}^3$ of $\mathfrak{sl}_3(\mathbb{C})$ is thus a three-dimensional representation. Let $\varphi_1, \varphi_2, \varphi_3 \in V^*$ be the dual basis to the standard basis $e_1, e_2, e_3 \in V$, i.e. $\varphi_j(e_i) = \delta_{i,j}$ for all $i, j \in \{1, 2, 3\}$. If $H \in \mathfrak{h}$, then

$$\big(H.\varphi_j\big)(e_i) = -\varphi_j(H.e_i) = -\varphi_j\big(\eta^i(H)\, e_i\big) = -\eta^i(H)\,\delta_{i,j} = -\eta^j(H)\,\varphi_j(e_i),$$

which implies that $H.\varphi_j = -\eta^j(H)\,\varphi_j$. The basis vectors $\varphi_1, \varphi_2, \varphi_3$ are thus weight vectors, with respective weights $-\eta^1, -\eta^2, -\eta^3$, and the weight space decomposition of the dual of the standard representation of $\mathfrak{sl}_3(\mathbb{C})$ is

$$V^* = \mathbb{C}\varphi_1 \oplus \mathbb{C}\varphi_2 \oplus \mathbb{C}\varphi_3 = (V^*)_{-\eta^1} \oplus (V^*)_{-\eta^2} \oplus (V^*)_{-\eta^3}.$$

In particular (unlike for $\mathfrak{sl}_2(\mathbb{C})$), a representation of $\mathfrak{sl}_3(\mathbb{C})$ and its dual are generally not isomorphic to each other (even the weights in V and V^* are different).

Example II.30. The adjoint representation of $\mathfrak{sl}_3(\mathbb{C})$ is the vector space $V = \mathfrak{sl}_3(\mathbb{C})$ equipped with the adjoint action: for $X \in \mathfrak{sl}_3(\mathbb{C})$ and $Y \in V = \mathfrak{sl}_3(\mathbb{C})$, we set

$$\mathrm{ad}_X(Y) = [X, Y], \qquad \text{which defines} \qquad \mathrm{ad}\colon \mathfrak{sl}_3(\mathbb{C}) \to \mathrm{End}\big(\mathfrak{sl}_3(\mathbb{C})\big).$$

This is an eight-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$.

We will next address the weight space decomposition in this case.

6.2. Representations of $\mathfrak{sl}_3(\mathbb{C})$

We will use the following two facts about finite dimensional representations of $\mathfrak{sl}_3(\mathbb{C})$.

Fact II.18. On any finite dimensional representation V of $\mathfrak{sl}_3(\mathbb{C})$, the actions of all $H \in \mathfrak{h} \subset \mathfrak{sl}_3(\mathbb{C})$ are simultaneously diagonalizable.

The proof of this fact follows from general theory of semisimple Lie algebras, but it is also not difficult to deduce from the corresponding fact for $\mathfrak{sl}_2(\mathbb{C})$. The simultaneous eigenspaces are the weight spaces (II.16) in the decomposition (II.15).

Fact II.19. Any finite-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$ is a direct sum of its irreducible subrepresentations.

This fact follows from general theory of semisimple Lie algebras, which we will treat later.

6.2.1. *The adjoint representation and roots for* $\mathfrak{sl}_3(\mathbb{C})$

In particular, the adjoint representation $V = \mathfrak{sl}_3(\mathbb{C})$ admits a decomposition to weight spaces

$$\mathfrak{sl}_3(\mathbb{C}) = \bigoplus_{\mu} \big(\mathfrak{sl}_3(\mathbb{C})\big)_{\mu}$$

as we will verify now. The (abelian) subalgebra of diagonal matrices clearly consists of vectors that have eigenvalue 0 for the adjoint action of any other diagonal matrix, so we have $\mathfrak{h} \subset \big(\mathfrak{sl}_3(\mathbb{C})\big)_0$. For an elementary matrix E^{ij}, and diagonal matrix $H = \sum_k a_k E^{kk}$, we calculate

$$\begin{aligned}[H, E^{ij}] = \sum_k a_k \, [E^{kk}, E^{ij}] &= \sum_k a_k \, (\delta_{ki} E^{kj} - \delta_{jk} E^{ik}) \\ &= (a_i - a_j) \, E^{ij},\end{aligned} \tag{II.17}$$

which shows that the one-dimensional subspace $\mathbb{C}E^{ij}$, for $i \neq j$, is a simultaneous eigenspace for all $H \in \mathfrak{h}$, with eigenvalues given by the weight $\eta^i - \eta^j \in \mathfrak{h}^*$. This in fact concludes the weight space decomposition: the eight-dimensional space $\mathfrak{sl}_3(\mathbb{C})$ has six one-dimensional weight spaces of different non-zero weights, and the two-dimensional subspace $\mathfrak{h}$ of zero weight:

$$\mathfrak{sl}_3(\mathbb{C}) = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{C}E^{ij}. \tag{II.18}$$

The non-zero weights appearing in the adjoint representation are called *roots*, and denoted traditionally by α. The set of roots is denoted by Φ: for $\mathfrak{sl}_3(\mathbb{C})$ we have

$$\Phi = \left\{ \eta^1 - \eta^2, \eta^1 - \eta^3, \eta^2 - \eta^3, \eta^2 - \eta^1, \eta^3 - \eta^1, \eta^3 - \eta^2 \right\}. \tag{II.19}$$

For the adjoint representation, the weight spaces other than $\mathfrak{h}$ are called *root spaces*. The decomposition (II.18) is also called the root space decomposition.

6.2.2. *Irreducible representations of* $\mathfrak{sl}_3(\mathbb{C})$

Let again V be a finite-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$, and assume moreover that it is irreducible. The decomposition $V = \bigoplus_\mu V_\mu$ to weight spaces

$$V_\mu = \left\{ v \in V \;\middle|\; \forall H \in \mathfrak{h} : Hv = \mu(H) v \right\}$$

tells exactly how any $H \in \mathfrak{h}$ acts on V. In view of the root space decomposition (II.18) of $\mathfrak{sl}_3(\mathbb{C})$, the remaining task is to describe how the root vectors E^{ij}, $i \neq j$, act on V.

Let now $v \in V_\mu$ be a weight vector of weight $\mu \in \mathfrak{h}^*$, and consider the action of E^{ij} on v. Denote by $\alpha^{ij} = \eta^i - \eta^j$ the corresponding root, and let $H \in \mathfrak{h}$.

Fundamental calculation (second time):

$$\begin{aligned} H(E^{ij} v) &= E^{ij}(Hv) + [H, E^{ij}] v \\ &= E^{ij}\big(\mu(H) v\big) + \alpha^{ij}(H) E^{ij} v \\ &= \big(\mu + \alpha^{ij}\big)(H) \, E^{ij} v. \end{aligned}$$

This calculation shows that if v is a weight vector with weight μ, then $E^{ij} v$ is a weight vector with weight $\mu + \alpha^{ij}$ (although not necessarily a non-zero vector). In other words, for any μ and for any $i \neq j$ we have

$$E^{ij} \colon V_\mu \to V_{\mu + \alpha^{ij}}.$$

As with $\mathfrak{sl}_2(\mathbb{C})$ we can immediately conclude something about the differences of any two weights appearing in an irreducible representation.

Observation II.20. In an irreducible representation of $\mathfrak{sl}_3(\mathbb{C})$, any two weights μ, μ' differ by an integer linear combination of roots, $\mu' = \mu + \sum_{i \neq j} n_{ij} \, \alpha^{ij}$ with some $n_{ij} \in \mathbb{Z}$.

