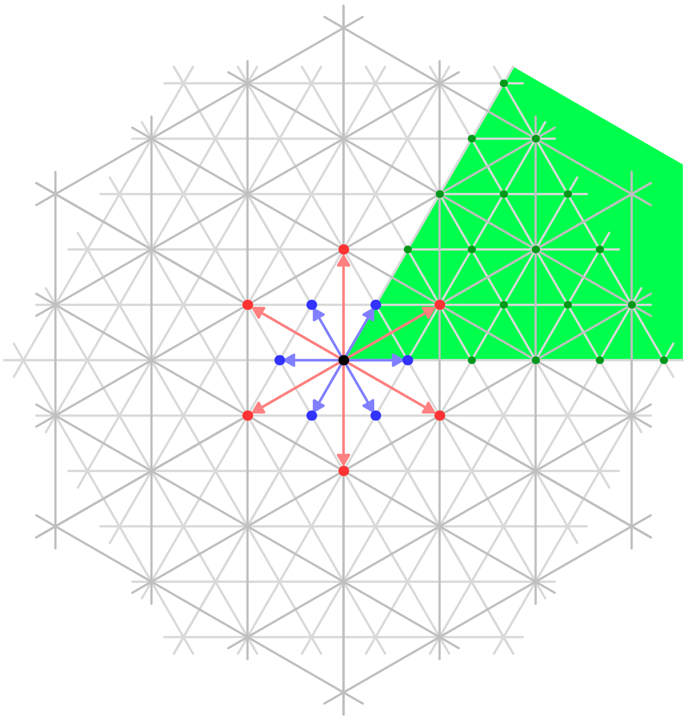
An introduction to
Lie groups and Lie algebras
and representation theory



Kalle Kytölä

February 23, 2026

Contents

Foreword

These lecture notes are primarily intended for the course **MS-E1200** *Lie groups and Lie algebras* at Aalto University.

The notes are still in a preliminary and incomplete form, and I plan to frequently update them during the current course. You could help me — or perhaps more importantly the students who will use this material — by sending comments about mistakes, misprints, needs for clarification, etc., to me (`kalle.kytola@aalto.fi`).

Appendix A contains some relevant background material on linear algebra. In particular, Appendix A.1 discusses tensor products of vector spaces, and Appendix A.2 discusses the diagonalization of matrices, the Jordan normal form, and the Cayley-Hamilton theorem. In the main material such background is assumed.

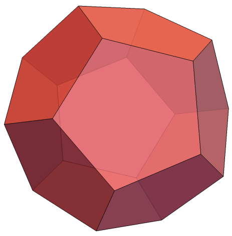
As the two main textbooks for the course, we recommend [**FH91**] and [**Sim96**].

Introduction: the mathematics of symmetries

On a general level, this course is about the mathematics of symmetries. So let us start by discussing what is meant by symmetry, and by describing how to mathematically study symmetries.

A symmetry is always a symmetry of something, of some specific (but possibly abstract) object. Familiar, concrete types of objects which have symmetries include, e.g., physical objects or geometric shapes. More abstract types of objects with symmetries could include mathematical equations, physical theories, spaces (e.g. topological spaces, vector spaces, ...), etc. In principle just about any object can have symmetries! As a few examples that may be helpful to keep in mind during this discussion, consider

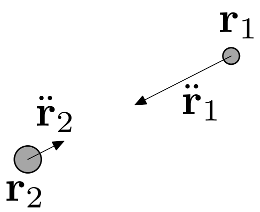
(i) the regular dodecahedron (one of the Platonic solids),



,

FIGURE O.1. Dodecahedron.

(ii) Newton's law of gravitation (for two bodies in the three-space $\mathbb{R}^3$)



$$\begin{cases} \ddot{\mathbf{r}}_1 = m_2 G \, \frac{\mathbf{r}_2 - \mathbf{r}_1}{\|\mathbf{r}_2 - \mathbf{r}_1\|^3} \\ \ddot{\mathbf{r}}_2 = m_1 G \, \frac{\mathbf{r}_1 - \mathbf{r}_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3}, \end{cases} \tag{O.1}$$

(iii) the complex upper half plane

$$\mathbb{H} := \left\{ z \in \mathbb{C} \,\middle|\, \mathfrak{Im}(z) > 0 \right\} \tag{O.2}$$

equipped with the Riemannian metric such that the length of a smooth path $\gamma : [0, 1] \to \mathbb{H}$ is given by

$$\ell(\gamma) := \int_0^1 \frac{|\dot{\gamma}(t)|}{\mathfrak{Im}(\gamma(t))} \, \mathrm{d}t. \tag{O.3}$$

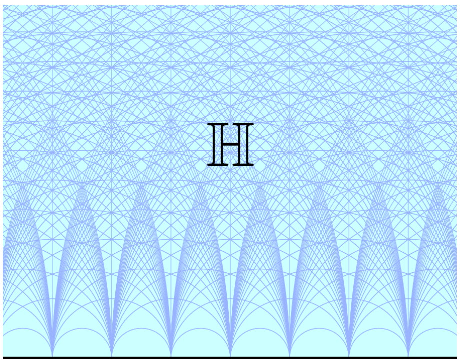


FIGURE O.2. The upper half plane $\mathbb{H}$.

Each of these objects has interesting symmetries, some more apparent than others.