This can be reformulated as saying that the weights in an irreducible lie in some translate of the *root lattice*

$$\Lambda_{\mathrm{R}} = \sum_{\alpha \in \Phi} \mathbb{Z}\alpha = \mathbb{Z}\alpha^{12} \oplus \mathbb{Z}\alpha^{23}. \tag{II.20}$$

For the latter expression we used the fact that $\alpha^{13} = \alpha^{12} + \alpha^{23}$, by virtue of which all roots can in fact be expressed as integer linear combinations of α^{12} and α^{23}. We call these α^{12} and α^{23} *simple roots* (a choice has been made here). The set $\Delta = \{\alpha^{12}, \alpha^{23}\}$ of simple roots forms a $\mathbb{Z}$-basis of the root lattice Λ_{R}. Roots which are non-negative (resp. non-positive) integer linear combinations of simple roots are called positive roots (resp. negative roots), and their set is denoted by

$$\Phi^+ = \Phi \,\cap\, \bigoplus_{\alpha \in \Delta} \mathbb{Z}_{\geq 0}\, \alpha \qquad (\text{resp. } \Phi^- = -\Phi^+).$$

Concretely, here we have $\Phi^+ = \{\alpha^{12}, \alpha^{23}, \alpha^{13}\} = \{\alpha^{ij} \,|\, i < j\}$.

To continue with comparisons to the case of $\mathfrak{sl}_2(\mathbb{C})$, recall that at this stage we showed that in an irreducible representation, any non-zero vector from the H-eigenspace with maximal eigenvalue λ generated the entire representation, which was in fact determined by λ. Such a vector v satisfied $Ev = 0$ and then successive action by F on v was enough to span the representation. What is the correct generalization to the present situation?

The eigenvalues have been replaced by weights $\mu \in \mathfrak{h}^*$, and it is not a priori clear which should be though of as maximal. Let us make an arbitrary looking choice: choose numbers $r_1 > r_2 > r_3$ such that $r_1 + r_2 + r_3 = 0$, and define a linear functional ℓ on $\mathfrak{h}^*$ by

$$\ell(a_1\eta^1 + a_2\eta^2 + a_3\eta^3) = a_1 r_1 + a_2 r_2 + a_3 r_3.$$

The choice made above is such that the positive roots evaluate to positive numbers, in particular for the two simple roots we have $\ell(\alpha^{12}) = r_1 - r_2 > 0$ and $\ell(\alpha^{23}) = r_2 - r_3 > 0$. Let us agree to say that a maximal weight is the one with the largest value of (the real part of) ℓ. To ensure that there is a unique maximal choice, we assume furthermore r_1, r_2, r_3 chosen so that $\ell \colon \Lambda_{\mathrm{R}} \to \mathbb{R}$ has a trivial kernel (ℓ is irrational with respect to the lattice Λ_{R}).

Then in a finite-dimensional representation V there exists a unique maximal weight, denote it by λ. Note that since $\ell(\alpha^{ij}) > 0$ for all $i < j$, we must have $E^{ij} V_\lambda = 0$. The root spaces of the positive roots thus annihilate the weight space with maximal weight. We introduce some terminology:

Definition II.21. If V is any representation of $\mathfrak{sl}_3(\mathbb{C})$, then a (non-zero) vector $v \in V$ which satisfies $E^{ij} v = 0$ for all $i < j$, and $Hv = \mu(H)v$ for all $H \in \mathfrak{h}$ and some $\lambda \in \mathfrak{h}^*$ is called a *highest weight vector*, and the weight $\lambda \in \mathfrak{h}^*$ is called its *highest weight*.

Observation II.22. In any irreducible finite-dimensional representation $V \neq 0$ of $\mathfrak{sl}_3(\mathbb{C})$, there exists a non-zero highest weight vector.

Proof. Take λ the maximal weight in $V = \bigoplus_\mu V_\mu$, and choose a non-zero $v \in V_\lambda$. □

Example II.31. In the standard representation $V = \mathbb{C}^3$ of $\mathfrak{sl}_3(\mathbb{C})$, the vector e_1 a highest weight vector of highest weight η^1.

Example II.32. In the dual V^* of the standard representation the vector φ_3 a highest weight vector of highest weight $-\eta^3$.

Example II.33. In the adjoint representation $\mathfrak{sl}_3(\mathbb{C})$, by Equations (II.12) and (II.17), the vector E^{13} a highest weight vector of highest weight $\alpha^{13} = \eta^1 - \eta^3$.

A highest weight vector $v \in V_\lambda$ is annihilated by half of the root vectors, and like for $\mathfrak{sl}_2(\mathbb{C})$, applying repeatedly on it the other half of the root vectors, we generate the entire irreducible representation.

Claim II.23. Let $0 \neq v \in V_\lambda$. Then V is spanned by the vectors obtained by successively applying E^{21}, E^{32}, and E^{31} on v.

Proof. Let W be the linear span of vectors obtained by successively applying E^{21}, E^{32}, and E^{31} on v. Note that since $E^{31} = -[E^{21}, E^{32}]$, alternatively W could have been defined as the linear span of vectors obtained by successively applying only E^{21} and E^{32} on v. For an inductive argument, let W_n denote the linear span of vectors obtained by successively applying on v a word of at most n letters, each equal to E^{21} or E^{32}. Then W is the sum of W_n, as n ranges over natural numbers. By definition we have $E^{21}W_n \subset W_{n+1}$ and $E^{32}W_n \subset W_{n+1}$, and then using the fact that $E^{31} = -[E^{21}, E^{32}]$ we get that $E^{31}W_n \subset W_{n+2}$. Also for any $H \in \mathfrak{h}$ we have $HW_n \subset W_n$, since the vector obtained by applying a word on the highest weight vector, is a weight vector (of weight λ plus the sum of the negative roots corresponding to the letters of the word), and such vectors span W_n. It follows that $W = \sum_n W_n$ is an invariant subspace for the action of all $H \in \mathfrak{h}$ and E^{21}, E^{32}, and E^{31}. It remains to see what the positive root vectors E^{12}, E^{23}, and E^{13} do to W_n. Moreover, since $E^{13} = [E^{12}, E^{23}]$, it in fact suffices to consider E^{12} and E^{23}.

We claim that $E^{12}W_n \subset W_{n-1}$ and $E^{23}W_n \subset W_{n-1}$. The proofs are entirely similar, so consider the first case. The case $n = 0$ is clear, since $W_0 = \mathbb{C}v$ is the one-dimensional space spanned by the highest weight vector, which is annihilated by E^{12} and E^{23}. Proceed by induction on n. Suppose that w is a vector obtained by applying on v a word of n letters, each equal to E^{21} or E^{32}. Depending on the last letter, we have either $w = E^{21}w'$ or $w = E^{32}w'$, with $w' \in W_{n-1}$. Consider first the first case. Then

$$\begin{aligned} E^{12}w &= E^{12}E^{21}w' = \big(E^{21}E^{12} + [E^{12}, E^{21}]\big)w' = \big(E^{21}E^{12} + H^{12}\big)w' \\ &= E^{21}E^{12}w' + H^{12}w' \ \in \ E^{21}W_{n-2} + W_{n-1} \subset W_{n-1} \end{aligned}$$

where we used the induction assumption $E^{12}W_{n-1} \subset W_{n-2}$ and the fact that $\mathfrak{h}$ preserves W_{n-1}. In the second case,

$$\begin{aligned} E^{12}w &= E^{12}E^{32}w' = \big(E^{32}E^{12} + [E^{12}, E^{32}]\big)w' = \big(E^{32}E^{12} + 0\big)w' \\ &= E^{32}E^{12}w' \ \in \ E^{32}W_{n-2} \subset W_{n-1}, \end{aligned}$$

where we again used the induction assumption $E^{12}W_{n-1} \subset W_{n-2}$. By induction, we thus establish that $E^{12}W_n \subset W_{n-1}$ and $E^{23}W_n \subset W_{n-1}$, and as a consequence also $E^{13}W_n \subset W_{n-2}$. Therefore $W = \sum_n W_n$ is invariant also for E^{12}, E^{23}, and E^{13}, and is therefore a subrepresentation. $\square$

Observation II.24. The weights μ appearing in an irreducible finite-dimensional representation V of $\mathfrak{sl}_3(\mathbb{C})$ lie in a cone (a $\frac{1}{3}$-plane, in fact) seen from the maximal weight λ, namely in

$$\lambda - (\mathbb{R}_{\geq 0}\alpha^{12} + \mathbb{R}_{\geq 0}\alpha^{23}).$$

By Observation II.22 any irreducible representation contains a highest weight vector, and by Claim II.23 the subspace spanned by vectors obtained by successively applying E^{21}, E^{32}, and E^{31} on the highest weight vector is a subrepresentation — in particular an irreducible representation is generated by successively applying E^{21}, E^{32}, and E^{31} on a highest weight vector. Actually a little more is true:

Proposition II.34. *If V is any representation of $\mathfrak{sl}_3(\mathbb{C})$, and $v \in V$ is a non-zero highest weight vector, then the subspace $W \subset V$ spanned by vectors obtained by successively applying E^{21}, E^{32}, and E^{31} on v is an irreducible subrepresentation.*

Proof. Let λ be the highest weight of v, i.e. $v \in V_\lambda$. We have shown that $W \subset V$ is a subrepresentation, and clearly W_λ is one-dimensional, $W_\lambda = \mathbb{C}v$. If W would not be irreducible, then by complete reducibility (Fact II.19) we would have $W = W' \oplus W''$, with W' and W'' non-zero subrepresentations. But since the projections to W' and W'' commute with the action of $\mathfrak{h}$, we have $W_\lambda = W'_\lambda \oplus W''_\lambda$. By one-dimensionality, one of these has to be zero, and so v belongs to either W' or W'', and thus W is either W' or W''. □

Corollary II.35. *The highest weight λ of an irreducible representation is uniquely determined, and the highest weight vector is unique up to a multiplicative constant.*

Proof. If an irreducible representation V would contain a (non-zero) highest weight vector of highest weight λ' other than the maximal weight λ (according to the ordering given by the real part of $\ell\colon \mathfrak{h} \to \mathbb{C}$), then the subrepresentation W generated by it could not contain vectors in V_λ, and thus would be a proper subrepresentation. This shows the uniqueness of the highest weight. The uniqueness up to constants of a highest weight vector follows from Claim II.23. □

Corollary II.36. *An irreducible representation of $\mathfrak{sl}_3(\mathbb{C})$ is determined by its highest weight.*

Proof. Suppose that V and W are irreducible representations with the same highest weight λ. Take non-zero highest weight vectors $v \in W$ and $w \in W$. Consider the representation $V \oplus W$, and the subrepresentation $U \subset V \oplus W$ generated by the vector $v + w$. Since $v + w$ is a highest weight vector, U is an irreducible representation by Proposition II.34. Let $\pi_V : V \oplus W \to V$ be the projection to V. Since $\pi_V(v + w) = v \neq 0$, by Schur's lemma we have $U \cong V$. Similarly one shows $U \cong W$. This shows $V \cong W$. □

Let us summarize what we know up to now about irreducible finite-dimensional representations of $\mathfrak{sl}_3(\mathbb{C})$. By Observation II.22 we know that an irreducible representation contains highest weight vectors, by Corollary II.35 we know that they have a unique highest weight, and by Corollary II.36 we know that the irreducible representation is determined by the highest weight. Thus the classification of irreducible representations has been reduced to answering:

Which elements $\lambda \in \mathfrak{h}^*$ can serve as highest weights of finite-dimensional irreducible representations?

Also we should obtain a more detailed and concrete understanding of the representation. Let us first proceed with the study of the weights and their multiplicities in a finite-dimensional irreducible highest weight representation with highest weight λ.

From Corollary II.35 we know that the multiplicity of the highest weight λ is one in an irreducible representation. Let us continue looking at the weights along the borders of the cone in which all weights of the representation are known to reside by Observation II.24. The weight space $V_{\lambda - k\,\alpha^{12}}$, $k \in \mathbb{Z}_{\geq 0}$, is necessarily spanned by $(E^{21})^k v$ (any application of E^{32} or E^{31} would move the weight away from that border of the cone). In particular, among weights of the form $\lambda - k\alpha^{12}$, there is one uninterrupted string, with $k = 0, 1, 2, \ldots, d-1$, where d is the smallest positive integer such that $(E^{21})^d v = 0$.

We can actually now apply our knowledge of $\mathfrak{sl}_2(\mathbb{C})$. Denote $H^{12} = E^{11} - E^{22} \in \mathfrak{h}$, and recall calculations (II.12) and (II.17), which give

$$[H^{12}, E^{12}] = 2\,E^{12}, \qquad [H^{12}, E^{21}] = -2\,E^{21}, \qquad [E^{12}, E^{21}] = H^{12}.$$

In other words, the span of the three elements $E^{12}, H^{12}, , E^{21}$ is a Lie subalgebra $\mathfrak{s}^{12} \subset \mathfrak{sl}_3(\mathbb{C})$ which is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.

The action of the subalgebra $\mathfrak{s}^{12} \subset \mathfrak{sl}_3(\mathbb{C})$ only shifts weights in the directions $\pm\alpha^{12}$, and the sum of weight spaces

$$\bigoplus_{k=0}^{d-1} V_{\lambda - k\,\alpha^{12}}$$

is a representation of $\mathfrak{s}^{12} \cong \mathfrak{sl}_2(\mathbb{C})$. From the previous lecture, we then know that the dimension d of it relates to the maximal H^{12} eigenvalue $\lambda(H^{12}) \in \mathbb{Z}_{\geq 0}$ by $d = \lambda(H^{12}) + 1$.

Observation II.25. The highest weight $\lambda \in \mathfrak{h}^*$ of an irreducible finite-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$ takes non-negative integer values on the basis H^{12}, H^{23} of $\mathfrak{h} \subset \mathfrak{sl}_3(\mathbb{C})$:

$$\lambda(H^{12}) = a \in \mathbb{Z}_{\geq 0}, \quad \lambda(H^{23}) = b \in \mathbb{Z}_{\geq 0},$$

and consequently also on $H^{13} = H^{12} + H^{23}$:

$$\lambda(H^{13}) = \lambda(H^{12}) + \lambda(H^{23}) = a + b \in \mathbb{Z}_{\geq 0}.$$

Proof. Indeed, from above we see that $\lambda(H^{12}) = d - 1 =: a$, where d is the dimension of the representation of $\mathfrak{s}^{12} \cong \mathfrak{sl}_2(\mathbb{C})$ consisting of weight spaces along one border of the cone in weight space. Similarly, by looking at another border of the cone and the subalgebra $\mathfrak{s}^{23} \cong \mathfrak{sl}_2(\mathbb{C})$ spanned by E^{23}, H^{23}, E^{32}, one concludes that $\lambda(H^{23})$ is a non-negative integer. □

This gives a necessary condition for an element $\lambda \in \mathfrak{h}^*$ to be the highest weight of an irreducible finite-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$. In fact, it turns out that the condition is also sufficient.