So what is a symmetry? You may notice that there are certain transformations that you can do to the objects in the above examples without altering their essential features. Therefore it seems natural to say that a symmetry is a collection of transformations of the object, which leaves some relevant property of the object unchanged. Typically the collection of transformations forms a group, and a property that is unchanged by the transformations is called an invariant (for the collection of transformations in question). It is useful, though, to allow the transformations to act not necessarily on the object itself, but possibly on something else related to the original object. Let us describe a few examples:

- The symmetry group of a regular polyhedron acts by bijective maps of the sets of vertices/edges/faces of the polyhedron. In the example of the dodecahedron, Figure O.1, the symmetry group[1] thus acts either on the set of 12 faces, on the set of 30 edges, or on the set of 20 vertices of the dodecahedron.
- If a pair of trajectories $\mathbf{r}_1 \colon [0,\infty) \to \mathbb{R}^3$, $\mathbf{r}_2 \colon [0,\infty) \to \mathbb{R}^3$ satisfy Newton's law of gravitation (O.1), and $f \colon \mathbb{R}^3 \to \mathbb{R}^3$ is an Euclidean motion $f(\mathbf{x}) = \mathbf{c} + R\mathbf{x}$ with R a rotation matrix, then also the pair $\tilde{\mathbf{r}}_1 := f \circ \mathbf{r}_1$, $\tilde{\mathbf{r}}_2 := f \circ \mathbf{r}_2$ can be seen to satisfy (O.1). In other words, the group of Euclidean motions[2] is a symmetry of (the space of solutions to) Newton's law of gravitation.
- Any Möbius transformation $z \mapsto \frac{az+b}{cz+d}$ with $a,b,c,d \in \mathbb{R}$ and $ad - bc > 0$ maps the upper half plane $\mathbb{H}$ to itself. The group of such Möbius transformations[3] thus acts on the points of $\mathbb{H}$. It moreover acts on many associated objects. As an example, it acts of continuously differentiable paths $\gamma \colon [0,1] \to \mathbb{H}$ by post-composition, transforming γ to $t \mapsto \frac{a\gamma(t)+b}{c\gamma(t)+d}$. This action can straightforwardly be checked to preserve the lengths (O.3) of paths. Also, it acts on vector fields on $\mathbb{H}$ by push-forward, differential forms by pull-backs, etc. This group of Möbius transformations can indeed be interpreted as the group of symmetries of the upper half plane $\mathbb{H}$ viewed as a Riemannian manifold with the hyperbolic metric.

[1] The symmetry group of the dodecahedron is isomorphic to a semidirect product $\mathfrak{A}_5 \rtimes (\mathbb{Z}/2\mathbb{Z})$.
[2] The group of Euclidean motions is (isomorphic to) a semidirect product $\mathbb{R}^3 \rtimes \mathrm{SO}(3)$.
[3] The group of Möbius transformations of the upper half plane is (isomorphic to) $\mathrm{PSL}(2,\mathbb{R})$.

- On many resonable spaces of functions of n variables, the group of permutations of the variables[4] acts naturally. As a special case, for functions f of two variables, we have the transformation $f \mapsto \tau.f$ defined by transposition of the variables
$$(\tau.f)(x_1, x_2) := f(x_2, x_1).$$
If the functions take values in a vector space[5] (e.g., the familiar cases of real-valued or complex-valued functions), then we could consider separately functions which are symmetric,
$$f(x_2, x_1) = f(x_1, x_2),$$
and functions which are antisymmetric,
$$f(x_2, x_1) = -f(x_1, x_2).$$
One many note that any function decomposes as a sum of its symmetric and antisymmetric part. Can you think of generalizations of this to more than two variables?

Representation theory is concerned with the case when the symmetry transformations act linearly on a vector space. At first, this may appear as a restrictive special case, so the question is — why is it worthwhile to study on its own right? Here are a few possible answers:

- If we manage to make the symmetries act on vector spaces, then we obtain concrete implementations of the symmetries as linear operators or matrices.
- It is often *naturally* the case that the transformations act on a vector space: think of for example
  - if transformations a priori act on the object itself, then it is easy to see that they also act on the vector space of $\mathbb{R}$-valued or $\mathbb{C}$-valued functions defined on the object, and this action is linear;
  - physical states of a quantum mechanical system are vectors in a Hilbert space, so physical symmetries should act on these states;
  - transition probabilities of a Markov process are encoded in a matrix acting on a vector space;
  - etc, etc. . .
- In this case of linear actions on vector spaces we can develop a powerful mathematical theory which has many applications!

The introductory discussion here has intentionally been rather vague — the goal was to just offer some general perspective. It is now time to start addressing the topic more precisely.

[4]The group of permutations is the symmetric group $\mathfrak{S}_n$ (recalled precisely in Example I.3).
[5]For simplicity, assume that the vector space is over a field of characteristic different from 2.

Part I

Representations of finite groups

1. Review of groups and related concepts

Let us begin with a quick review of a few key concepts of group theory.

1.1. The definition of a group

Definition I.1 (Group).
A **group** is a pair $(G, *)$, where G is a set and $*$ is a binary operation on G

$$*\colon G \times G \to G \qquad (g, h) \mapsto g * h$$

such that the following hold

associativity:

$$g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3 \qquad \text{for all } g_1, g_2, g_3 \in G$$

neutral element: there exists an $e \in G$ such that

$$g * e = g \quad \text{and} \quad e * g = g \qquad \text{for all } g \in G$$

inverse elements: for all $g \in G$ there exists a $g^{-1} \in G$ such that

$$g * g^{-1} = e \quad \text{and} \quad g^{-1} * g = e.$$

A group $(G, *)$ is said to be finite if its **order** $|G|$ (that is the cardinality of the set G) is finite.

We usually omit the notation for the binary operation $*$ and write simply $gh := g*h$. The additive symbol $+$ is sometimes used instead for binary operations in **abelian** (i.e. **commutative**) groups.[1] We also usually abbreviate and talk about a group G instead of $(G, *)$, assuming that the binary operation in clear from the context.