Theorem II.37. *For any $a, b \in \mathbb{Z}_{\geq 0}$, let $\lambda_{a,b} = a\,\eta^1 - b\,\eta^3 \in \mathfrak{h}^*$, i.e., $\lambda(H^{12}) = a$, $\lambda(H^{23}) = b$. Then there exists a unique irreducible finite-dimensional representation $L(\lambda_{a,b})$ of $\mathfrak{sl}_3(\mathbb{C})$ with highest weight $\lambda_{a,b}$. Moreover, any irreducible* *finite-dimensional representation of* $\mathfrak{sl}_3(\mathbb{C})$ *is isomorphic to* $L(\lambda_{a,b})$ *for some* $a, b \in \mathbb{Z}_{\geq 0}$.

Proof. We have shown all other parts of the assertion except the existence of a finite-dimensional representation with highest weight $\lambda_{a,b}$. Recall from Example **??** that the highest weight of the standard representation $V = \mathbb{C}^3$ is $\eta^1 = \lambda_{1,0}$, and the highest weight of the dual V^* is $-\eta^3 = \lambda_{0,1}$. Consider the tensor product

$$\underbrace{V \otimes \cdots \otimes V}_{a \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{b \text{ times}}$$

of a copies of V and b copies of V^*. In it, the vector

$$e_1 \otimes \cdots \otimes e_1 \otimes \varphi_3 \otimes \cdots \otimes \varphi_3$$

is annihilated by E^{12}, E^{23}, and E^{13}, and it is an eigenvector of any $H \in \mathfrak{h}$, with eigenvalue $a\,\eta^1(H) - b\,\eta^3(H)$. Therefore, this vector is a highest weight vector of highest weight $\lambda_{a,b}$, and the subrepresentation generated by it is an irreducible highest weight representation of dimension at most 3^{a+b}, the dimension of the tensor product. □

We have thus in principle classified all irreducible finite-dimensional representations, but our description of them is so far not satisfactory in terms of explicitness — we have not for example told what are the different weights appearing in the irreducible $L(\lambda_{a,b})$, or what is its dimension.

6.2.3. *More about the weights in irreducible representations*

We found that the irreducible representations were labeled by their highest weights, the possible values of which form the set of *dominant weights*

$$\Lambda_{\mathrm{W}}^{+} = \left\{ \mu \in \mathfrak{h}^* \;\middle|\; \mu(H^{12}) \in \mathbb{Z}_{\geq 0},\ \mu(H^{23}) \in \mathbb{Z}_{\geq 0} \right\}. \tag{II.22}$$

All weights must be obtained from these by translating by some integer linear combinations of roots. Since the roots α^{ij} satisfy $\alpha^{ij}(H^{12}) \in \mathbb{Z}$ and $\alpha^{ij}(H^{23}) \in \mathbb{Z}$, we see that all weights of any finite-dimensional representations of $\mathfrak{sl}_3(\mathbb{C})$ must belong to the *weight lattice*

$$\Lambda_{\mathrm{W}} = \left\{ \mu \in \mathfrak{h}^* \;\middle|\; \mu(H^{12}) \in \mathbb{Z},\ \mu(H^{23}) \in \mathbb{Z} \right\}. \tag{II.23}$$

It is useful to have in mind the picture of $\mathfrak{h}^*$ with the discrete set Φ of roots, the lattice Λ_{R} generated by them, the lattice Λ_{W} of possible weights which refines the root lattice Λ_{R}, and the cone Λ_{W}^{+} of dominant weights

$$\Phi \subset \Lambda_{\mathrm{R}} \subset \Lambda_{\mathrm{W}} \subset \mathfrak{h}^*, \qquad \text{and} \qquad \Lambda_{\mathrm{W}}^{+} \subset \Lambda_{\mathrm{W}} \subset \mathfrak{h}^*.$$

Recall also that in the irreducible representation $L(\lambda)$ with highest weight $\lambda \in \Lambda_{\mathrm{W}}^{+}$, all weights are known to lie in the cone (in fact a $\frac{1}{3}$-plane)

$$\lambda - (\mathbb{Z}_{\geq 0}\alpha^{12} + \mathbb{Z}_{\geq 0}\alpha^{23}),$$

by Observation II.24.

By Claim II.23 we got that along the borders of that cone the multiplicities of weights are equal to one, until at some point they terminate

$$\dim(L(\lambda)_{\lambda - k\alpha^{12}}) = 1, \quad \text{for } k = 0, 1, \ldots, k_{\max}$$
$$\dim(L(\lambda)_{\lambda - k\alpha^{23}}) = 1, \quad \text{for } k = 0, 1, \ldots, k'_{\max},$$

but as already indicated, we do in fact get more precise information by making use of the subalgebras $\mathfrak{s}^{12} \subset \mathfrak{sl}_3(\mathbb{C})$ and $\mathfrak{s}^{23} \subset \mathfrak{sl}_3(\mathbb{C})$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. We now turn to that.

Consider thus again for example the subalgebra $\mathfrak{s}^{12} \cong \mathfrak{sl}_2(\mathbb{C})$, spanned by E^{12}, H^{12}, and E^{21} in $\mathfrak{sl}_3(\mathbb{C})$. Suppose that $\mu \in \mathfrak{h}^*$ is a weight appearing in a representation V. The subspace

$$\bigoplus_{k \in \mathbb{Z}} V_{\mu + k\alpha^{12}}$$

consisting of the weight spaces with weight μ translated by an integer multiple of the root α^{12}, is a representation of $\mathfrak{s}^{12}$, by virtue of our "fundamental calculation" (II.20). Let us apply the symmetry of eigenvalues of $\mathfrak{sl}_2(\mathbb{C})$ to this representation. Note that the H^{12}-eigenvalue of any $w \in V_\mu$ is the integer $\tilde{\mu} = \mu(H^{12})$. Similarly, the H^{12}-eigenvalue of $w \in V_{\mu+k\alpha^{12}}$ is is $\mu(H^{12}) + k\alpha^{12}(H^{12}) = \tilde{\mu} + 2k$. The reflected weight, $-\tilde{\mu}$, in particular, is obtained by setting $k = -\tilde{\mu} = -\mu(H^{12})$. This H^{12}-eigenspace in the above representation of $\mathfrak{s}^{12}$ is the weight space with weight $\mu - \mu(H^{12})\,\alpha^{12}$. We define the operation

$$\sigma_{12} \colon \mathfrak{h}^* \to \mathfrak{h}^*, \qquad \sigma_{12}(\mu) = \mu - \mu(H^{12})\,\alpha^{12}.$$

The dimensions of the weight spaces $L(\lambda)_\mu$ and $L(\lambda)_{\sigma_{12}(\mu)}$ must be equal.

An entirely similar analysis of the subalgebra $\mathfrak{s}^{23} \cong \mathfrak{sl}_2(\mathbb{C})$, spanned by E^{23}, H^{23}, E^{32}, and of the subalgebra $\mathfrak{s}^{13} \cong \mathfrak{sl}_2(\mathbb{C})$, spanned by E^{13}, H^{13}, E^{31}, shows that the dimensions of the weight spaces are also unchanged by the operations

$$\begin{aligned} \sigma_{23} \colon \ \mathfrak{h}^* &\to \mathfrak{h}^*, & \sigma_{23}(\mu) &= \mu - \mu(H^{23})\,\alpha^{23} \\ \sigma_{13} \colon \ \mathfrak{h}^* &\to \mathfrak{h}^*, & \sigma_{13}(\mu) &= \mu - \mu(H^{13})\,\alpha^{13} \end{aligned}$$

on weights. Let $\mathcal{W}$ be the group generated by $\sigma_{12}, \sigma_{23}, \sigma_{13}$, the *Weyl group*. Then the multiplicities of weights in any representation V of $\mathfrak{sl}_3(\mathbb{C})$, are symmetric under the action of the Weyl group

$$\dim(V_\mu) = \dim(V_{\sigma(\mu)}) \qquad \text{for any } \sigma \in \mathcal{W}.$$

Note that the operation σ_{12} on $\mathfrak{h}^*$ is actually a reflection across the line determined by $\Omega_{12} = \{\mu \in \mathfrak{h}^* \mid \mu(H^{12}) = 0\}$, and similarly σ_{23} and σ_{13} are reflections across lines $\Omega_{23} = \{\mu \mid \mu(H^{23}) = 0\}$ and $\Omega_{13} = \{\mu \mid \mu(H^{13}) = 0\}$, respectively. Applying the invariance of weight multiplicities under Weyl group $\mathcal{W}$ to the adjoint representation, we find that each $\sigma \in \mathcal{W}$ permutes the set Φ of roots. As an example, we calculate

$$\begin{aligned} \sigma_{12}(\alpha^{12}) &= \alpha^{12} - \alpha^{12}(H^{12})\alpha^{12} = \alpha^{12} - 2\alpha^{12} = -\alpha^{12} \\ \sigma_{12}(\alpha^{23}) &= \alpha^{23} - \alpha^{23}(H^{12})\alpha^{12} = \alpha^{23} + \alpha^{12} = \alpha^{13} \\ \sigma_{12}(\alpha^{13}) &= \alpha^{13} - \alpha^{13}(H^{12})\alpha^{12} = \alpha^{13} - \alpha^{12} = \alpha^{23}. \end{aligned} \tag{II.24}$$

Exercise II.4. Show that the group $\mathcal{W}$ is isomorphic to the symmetric group $\mathfrak{S}_3$ on three letters.

Let us return to the analysis of the irreducible representation $V = L(\lambda)$ with highest weight $\lambda \in \Lambda^+_{\mathrm{W}}$. The highest weight vector v in the one-dimensional weight space V_λ is annihilated by E^{12} and E^{23} and consequently also by $E^{13} = [E^{12}, E^{23}]$. This lead to Observation II.24 that all weights of $L(\lambda)$ must lie in the cone

$$\lambda - (\mathbb{Z}_{\geq 0}\alpha^{12} + \mathbb{Z}_{\geq 0}\alpha^{23}).$$

Consider then a vector v' in the one-dimensional weight-space $V_{\lambda'}$ with the reflected weight $\lambda' = \sigma_{12}(\lambda)$. Note that $\sigma_{12} \circ \sigma_{12} = \mathrm{id}_{\mathfrak{h}^*}$, and therefore $\sigma_{12}(\lambda') = \lambda$. The first calculation in Equation (II.24) then has the significant consequence that

$$\sigma_{12}(\lambda' - \alpha^{12}) = \sigma_{12}(\lambda') - \sigma_{12}(\alpha^{12}) = \lambda + \alpha^{12},$$

from which we infer that $\dim(V_{\lambda'-\alpha^{12}}) = \dim(V_{\lambda+\alpha^{12}}) = \{0\}$, and in particular that $E^{21}v' = 0$. Similarly, the third calculation in Equation (II.24) implies that

$$\sigma_{12}(\lambda' + \alpha^{13}) = \sigma_{12}(\lambda') + \sigma_{12}(\alpha^{13}) = \lambda + \alpha^{23}$$

from which we infer that $\dim(V_{\lambda'+\alpha^{13}}) = \dim(V_{\lambda+\alpha^{23}}) = \{0\}$, and in particular that $E^{13}v' = 0$. Consequently we also get $E^{23}v' = 0$, since $E^{23} = [E^{21}, E^{13}]$. This makes $v' \in V_{\lambda'}$ something like a highest weight vector, but with respect to a different choice of what is meant by the maximal weight. By performing an analysis similar to Claim II.23, we may conclude that the weights of V lie in a certain cone seen from λ', namely

$$\lambda - (\mathbb{Z}_{\geq 0}\alpha^{21} + \mathbb{Z}_{\geq 0}\alpha^{13}).$$

We can play a similar game with each of the 6 elements of the Weyl group $\mathcal{W}$. Vectors v_σ, $\sigma \in \mathcal{W}$, in the one-dimensional weight spaces $V_{\sigma(\lambda)}$ are annihilated by three of the six root spaces, and all weights must lie in a cone seen from $\sigma(\lambda)$. The intersection of these cones is a hexagon in the weight lattice, whose corners are the images of the highest weight λ under the action of the Weyl group, i.e., $\sigma(\lambda)$ with $\sigma \in \mathcal{W}$. Although the general case is a genuine hexagon, note that some side length may degenerate to zero if λ lies on one of the lines across which the Weyl group generators reflect weights. This happens if either $\lambda(H^{12}) = a$ or $\lambda(H^{23}) = b$ vanishes. The hexagon then degenerates to a triangle, or even a single point in the particular case of the trivial representation $L(0) = L(\lambda_{0,0})$.

What we just did for $\mathfrak{sl}_3(\mathbb{C})$ works very similarly for all semisimple Lie algebras.

The general definition of semisimple Lie algebra is given in the exercises. However, there turns out to be a convenient equivalent characterization, which makes them very concrete.

<u>Def</u>: A Lie algebra $\mathfrak{g}$ is <u>simple</u> if it is not abelian (i.e. $[\mathfrak{g},\mathfrak{g}] \neq \{0\}$) and it has no other ideals except $\{0\}$ and $\mathfrak{g}$.

<u>Fact</u>: The simple Lie algebras are (up to isomorphism):

- $\mathfrak{sl}_n(\mathbb{C})$, $n = 2, 3, \dots$
- $\mathfrak{sp}_{2n}(\mathbb{C}) = \{ X \in \mathbb{C}^{2n \times 2n} \mid X^T J + J X = 0 \}$ $n = 1, 2, \dots$

  where $J = \begin{bmatrix} 0_{n\times n} & \mathbb{1}_{n\times n} \\ -\mathbb{1}_{n\times n} & 0_{n\times n} \end{bmatrix}$
- $\mathfrak{so}_n(\mathbb{C}) = \{ X \in \mathbb{C}^{n\times n} \mid X^T + X = 0 \}$ $n = 8, 9, 10, \dots$ or $n = 5$ or $n = 7$
- five "exceptional" simple Lie algebras called $\mathfrak{e}_6, \mathfrak{e}_7, \mathfrak{e}_8, \mathfrak{f}_4, \mathfrak{g}_2$.

<u>Def</u> (Equivalent to the standard definition of semisimple Lie algebras given in the exercises)

A Lie algebra $\mathfrak{g}$ is <u>semisimple</u> if it is a direct sum of simple Lie subalgebras.

ANALYZING SEMISIMPLE LIE ALGEBRAS IN GENERAL

We give a "receipe" to analyze a semisimple Lie algebra $\underline{g}$, very similar to what we did with $\underline{sl}_3(\mathbb{C})$. One can carry out the analysis concretely for each of the possible cases listed above, or one can develop general theory to show that the "receipe" will have to work.