1.2. Examples of groups

Example I.2 (Some abelian groups).
The following are abelian groups:

- A vector space V, with the binary operation $+$ of vector addition.
- The set $\mathbb{K} \setminus \{0\}$ of nonzero numbers in a field $\mathbb{K}$, with the binary operation of multiplication.
- The **infinite cyclic group** $\mathbb{Z}$ of integers, with the binary operation of addition.

[1] A group $(G, *)$ is said to be abelian (or commutative) if $g * h = h * g$ for all $g, h \in G$.

- For a given $N \in \mathbb{Z}_{>0}$, the group of all complex N^{th} roots of unity
$$\left\{ e^{2\pi \mathbb{i} k/N} \;\middle|\; k = 0, 1, 2, \ldots, N-1 \right\} \subset \mathbb{C},$$
with the binary operation of complex multiplication. This finite group is isomorphic to the **cyclic group** $\mathbb{Z}/N\mathbb{Z}$ of order N.

Example I.3 (Symmetric groups)**.**
Let X be a set. The **symmetric group** on X is the set
$$\mathfrak{S}(X) := \{\sigma \colon X \to X \text{ bijective}\},$$
with composition of functions as the binary operation.

In the case $X = \{1, 2, \ldots, n\}$ we denote the symmetric group by $\mathfrak{S}_n$, and we often refer to this case as the **symmetric group on** n **letters**, and the elements $\sigma \in \mathfrak{S}_n$ as **permutations** (of n letters). The group $\mathfrak{S}_n$ is a finite group of order $|\mathfrak{S}_n| = n!$.

Example I.4 (General linear groups)**.**
Let $\mathbb{K}$ be a field and let $n \in \mathbb{Z}_{>0}$. The set
$$\mathrm{GL}_n(\mathbb{K}) := \left\{ M \in \mathbb{K}^{n \times n} \;\middle|\; \det(M) \neq 0 \right\} \tag{I.1}$$
of invertible $n \times n$ matrices with entries in $\mathbb{K}$ is a group under the binary operation of matrix multiplication. It is called the **general linear group** in dimension n over the field $\mathbb{K}$.

Example I.5 (Automorphism groups of vector spaces)**.**
Let V be a vector space over a field $\mathbb{K}$ and let
$$\mathrm{Aut}(V) = \{A \colon V \to V \text{ linear bijection}\} \tag{I.2}$$
with composition of functions as the binary operation. Then $\mathrm{Aut}(V)$ is a group, called the **automorphism group of the vector space** V.

When V is a finite dimensional vector space over $\mathbb{K}$ of dimension $\dim_{\mathbb{K}}(V) = n$, and a basis of V has been chosen, then V can be identified with the vector space $\mathbb{K}^n$, and $\mathrm{Aut}(V)$ can be identified with the group $\mathrm{GL}_n(\mathbb{K})$ of invertible $n \times n$ matrices described in Example I.4. Therefore we also sometimes call $\mathrm{Aut}(V)$ the **general linear group** of V, and occasionally denote it by $\mathrm{GL}(V)$.

Example I.6 (Dihedral groups)**.**
Consider a regular polygon with n sides: a triangle, square, pentagon, hexagon, ... — generally called an n-gon.

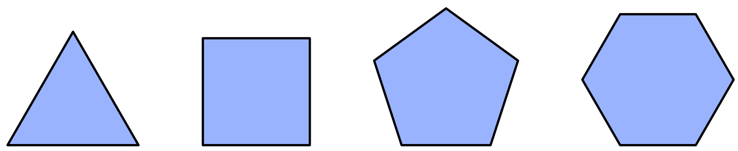


FIGURE I.1. Regular polygons.

The group D_n of symmetries of the polygon, or the **dihedral group** of order $2n$, is the group with two generators
$$r: \text{ “rotation by } 2\pi/n\text{”}, \qquad m: \text{ “reflection”}$$
and relations
$$r^n = e, \qquad m^2 = e, \qquad rmrm = e.$$

The following group has the interpretation the group of rotations in the Euclidean space $\mathbb{R}^3$.

Exercise I.7 (The special orthogonal group in three dimensions).
Show that the set

$$\mathrm{SO}_3 = \big\{ M \in \mathbb{R}^{3\times 3} \;\big|\; M^\top M = \mathbb{I}_3,\ \det(M) = 1 \big\}$$

of orthogonal matrices with determinant one is a group, with matrix multiplication as the group operation.

Regular polyhedra, in particular all the Platonic solids (such as the dodecahedron of Figure O.1) have symmetry groups. The following describes the rotational symmetries of one of them — the cube. Indeed the set F in the following exercise can be interpreted as the set of the six faces of a cube centered at the origin.

Exercise I.8 (Orientation preserving symmetries of the cube).
Let u_1, u_2, u_3 denote the standard basis vectors of $\mathbb{R}^3$. Show the subset $G \subset \mathrm{SO}_3$ consisting of those $M \in \mathrm{SO}_3$ which map the set $F = \{u_1, -u_1, u_2, -u_2, u_3, -u_3\}$ to itself is a finite subgroup of SO_3 of order $|G| = 24$.

1.3. Group homomorphisms

Maps between groups which respect the structure given by the binary operations are called group homomorphisms, or just homomorphisms, when the context is clear.

Definition I.9 (Group homomorphism).
Let $(G_1, *_1)$ and $(G_2, *_2)$ be groups. A mapping $f\colon G_1 \to G_2$ is said to be a **homomorphism** if for all $g, h \in G_1$

$$f\big(g *_1 h\big) \;=\; f(g) *_2 f(h). \tag{I.3}$$

Example I.10 (Determinant is a homomorphism).
The determinant function $A \mapsto \det(A)$ from the group $\mathrm{GL}_n(\mathbb{C})$ of invertible $n \times n$ complex matrices to the multiplicative group of non-zero complex numbers, is a homomorphism since

$$\det(A\,B) = \det(A)\,\det(B).$$

The reader should be familiar with the notions of subgroup, normal subgroup, quotient group, canonical projection, kernel, isomorphism etc.