Step 1 Find a Cartan subalgebra $\underline{h} \subset \underline{g}$.

- $\underline{h}$ should be a maximal abelian Lie subalgebra in $\underline{g}$, which acts diagonalizably in some (in fact all) faithful representation.

Step 2 Decompose the adjoint representation to weight spaces for $\underline{h}$: weight $\mu \in \underline{h}^*$

$$\underline{g}_\mu = \{ X \in \underline{g} \mid \forall H \in \underline{h} : [H, X] = \mu(H) X \}.$$

- $\underline{g}_0 = \underline{h}$ (indeed $\underline{h} \subset \underline{g}_0$ by abelianity, and then by maximality $\underline{h} = \underline{g}_0$)
- $\Phi := \{ \alpha \in \underline{h}^* \setminus \{0\} \mid \underline{g}_\alpha \neq \{0\} \}$ the set of roots i.e. non-zero weights in the adjoint rep.
- Facts: * $\dim(\underline{g}_\alpha) = 1$ for all $\alpha \in \Phi$.

  * $\Lambda_R = \{ \sum_{\alpha \in \Phi} n_\alpha \cdot \alpha \mid n_\alpha \in \mathbb{Z} \}$ root lattice is a lattice of rank equal to $\dim(\underline{h})$

  Terminology: $\dim(\underline{h}) = \operatorname{rank}(\Lambda_R) =$ the rank of $\underline{g}$

  * if $\alpha \in \Phi$ then $-\alpha \in \Phi$ also.

$\rightsquigarrow$ root space decomposition $\underline{g} = \underline{h} \oplus \left(\bigoplus_{\alpha \in \Phi} \underline{g}_\alpha \right)$

Remark: At this stage we can conclude that if a rep. V of $\mathfrak{g}$ has the weight space decomposition

$$V = \bigoplus_{\mu} V_\mu \quad , \quad V_\mu := \{ v \in V \mid H.v = \mu(H)\, v \quad \forall H \in \underline{h} \} \quad (\mu \in \underline{h}^*)$$

then root vectors $E_\alpha \in \mathfrak{g}_\alpha$ map

$$E_\alpha : V_\mu \longrightarrow V_{\mu+\alpha} \qquad (\text{"fundamental calculation"}).$$

In particular, for any weight μ of V, the subspace

$$\bigoplus_{\beta \in \Lambda_R} V_{\mu+\beta} \subset V$$

is a subrepresentation. In an irreducible representation, therefore, weights differ by elements in the root lattice Λ_R.

Step 3 Find the distinguished subalgebras $\underline{s}_\alpha \subset \mathfrak{g}$, $s_\alpha \cong \underline{sl}_2(\mathbb{C})$.

- Note: $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$ and in particular $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \underline{h}$
- Fact $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \neq \{0\}$

  Thus one can find $E_\alpha \in \mathfrak{g}_\alpha$, $F_\alpha \in \mathfrak{g}_{-\alpha}$, $H_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$

  s.t. $[H_\alpha, E_\alpha] = 2E_\alpha$, $[H_\alpha, F_\alpha] = -2F_\alpha$, $[E_\alpha, F_\alpha] = H_\alpha$.

$$\underline{s}_\alpha = \operatorname{span}\{E_\alpha, H_\alpha, F_\alpha\} \cong \underline{sl}_2(\mathbb{C}).$$

Remark: For any $\alpha \in \Phi$, then, and any weight μ in a representation $V = \bigoplus_{\mu} V_\mu$, the subspace

$$\bigoplus_{k \in \mathbb{Z}} V_{\mu+k\alpha} \quad \text{is a rep. of } \underline{s}_\alpha \cong \underline{sl}_2(\mathbb{C}).$$

In particular the H_α eigenvalues $\mu(H_\alpha)$ are integers.

The weight lattice Λ_W is the set of such weights

$$\Lambda_W = \{ \mu \in \underline{h}^* \mid \mu(H_\alpha) \in \mathbb{Z} \quad \forall \alpha \in \Phi \}.$$

We have $\Lambda_R \subset \Lambda_W \subset \underline{h}^*$, and the weights μ appearing in any representation must satisfy $\mu \in \Lambda_W$.

Step 5: Use the symmetry of the eigenvalues of H_α

- The eigenvalues $\tilde{\mu} = \mu(H_\alpha)$ of H_α in $V = \bigoplus_\mu V_\mu$ must be symmetric w.r.t. origin $(\tilde{\mu} \mapsto -\tilde{\mu})$ and their multiplicities are also symmetric. The operation $\sigma_\alpha : \underline{h}^* \to \underline{h}^*$ defined by

$$\sigma_\alpha(\mu) = \mu - \mu(H_\alpha) \cdot \alpha$$

implements this reflection. The group W generated by σ_α, $\alpha \in \Phi$, is called the Weyl group.
We have $\dim(V_\mu) = \dim(V_{\sigma(\mu)})$ for any $\sigma \in W$.

Step 6 Draw a picture of $\Phi \subset \Lambda_R \subset \Lambda_W \subset \underline{h}^*$.

Step 7 Choose a direction in $\underline{h}^*$.

- Functional $\ell : \underline{h}^* \to \mathbb{C}$ (real on Λ_R) s.t. $\ell|_{\Lambda_R}$ has trivial kernel.

Remark If λ is the weight of V with maximal $\mathrm{Re}(\lambda)$, then any vector $v \in V_\lambda \subset V$ is a highest weight vector:

$$H.v = \lambda(H)\, v \qquad \forall H \in \underline{h}$$
$$E_\alpha . v = 0 \qquad \forall \alpha \in \Phi^+$$

where $\Phi^+ = \{\alpha \in \Phi \mid \ell(\alpha) > 0\}$.

Proposition: (i) Every finite-dim. rep. V of $\underline{g}$ contains a non-zero highest weight vector.

(ii) The subspace W spanned by vectors obtained by successive application of F_α, $\alpha \in \Phi^+$, on a h.w. vect. is an irreducible subrep.

(iii) All irreducible representations possess a unique h.w. vect up to scalar multiples.

Remark: A highest weight λ must satisfy $\lambda(H_\alpha) \in \mathbb{Z}_{\geq 0}$.

Step 8: Construct the irreducible representations with highest weights λ s.t. $\lambda(H_\alpha) \in \mathbb{Z}_{\geq 0}$ $\forall \alpha \in \Phi$.

COMPLETE REDUCIBILITY

Consider representation V and W of $\mathfrak{g}$.

Recall:

- $Hom(V,W)$ is a rep. of $\mathfrak{g}$ by
$$(X.T)(v) = X.(T(v)) - T(X.v)$$
- $$\underbrace{Hom_{\mathfrak{g}}(V,W)}_{\text{intertwining maps}} = \underbrace{Hom(V,W)^{\mathfrak{g}}}_{\{T \text{ s.t. } X.T=0 \ \forall X\in\mathfrak{g}\}}$$
$\quad = \text{invariant (trivial) part of the rep. } Hom(V,W)$

In the proof of complete reducibility, we end up considering the situation treated in the next Lemma.

Lemma Let V be a rep. of $\mathfrak{g}$, and $W \subset V$ a subrepresentation. Denote by
$$r : Hom(V,W) \longrightarrow Hom(W,W)$$
the restriction map: $r(T) = T_{|W}$ for $T: V \to W$.
Denote by $R \subset Hom(V,W)$ the space of maps whose restriction is a multiple of identity
$$R = r^{-1}(\mathbb{C}\cdot id_W)$$
and by $R_0 = Ker(r)$ those with zero restriction.
Then we have

(a) $Im(r_{|R}) = \mathbb{C}\, id_W$
(b) $R \subset Hom(V,W)$ is a subrepresentation
(c) $R_0 \subset R$ is a subrepresentation
(d) R/R_0 is a trivial one-dimensional rep.

Proof: By def. $\mathrm{Im}(r|_R) \subset \mathbb{C}\,\mathrm{id}_W$ and the image of any projection $V \to W$ is id_W. This shows (a). We also get

$$\dim(R/R_0) = \dim(R/\mathrm{Ker}(r|_R)) = \dim(\mathrm{Im}(r|_R)) = 1.$$

Now calculate, for $T \in R$, $T|_W = \lambda \cdot \mathrm{id}_W$,

$$(X.T)(w) = X.(\underbrace{T(w)}_{=\lambda w}) - \underbrace{T(\underbrace{X.w}_{\in W})}_{=\lambda X.w} = 0 \qquad \forall X \in \mathfrak{g} \;\; \forall w \in W.$$

This directly implies (b), (c), and (d). $\square$

With the help of this Lemma, we find a convenient criterion which ensures complete reducibility.

Proposition Suppose that $\mathfrak{g}$ has the following property:

(⊛) Whenever U is a rep. of $\mathfrak{g}$ and $U_0 \subset U$ is a subrep. such that U/U_0 is one-dim. and trivial, then $U = U_0 \oplus P$ (as rep.)

Then any finite dimensional representation of $\mathfrak{g}$ is completely reducible. (direct sum of irreducibles)

Proof: As usual, it suffices to show that in a finite-dim. rep. V, any subrep. W has a complementary subrep. W' s.t. $V = W \oplus W'$, which in turn is equivalent to the existence of a projection $p: V \to W$ which is $\mathfrak{g}$-intertwining.

So let $W \subset V$ as above, and construct $R_0 \subset R \subset \mathrm{Hom}(V,W)$ as in the Lemma. The property (⊛) applies to $R = U$, $R_0 = U_0$, and shows that $R = R_0 \oplus P$, with P one-dimensional and trivial. Choose a non-zero $p \in P$, normalized such that $p|_W = \mathrm{id}_W$. This projection $p: V \to W$ is $\mathfrak{g}$-intertwining since P is trivial. $\square$

The quadratic Casimir element

Suppose that a Lie algebra $\mathfrak{g}$ has a bilinear form

$$B : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{C}$$

such that

- $B(X,Y) = B(Y,X) \quad \forall X,Y \in \mathfrak{g}$ ("symmetric")
- if $B(X,Y)=0 \quad \forall Y \in \mathfrak{g}$ then $X=0$ ("non-degenerate")
- $B([Z,X],Y) + B(X,[Z,Y]) = 0 \quad \forall X,Y,Z \in \mathfrak{g}$ ("invariant").

Let $X_1, \dots, X_d$ be a basis of $\mathfrak{g}$ and $X_1', \dots, X_d'$ the dual basis with respect to B, i.e. $B(X_i, X_j') = \delta_{ij} \quad \forall i,j = 1, \dots, d$.

Lemma If V is a representation of $\mathfrak{g}$ and we set $Q_V(v) = \sum_{j=1}^{d} X_j X_j' v$ for all $v \in V$, then $Q_V : V \to V$ is a $\mathfrak{g}$-intertwining map.

Proof: Let us compute, for a given $v \in V$ and $Z \in \mathfrak{g}$

$$Z.Q_V(v) - Q_V(Z.v) = \sum_j (Z X_j X_j' - X_j X_j' Z) v$$

$$= \sum_j \left([Z,X_j] X_j' + X_j [Z, X_j'] \right) v .$$

If $[Z,X_j] = \sum_k \xi_{kj} X_k$ and $[Z, X_i'] = \sum_\ell \eta_{\ell i} X_\ell'$

then $\xi_{mj} = \sum_k \xi_{kj} B(X_k, X_m') = B([Z,X_j], X_m')$

$$= -B(X_j, [Z, X_m']) = -\sum_\ell \eta_{\ell m} B(X_j, X_\ell') = -\eta_{jm} .$$

Therefore,

$$Z.Q_V(v) - Q_V(Z.v) = \sum_j \left(\sum_k \xi_{kj} X_k X_j' + \sum_\ell \eta_{\ell j} X_j X_\ell' \right) v$$

$$= -\sum_{j,k} \eta_{jk} X_k X_j' v + \sum_{j,\ell} \eta_{\ell j} X_j X_\ell' v = 0 \qquad \square$$

<u>Examples</u>:

If V is any representation of $\mathfrak{g}$, $\rho : \mathfrak{g} \to \mathrm{End}(V)$ and we define $B_V : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{C}$ by

$$B_V(X,Y) = \mathrm{tr}_V(\rho(X) \circ \rho(Y)),$$

then B_V is symmetric (by $\mathrm{tr}(AB) = \mathrm{tr}(BA)$) and invariant, since

$$B_V([Z,X],Y) = \mathrm{tr}_V(\rho(Z)\rho(X)\rho(Y) - \rho(X)\rho(Z)\rho(Y))$$
$$= \mathrm{tr}_V(\rho(X)\rho(Y)\rho(Z) - \rho(X)\rho(Z)\rho(Y)) = B_V(X,[Y,Z]).$$

The question is whether for some choice of V the form B_V is non-degenerate.

<u>Remark</u>: For B_V to be non-degenerate, it is at least necessary for the representation $\rho : \mathfrak{g} \to \mathrm{End}(V)$ to be faithful, i.e. $\mathrm{Ker}(\rho) = \{0\}$.

Usually the good choices of V are the standard representation or the adjoint representation.

<u>Example</u> The Lie algebra $\underline{su}_2$ had a basis

$$S_x = -\frac{i}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad S_y = -\frac{i}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad S_z = -\frac{i}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

In the standard representation $\mathbb{C}^2$ of $\underline{su}_2$ we have

$$B_{\mathbb{C}^2}(S_x,S_x) = B_{\mathbb{C}^2}(S_y,S_y) = B_{\mathbb{C}^2}(S_z,S_z) = -\frac{1}{2},$$

and $B_{\mathbb{C}^2}(S_x,S_y) = B_{\mathbb{C}^2}(S_x,S_z) = B_{\mathbb{C}^2}(S_y,S_z) = 0.$

The form $B_{\mathbb{C}^2}$ is negative-definite on the real Lie algebra $\underline{su}_2$, and in particular non-degenerate.

In the adjoint representation $\underline{su}_2$ with basis S_x, S_y, S_z we had

$$ad_{S_x} = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}}_{=R_x} \qquad ad_{S_y} = \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}}_{=R_y} \qquad ad_{S_z} = \underbrace{\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{=R_z}$$

which also is a basis of $\underline{so}_3$. The form reads

$$B_{\underline{su}_2}(S_x, S_x) = B_{\underline{su}_2}(S_y, S_y) = B_{\underline{su}_2}(S_z, S_z) = -2$$

$$B_{\underline{su}_2}(S_x, S_y) = B_{\underline{su}_2}(S_x, S_z) = B_{\underline{su}_2}(S_y, S_z) = 0 .$$

We find that $B_{\underline{su}_2} = 4 \cdot B_{\mathbb{C}^2}$.

This latter calculation is also the form $B_{\mathbb{C}^3}$ of $\underline{so}_3$ in the standard representation $\mathbb{C}^3$.