One of the most fundamental recurrent principles in mathematics is isomorphism theorems. In the case of groups, the statement is the following.

Theorem I.11 (Isomorphism theorem for groups).
Let G and 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 group* $G/\mathrm{Ker}\,(f)$ *is isomorphic to* $\mathrm{Im}\,(f)$.

Property 3° above is more explicitly stated as follows. There exists a unique group homomorphism

$$\bar{f}\colon G/\mathrm{Ker}\,(f) \to H$$

such that the following diagram commutes

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

where $\pi\colon G \to G/\mathrm{Ker}\,(f)$ *is the canonical projection. The homomorphism* $\bar{f}$ *is injective,* $\mathrm{Ker}\,(\bar{f}) = \{e_{G/\mathrm{Ker}(f)}\}$*, and its image coincides with that of* f*,* $\mathrm{Im}\,(\bar{f}) = \mathrm{Im}\,(f) \subset H$*, so* $\bar{f}$ *provides an isomorphism* $G/\mathrm{Ker}\,(f) \cong \mathrm{Im}\,(f)$*.*

You have probably encountered isomorphism theorems for several algebraic structures already — the following table summarizes the corresponding concepts in a few familiar cases

Structure	**Morphism** f	**Image** $\mathrm{Im}\,(f)$	**Kernel** $\mathrm{Ker}\,(f)$
group	group homomorphism	subgroup	normal subgroup
vector space	linear map	vector subspace	vector subspace
ring	ring homomorphism	subring	ideal
$\vdots$	$\vdots$	$\vdots$	$\vdots$

We will encounter isomorphism theorems for yet other algebraic structures during this course: representations (modules), Lie algebras, and Lie groups[2] in particular. The idea is always the same, and the proofs are mostly very similar.

1.4. Conjugacy and conjugacy classes

The notion of conjugacy will be important in representation theory.

Definition I.12 (Conjugacy)**.**

Let G be a group. Two elements $g_1, g_2 \in G$ are said to be **conjugates** if there exists an element $h \in G$ such that $g_2 = h\, g_1\, h^{-1}$. Being conjugate is an equivalence relation in G, and the equivalence classes are called **conjugacy classes** of G.

Exercise I.13 (Conjugacy classes in the symmetric group on three letters)**.**

Find the conjugacy classes of the symmetric group $\mathfrak{S}_3$ on three letters.

Hint: Recall that if $g_1, g_2 \in G$ are conjugate elements, then, e.g., their orders are equal.

Exercise I.14 (Conjugacy classes in the group of symmetries of the cube)**.**

Find the conjugacy classes of the group of orientation preserving symmetries of a cube (see Exercise I.8).

Hint: Note that if $M_1, M_2 \in \mathbb{C}^{n\times n}$ are conjugate matrices, then, e.g., their eigenvalues coincide.

[2]Lie groups are groups, but with the extra structure allowing us to do calculus on them, so homomorphisms must also preserve this additional structure.

1.5. Group actions

Definition I.15 (Action of a group).

Let G be a group and X a set. An **action** of G on X is a group homomorphism

$$\alpha\colon G \to \mathfrak{S}(X).$$

In other words, if α is an action of G on X, then any group element $g \in G$ acts by a bijection $\alpha(g)\colon X \to X$ of X, and the compositions of these bijections respect the product in the group (by the homomorphism requirement).

Example I.16 (The defining action of the symmetric group).

The symmetric group $\mathfrak{S}_n$ consists of bijective maps of $\{1, \ldots, n\}$ to itself, so tautologically (by setting $\alpha(\sigma) := \sigma$ for all $\sigma \in \mathfrak{S}_n$) it we get an action of $\mathfrak{S}_n$ on $\{1, \ldots, n\}$.

Example I.17 (Action of a group on itself by left multiplication).

Let G be any group. Then we can define an action α of G on itself, which we for clarity denote by $\alpha(g) =: \alpha_g$, by

$$\alpha_g(h) = gh$$

for all $g \in G$ and $h \in G$.[3] The homomorphism property of α is a consequence of associativity:

$$\alpha_{g_1 g_2}(h) = g_1 g_2 h = \alpha_{g_1}(g_2 h) = \alpha_{g_1}\big(\alpha_{g_2}(h)\big) = (\alpha_{g_1} \circ \alpha_{g_2})(h).$$

Exercise I.18 (Action of a group on itself by right multiplication).

Let G be any group and for $g \in G$ define $\tilde{\alpha}_g\colon G \to G$ by the formula

$$\tilde{\alpha}_g(h) = hg^{-1.}$$

Show that $g \mapsto \tilde{\alpha}_g$ is an action of the group G on itself.

Exercise I.19 (Action of a group on itself by conjugation).

Let G be any group and for $g \in G$ define $\gamma_g\colon G \to G$ by the formula

$$\gamma_g(h) = ghg^{-1.}$$

Show that $g \mapsto \gamma_g$ is an action of the group G on itself.

It is also not difficult to envision the ways in which the (abstract) symmetry group of a regular polyhedron (such as the dodecahedron of Figure O.1) acts on the set of vertices, on the set of edges, or on the set of faces of the polyhedron. Indeed, in Exercise I.8 you have basically seen the action of the group of orientation preserving symmetries of the cube on the set of faces of the cube.

Before the next exercise pertaining to another one of the examples in the introduction, let us introduce a few notions related to group actions.

Suppose that $\alpha\colon G \to \mathfrak{S}(X)$ is an action of a group G on a set X, denoted by $g \mapsto \alpha_g$. Let $x \in X$. The **stabilizer** of x is the subgroup[4]

$$\big\{ g \in G \;\big|\; \alpha_g(x) = x \big\} \subset G.$$

The **orbit** of x is the subset

$$\big\{ y \in X \;\big|\; \exists g \in G :\; \alpha_g(x) = y \big\} \subset X.$$

[3]Note that here g is an element of the group G which acts, whereas h is an element of the set G upon which the group acts.

[4]It is an easy but worthwhile exercise to check that the formula for the stabilizer indeed defines a subgroup.

The action α is said to be **transitive** if for all $x, y \in X$ we have $\alpha_g(x) = y$, i.e., any two points are on the same orbit.

Exercise I.20 (Action of Möbius transformations on the upper half-plane).
Let $\mathbb{H} = \big\{ z \in \mathbb{C} \;\big|\; \mathfrak{Im}(z) > 0 \big\}$ be the upper half-plane, interpreted as a subset of the complex plane $\mathbb{C}$. Let $\mathrm{SL}_2(\mathbb{R})$ be the group of 2×2 matrices with real entries and determinant one.

(a) For any
$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{R}),$$
define a function α_M on the complex plane by $\alpha_M(z) = \frac{az+b}{cz+d}$. Show that $M \mapsto \alpha_M$ defines an action of the group $\mathrm{SL}_2(\mathbb{R})$ on the set $\mathbb{H}$. What is the kernel of the homomorphism $\alpha \colon \mathrm{SL}_2(\mathbb{R}) \to \mathfrak{S}(\mathbb{H})$?

(b) Show that the action α in part (a) is transitive.

(c) Show that the stabilizer of $z_0 = \mathbb{i}$ is the special orthogonal group $\mathrm{SO}_2 \subset \mathrm{SL}_2(\mathbb{R})$, i.e., the group of 2×2 orthogonal matrices with determinant one.

Although we will not elaborate on this in detail, we note that the preceding exercise allows us to realize the upper half plane $\mathbb{H}$ as the "quotient" $\mathrm{SL}_2(\mathbb{R}) \,/\, \mathrm{SO}_2$ (the set of left cosets).

Group actions give a literal meaning to the idea of symmetry transformations — but recall that we plan to focus on the case of linear actions on a vector space, i.e., group representations!

2. Basic theory of representations of finite groups

Before taking on the subject of Lie groups and their representations, we first take a look at the simpler case of finite groups. This lets us introduce key notions in easier concrete examples, and without too much effort we obtain a clear theory which serves as a model for representation theory in more involved contexts.

Our main objective for this part is to prove that there are only finitely many irreducible (complex) representations of a given finite group G, and any finite dimensional (complex) representation of G can be written as a direct sum of copies of these irreducible representations.

The excellent textbooks [**FH91**] and [**Sim96**] both cover the same basics as we do here, and much more.

2.1. Representations: Definition and first examples

Compare the following definition to group actions, Definition I.15.

Definition I.21 (Representation of a group)**.**
Let G be a group and V a vector space. A **representation** of G in V is a group homomorphism
$$\varrho\colon G \to \mathrm{Aut}(V).$$

For any $g \in G$, the image $\varrho(g)$ is an invertible linear map $V \to V$. When the representation ϱ is clear from the context[5], we denote the image of a vector $v \in V$ under this linear map simply by
$$g.v := \varrho(g)\, v \in V.$$
With this notation the requirement that ϱ is a homomorphism reads
$$(gh).v = g.(h.v)$$
for all $g, h \in G$, $v \in V$. It is convenient to interpret this as a left multiplication of vectors $v \in V$ by elements g of the group G. Thus interpreted, we sometimes say that V is a **(left)** G**-module**.

Example I.22 (Trivial representation)**.**
Let V be a vector space over a field $\mathbb{K}$ and set $\varrho(g) = \mathrm{id}_V$ for all $g \in G$. In the module notation this becomes $g.v = v$ for all $g \in G$ and $v \in V$. This is called the **trivial representation** of G in V. If no other vector space V is specified, the trivial representation means the trivial representation in the one dimensional vector space $V = \mathbb{K}$.

Example I.23 (Alternating representations of symmetric groups)**.**
The symmetric group S_n for $n \geq 2$ has another one-dimensional representation called the **alternating representation**. This is the representation given by $\varrho(\sigma) = \mathrm{sgn}(\sigma)\ \mathrm{id}_{\mathbb{K}}$, where $\mathrm{sgn}(\sigma)$ is -1 when the permutation σ is the product of odd number of transpositions, and $+1$ when σ is the product of even number of transpositions.[6]

---

[5]. . . or when we are just too lazy to specify it. . .

[6]Recall that $\mathrm{sgn}(\sigma)$ is well defined by this condition: any permutation can be written as a product of transpositions, and a product of an odd number of transpositions is never equal to a product of an even number of transpositions. Recall, moreover, that $\mathrm{sgn}\colon S_n \to \{-1, +1\}$ thus defined is a group homomorphism.

The previous example is a particular case of the general fact that any group homomorphism to the multiplicative group of invertible scalars gives rise to a one-dimensional representation, and vice versa. In the next exercise you will prove this.

Exercise I.24 (One-dimensional representations).
Let G be a group and $\mathbb{K}$ a field, and denote by $\mathbb{K}^\times = \mathbb{K} \setminus \{0\}$ the multiplicative group of non-zero elements in the field $\mathbb{K}$. Show that there is a one-to-one correspondence between group homomorphisms from G to $\mathbb{K}^\times$, and representations of G in $\mathbb{K}$.
Hint: A homomorphism $f\colon G \to \mathbb{K}^\times$ corresponds to the representation defined by $\varrho(g) = f(g)\,\mathrm{id}_\mathbb{K}$.