<u>Example</u> Consider $\underline{sl}_n(\mathbb{C}) = \{ X \in \mathbb{C}^{n \times n} \mid tr(X) = 0 \}$, and the standard representation $\mathbb{C}^n$, giving the form $B_{\mathbb{C}^n}(X, Y) = tr(XY)$.

Recalling that $E^{ij} E^{kl} = \delta_{jk} E^{il}$ we see that

$$B_{\mathbb{C}^n}(E^{ij}, E^{kl}) = \delta_{jk} \delta_{il} .$$

The root vector E^{ij}, $i \neq j$, is "orthogonal" to everything except the root vector E^{ji} which has the opposite root, $\alpha^{ji} = -\alpha^{ij}$.

The Cartan subalgebra $\underline{h} \subset \underline{sl}_n(\mathbb{C})$ is orthogonal to all root spaces, and itself has the non-degenerate form $B_{\mathbb{C}^n}\left(\sum_k a_k E^{kk}, \sum_l b_l E^{ll}\right) = \sum_k a_k b_k$.

We see that $B_{\mathbb{C}^n}$ is non-degenerate.

Def: A Lie algebra $\mathfrak{g}$ is *simple* if it has no nontrivial ideals, and $\dim(\mathfrak{g}) > 1$.

Remark:
- Any non-trivial representation $\rho : \mathfrak{g} \to \mathrm{End}(V)$ of a simple Lie algebra is faithful, $\mathrm{Ker}(\rho) = \{0\}$ ($\mathrm{Ker}(\rho) \subset \mathfrak{g}$ is an ideal).
- In any representation V of a simple Lie algebra $\mathfrak{g}$, the form B_V is either non-degenerate or identically zero. Indeed,

$$\mathfrak{a} = \{X \in \mathfrak{g} \mid B_V(X,Y) = 0 \quad \forall Y \in \mathfrak{g}\}$$

is an ideal, since by invariance if $X \in \mathfrak{a}$ then for any $Z \in \mathfrak{g}$ and all $Y \in \mathfrak{g}$

$$B_V([Z,X], Y) = -B_V(X, [Z,Y]) = 0$$

so that $[Z,X] \in \mathfrak{a}$ also.

Exercise: Verify that $\mathfrak{sl}_n(\mathbb{C})$ $(n \geq 2)$ is simple.

Fact: Simple Lie algebras can be classified. They are the "classical" Lie algebras
- A_r : $\mathfrak{sl}_{r+1}(\mathbb{C})$, $r = 1,2,3,\ldots$
- B_r : $\mathfrak{so}_{2r+1}(\mathbb{C})$, $r = 2,3,4,\ldots$ "odd orthogonal Lie algebra"
- C_r : $\mathfrak{sp}_{2r}(\mathbb{C})$, $r = 3,4,5,\ldots$ "symplectic Lie algebra"
- D_r : $\mathfrak{so}_{2r}(\mathbb{C})$, $r = 4,5,6,\ldots$ "even orthogonal Lie algebra"

and five "exceptional" Lie algebras
- E_6, E_7, E_8
- F_4
- G_2

To finish the proof of complete reducibility, we will employ one general result, which is the topic of the next lecture:

<u>Cartan's solvability criterion</u> Let $\mathfrak{g} \subset \mathfrak{gl}(V)$ be a Lie algebra, and let $B_V : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ be the bilinear form associated with the "standard" representation V of $\mathfrak{g}$. Then $\mathfrak{g}$ is solvable if and only if $B_V(\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]) = 0$.

The way to use this is the following.

<u>Lemma</u> Let $\mathfrak{g}$ be a simple Lie algebra, and $\rho : \mathfrak{g} \to \mathrm{End}(V)$ a representation of $\mathfrak{g}$. Then either the representation is trivial $(\mathrm{Ker}(\rho) = \mathfrak{g})$ or the form B_V is non-degenerate.

<u>Proof</u>: $\mathrm{Ker}(\rho) \subset \mathfrak{g}$ is an ideal, so by simplicity of $\mathfrak{g}$ either ρ is trivial $(\mathrm{Ker}(\rho) = \mathfrak{g})$ or ρ is faithful $(\mathrm{Ker}(\rho) = \{0\})$. In the latter case we may realize $\mathfrak{g} \subset \mathfrak{gl}(V)$. Then the kernel of the form B_V, $\{X \in \mathfrak{g} \mid B_V(X, Y) = 0 \ \forall Y \in \mathfrak{g}\}$ is an ideal in $\mathfrak{g} \subset \mathfrak{gl}(V)$, which by Cartan's solvability criterion is solvable. However, the (semi)simple Lie algebra $\mathfrak{g}$ contains no non-zero solvable ideals, so the kernel of the form B_V is zero, i.e. the form is non-degenerate. $\square$

Lemma: Let $\mathfrak{g}$ be a simple Lie algebra. Suppose that V is a representation of $\mathfrak{g}$, and $W \subset V$ is an irreducible subrepresentation such that V/W is a trivial representation. Then $V = W \oplus W'$ (as a representation).

Proof: If V is trivial, the claim is obvious, so suppose V is non-trivial. Then the form $B_V(X,Y) = \mathrm{tr}_V(XY)$ is non-degenerate. Let $X_1, \ldots, X_d$ be a basis of $\mathfrak{g}$ and let $X_1', \ldots, X_d'$ be the dual basis w.r.t. B_V, i.e. $B_V(X_i, X_j') = \delta_{ij}$. Consider $Q_V : V \to V$ defined by

$$Q_V(v) = \sum_{i=1}^{d} X_i X_i' v \qquad \text{for } v \in V.$$

Then by triviality of V/W, $\mathrm{Im}(Q_V) \subset W$. Also Q_V is $\mathfrak{g}$-intertwining, so by Schur's lemma on W it acts on the irreducible subrep. $W \subset V$ as a scalar, $(Q_V)|_W = q \cdot \mathrm{id}_W$. By a calculation

$$q \cdot \dim(W) = \mathrm{tr}_V(Q_V) = \sum_{i=1}^{d} \mathrm{tr}_V(X_i X_i') = \sum_{i=1}^{d} B_V(X_i, X_i') = d = \dim(\mathfrak{g})$$

so the scalar is non-vanishing, $q \neq 0$. Thus $\frac{1}{q} Q_V : V \to W$ is a projection and $W' = \mathrm{Ker}(\frac{1}{q} Q_V)$ is a complementary subrep. $\square$

<u>Theorem</u> Let $\mathfrak{g}$ be a simple Lie algebra. Then any finite-dimensional representation of $\mathfrak{g}$ is completely reducible (direct sum of irreducible subrep.s).

<u>Proof</u>: It suffices to check the property in the Proposition:

(*) Whenever U is a rep. and U_0 is a subrep. such that U/U_0 is trivial 1-dim., then $U = U_0 \oplus P$.

By the previous lemma this holds if U_0 is irreducible. We do the case of general U_0 by induction on the dimension $\dim(U_0)$.
If U_0 is not irreducible, then choose an irreducible subrep. $W \subset U_0$, $W \neq \{0\}$, $W \neq U_0$.
Consider U/W and its subrep. U_0/W.
We have $(U/W)/(U_0/W)$ is 1-dim. trivial and $\dim(U_0/W) < \dim(U_0)$. By induction we can assume $U/W = U_0/W \oplus Y/W$ for some subrep. $Y \subset U$. But Y/W is trivial and W is irreducible so $Y = W \oplus W'$ by the previous lemma.
Then $U = U_0 \oplus W'$. $\square$

The proof of complete reducibility works almost without modification for semisimple $\mathfrak{g}$, but we only considered simple $\mathfrak{g}$ for clarity.