Of course one-dimensional representations are only a very particular type of representations in general.

The following allows us to associate to any action of a group G (Definition I.15) a representation of G (Definition I.21). It is one way to construct interesting representations.

Exercise I.25 (Permutation representations).
Suppose that G is a group acting on a set X via a group homomorphism

$$\alpha\colon G \to \mathfrak{S}(X),$$

denoted by $g \mapsto \alpha_g$, so that each $\alpha_g\colon X \to X$ is a bijection and $\alpha_g \circ \alpha_h = \alpha_{gh}$ for all $g, h \in G$. Form the vector space V with basis $(u_x)_{x\in X}$ indexed by the set X. For each $g \in G$, define $\varrho(g)\colon V \to V$ by linear extension of

$$\varrho(g)\, u_x := u_{\alpha_g(x)}$$

from the basis vectors u_x, $x \in X$. Show that ϱ is a representation of G on V.

The representation ϱ constructed in Exercise I.25 is called the **permutation representation** associated with the group action α. In the basis $(u_x)_{x\in X}$, the matrices of $\varrho(g)$ are permutation matrices: each row and each column has exactly one entry equal to 1, and all other entries are zeros. Below are some examples.

Example I.26 (The defining representation of a symmetric group).
The symmetric group $\mathfrak{S}_n$ on n letters naturally acts on the set $\{1, \ldots, n\}$ (see Example I.16). Consider the vector space $V = \mathbb{K}^n$, with standard basis $(u_i)_{i\in\{1,\ldots,n\}}$. Then by Exercise I.25 above, the space $\mathbb{K}^n$ becomes a representation of $\mathfrak{S}_n$ by linear extension of $\varrho(\sigma)u_i = u_{\sigma(i)}$, or in module notation

$$\sigma.u_i = u_{\sigma(i)} \qquad \text{for all } \sigma \in \mathfrak{S}_n,\ i \in \{1, \ldots, n\}\,.$$

This n-dimensional representation of the symmetric group $\mathfrak{S}_n$ on n letters is called the defining representation of the symmetric group.

As specific examples in the case of $n = 3$, if we identify $\mathrm{Aut}(\mathbb{K}^3) \cong \mathrm{GL}_3(\mathbb{K})$ through the choice of basis $(u_i)_{i\in\{1,2,3\}}$, the matrices of the transposition $(23) \in \mathfrak{S}_3$ and the three-cycle $(132) \in \mathfrak{S}_3$ become

$$\varrho\big((23)\big) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \varrho\big((132)\big) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Example I.27 (Regular representation of a group).
Let G be a group. Recall from Example I.17 that G acts on itself by left multiplication, i.e., by $\alpha\colon G \to \mathfrak{S}(G)$ defined by $\alpha_g(h) = gh$ for $g, h \in G$. Let us denote[7] by $\mathbb{K}[G]$ the vector

[7]This (at first strange) choice of notation will become more understandable, when we later realize that $\mathbb{K}[G]$ carries the structure of a $\mathbb{K}$-algebra — it is called the **group algebra** of G.

space with a basis $(u_g)_{g \in G}$ indexed by the elements of G. Then by Exercise I.25 above, the space $\mathbb{K}[G]$ becomes a representation of G by linear extension of $\varrho(g) u_h = u_{\alpha_g(h)} = u_{gh}$, or in module notation

$$g.u_h = u_{gh} \qquad \text{for all } g, h \in G.$$

This is called the **(left) regular representation** of G. If G is a finite group, then $\mathbb{K}[G]$ is a $|G|$-dimensional representation of G.

The following example of a representation should appear very natural.

Example I.28 (The defining representation of a dihedral group)**.**
Let D_3 be the dihedral group of order 6, with generators r, m and relations $r^3 = e$, $m^2 = e$, $rmrm = e$ (see Example I.6). This is the group of symmetries of an equilateral triangle. For concreteness, we can think of the equilateral triangle in the plane $\mathbb{R}^2$ with vertices $A = (1, 0)$, $B = (-1/2, \sqrt{3}/2)$, $C = (-1/2, -\sqrt{3}/2)$.

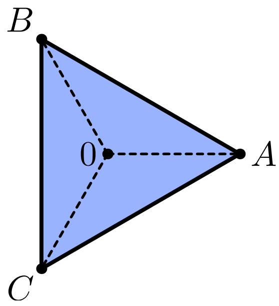


FIGURE I.2. Equilateral triangle centered at the origin.

Both of the matrices

$$R = \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}, \qquad M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

correspond to linear isometries of the plane $\mathbb{R}^2$, which preserve the set $\{A, B, C\}$ of vertices of the triangle (these isometries likewise preserve the set of edges of the triangle, and the set of points in the interior of the triangle, etc.). Since $R^3 = \mathbb{I}$, $M^2 = \mathbb{I}$, $RMRM = \mathbb{I}$, there exists a homomorphism

$$\varrho : D_3 \to \mathrm{GL}_2(\mathbb{R}) \cong \mathrm{Aut}(\mathbb{R}^2)$$

such that $\varrho(r) = R$, $\varrho(m) = M$. Such a homomorphism is unique since we have given the values of it on generators r, m of D_3. This way, it is very natural to represent the group D_3 (and similarly any dihedral group D_n) in a two dimensional vector space!

A representation ϱ is said to be **faithful** if it is injective, i.e., if $\mathrm{Ker}\,(\varrho) = \{e\} \subset G$. For a faithful representation ϱ, the image $\mathrm{Im}\,(\varrho) \subset \mathrm{Aut}(V)$ is isomorphic to the group itself (by Theorem I.11), so a faithful representation gives a concrete realization of the group as a group of linear transformations (or as matrices, if V is finite-dimensional and a basis is chosen). The representation of the symmetry group of the equilateral triangle in the last example is faithful, it could be taken as a defining representation of D_3.

Similarly to the above example, it is natural to define representations of the symmetry groups of regular polyhedra (such as the dodecahedron in Figure O.1) in the three-dimensional space $\mathbb{R}^3$. Indeed, in Exercise I.8 you have basically constructed a three-dimensional representation of the group of orientation preserving symmetries of the cube. There are, however, also other natural representations of such groups, and we return to one just after the next definition.

2.2. Invariant subspaces and subrepresentations

Definition I.29 (Invariant subspace)**.**
Let G be a group and $\varrho\colon G \to \mathrm{Aut}(V)$ a representation of it. An **invariant subspace** of the representation is a vector subspace $W \subset V$ such that for all $g \in G$ we have $\varrho(g)W \subset W$.

Let us now consider a six-dimensional representation of the symmetry group of the cube, and let you find some invariant subspaces yourself.

Exercise I.30 (Functions on the faces of a cube)**.**
As in Exercise I.8, let $F = \{u_1, -u_1, u_2, -u_2, u_3, -u_3\}$ (the positive and negative standard basis vectors in $\mathbb{R}^3$), and let G be the group of orientation preserving symmetries of the cube (a subgroup in the group of orthogonal 3×3 matrices). Let V be the space of complex valued functions on F, i.e.

$$V = \mathbb{C}^F = \{\phi\colon F \to \mathbb{C}\}\,.$$

(a) For a function $\phi\colon F \to \mathbb{C}$ and a group element $g \in G$, define $\varrho(g)\phi\colon F \to \mathbb{C}$ by

$$\big(\varrho(g)\phi\big)(u) = \phi(g^{-1}u) \qquad \text{for all } u \in F$$

(here we consider u as a vector and g^{-1} as a matrix, and $g^{-1}u$ is the multiplication of a vector by a matrix). Show that this defines a representation ϱ of G in $V = \mathbb{C}^F$.

(b) Find at least two examples of nontrivial[8] invariant subspaces $W \subset V$ of the representation in part (a).
Hint: To find invariant subspaces, you may want to take some sufficiently symmetric looking functions $\phi\colon F \to \mathbb{C}$, and see what is the subspace spanned by all $\varrho(g).\phi$ with $g \in G$.

Invariant subspaces naturally inherit representations from the whole space.

Definition I.31 (Subrepresentation)**.**
Suppose that $\varrho\colon G \to \mathrm{Aut}(V)$ is a representation of a group G on a vector space V, and $W \subset V$ is an invariant subspace. For each $g \in G$, the linear map $\varrho(g)\colon V \to V$ can be restricted to the subspace W, and by the invariance of this subspace, the restriction $\varrho(g)\big|_W =: \tilde{\varrho}(g)$ defines a map $\tilde{\varrho}(g)\colon W \to W$. This makes $\tilde{\varrho}\colon G \to \mathrm{Aut}(W)$ a representation[9], and we correspondingly say that W is a **subrepresentation** in V.

---

[8] The trivial cases are the zero subspace $W = \{0\}$ and the entire space $W = V$.

[9] The property $\tilde{\varrho}(gh) = \tilde{\varrho}(g) \circ \tilde{\varrho}(h)$ clearly follows from the corresponding property of ϱ, upon restriction to W. Also we observe

$$\tilde{\varrho}(e) = \varrho(e)\big|_W = (\mathrm{id}_V)\big|_W = \mathrm{id}_W.$$

Thus as a particular case of the property above, we get

$$\tilde{\varrho}(g) \circ \tilde{\varrho}(g^{-1}) = \tilde{\varrho}(gg^{-1}) = \tilde{\varrho}(e) = \mathrm{id}_W$$

and similarly $\tilde{\varrho}(g^{-1}) \circ \tilde{\varrho}(g) = \mathrm{id}_W$. This shows that $\tilde{\varrho}(g)\colon W \to W$ is invertible, so indeed $\tilde{\varrho}$ takes values in $\mathrm{Aut}(W)$.

2.3. Intertwining maps between representations

If V_1, V_2 are two vector spaces, then we denote by

$$\mathrm{Hom}(V_1, V_2) := \{T\colon V_1 \to V_2 \,\text{linear}\} \tag{I.4}$$

the space of linear maps from V_1 to V_2. This is itself a vector space, since linear combinations of linear maps are linear.

If V_1 and V_2 are moreover both representations of the same group G, then it makes sense to ask whether a linear map respects this structure as well.

Definition I.32 (Intertwining map)**.**
Let G be a group, and $\varrho_1\colon G \to \mathrm{Aut}(V_1)$ and $\varrho_2\colon G \to \mathrm{Aut}(V_2)$ two representations of G. A linear map $T\colon V_1 \to V_2$ is said to be an **intertwining map** of representations of G (or a G-module map) if for all $g \in G$ we have

$$\varrho_2(g) \circ T = T \circ \varrho_1(g).$$

We denote the space of such intertwining maps by $\mathrm{Hom}_G(V_1, V_2)$.

Clearly $\mathrm{Hom}_G(V_1, V_2) \subset \mathrm{Hom}(V_1, V_2)$ is a vector subspace.

In the module notation the requirement in Definition I.32 becomes simply[10]

$$T(g.v) = g.T(v) \qquad \text{for all } g \in G,\, v \in V.$$

Definition I.33 (Equivalence of representations)**.**
If an intertwining map $f \in \mathrm{Hom}_G(V_1, V_2)$ is bijective, we call it an **equivalence of representations** (or an **isomorphism of representations**), and we say that the representations V_1 and V_2 are equivalent (or isomorphic).

The basic question of representation theory is:

> Can we classify all representations of a given group G up to equivalence?

When we restrict attention to finite-dimensional complex representations of a finite group G, then we indeed achieve a very satisfactory classification later on in this first part of the course. But before we get there, we need to examine some operations by which we can build new representations out of given ones.

As for any other structures and maps that respect the structure, we have an isomorphism theorem for representations.

Exercise I.34 (Isomorphism theorem for representations)**.**
State and prove the isomorphism theorem for representations of G.
Hint: If $T\colon V_1 \to V_2$ is an intertwining map between representations V_1 and V_2 of G, then what can be said about $\mathrm{Ker}(T) \subset V_1$ and $\mathrm{Im}(T) \subset V_2$?

---

[10]Note, however, that the dot on the left hand side refers to the representation ϱ_1 on V_1, while the dot on the right hand side to ϱ_2 on V_2.

2.4. Operations on representations

Suppose that $\varrho_V : G \to \mathrm{Aut}(V)$ and $\varrho_W : G \to \mathrm{Aut}(W)$ are two representations of G, on vector spaces V and W over a field $\mathbb{K}$. Our goal here is to make sense of the following not only as vector spaces, but as representations of G:

- the direct sum $V \oplus W$ of two representations
- the tensor product $V \otimes W$ of two representations
- the space $\mathrm{Hom}(V, W)$ of linear maps between two representations.

The simplest one is direct sums. Recall first that the **direct sum of vector spaces** V and W over $\mathbb{K}$ is the set

$$\big\{(v, w) \;\big|\; v \in V,\, w \in W\big\} \tag{I.5}$$

of pairs of vectors in the two spaces[11], and this set of pairs equipped with the $\mathbb{K}$-vector space structure where addition and scalar multiplication are performed coordinatewise,

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2), \qquad \lambda(v, w) = (\lambda v, \lambda w)$$

for $v_1, v_2, v \in V$, $w_1, w_2, w \in W$, and $\lambda \in \mathbb{K}$. The vector space thus obtained is denoted $V \oplus W$.

Definition I.35 (Direct sum of representations)**.**
Suppose that $\varrho_V : G \to \mathrm{Aut}(V)$ and $\varrho_W : G \to \mathrm{Aut}(W)$ are two representations of G. For $g \in G$, define $\varrho(g) : V \oplus W \to V \oplus W$ by setting

$$\varrho(g)(v, w) := \big(\varrho_V(g)v,\; \varrho_W(g)w\big)$$

for $v \in V$, $w \in W$. It is easy to see that $\varrho : G \to \mathrm{Aut}(V \oplus W)$ is a representation of G on $V \oplus W$, and we call it the **direct sum of representations** V and W.

In the module notation, the formula in Definition I.35 simply reads[12]

$$g.(v, w) = (g.v, g.w).$$

Let us next consider tensor products. The notion of tensor product $V \otimes W$ of vector spaces V and W is reviewed in Appendix A.1, but let us recall that $V \otimes W$ comes equipped with a bilinear map $V \times W \to V \otimes W$ denoted by

$$(v, w) \mapsto v \otimes w$$

such that a basis of $V \otimes W$ can be obtained in the form $(v_i \otimes w_j)_{i \in I, j \in J}$, if $(v_i)_{i \in I}$ is a basis of V and $(w_j)_{j \in J}$ is a basis of W.

Definition I.36 (Tensor product of representations)**.**
Suppose that $\varrho_V : G \to \mathrm{Aut}(V)$ and $\varrho_W : G \to \mathrm{Aut}(W)$ are two representations of G. For $g \in G$, define $\varrho(g) : V \otimes W \to V \otimes W$ by linear extension of the formula

$$\varrho(g)(v \otimes w) := \big(\varrho_V(g)v\big) \otimes \big(\varrho_W(g)w\big)$$

[11] In other words the set is the Cartesian product $V \times W$, but we avoid this notation in the present context (since it does not correctly generalize to infinite direct sums).

[12] Note that each of the three dots is actually a shorthand for a different representation!

for $v \in V$, $w \in W$. It is easy to see that $\varrho\colon G \to \mathrm{Aut}(V \otimes W)$ is well-defined, and is a representation of G on $V \otimes W$. We call it the **tensor product of representations** V and W.

In the module notation, the formula in Definition I.36 simply reads[13]

$$g.(v \otimes w) = (g.v) \otimes (g.w).$$

The third construction we asked for pertained to spaces of linear maps (I.4). When both the domain and codomain there are not just vector spaces, but in addition representations of the same group, we want to equip the space of linear maps with the structure of a representation of the same group again.

Definition I.37 (The space of linear maps between representations)**.**
Suppose that $\varrho_V\colon G \to \mathrm{Aut}(V)$ and $\varrho_W\colon G \to \mathrm{Aut}(W)$ are two representations of G, and let $\mathrm{Hom}(V, W)$ be the space of linear maps $T\colon V \to W$. For $g \in G$, define $\varrho(g)\colon \mathrm{Hom}(V, W) \to \mathrm{Hom}(V, W)$ by

$$\varrho(g)T := \varrho_W(g) \circ T \circ \varrho_V(g^{-1}).$$

One can straightforwardly check that $\varrho\colon G \to \mathrm{Aut}(\mathrm{Hom}(V, W))$ is a representation of G on $\mathrm{Hom}(V, W)$. We call it the **space of linear maps between representations** V and W (or **Hom-space** of V and W).

In the module notation, Definition I.37 simply becomes[14]

$$(g.T)(v) = g.\big(T(g^{-1}.v)\big) \qquad \text{for any } v \in V.$$

Duals are a special case of spaces of linear maps. For a representation V, the dual V^* consists of all linear maps $V \to \mathbb{K}$, and we make it a representation by equipping the codomain $\mathbb{K}$ with the trivial representation. The example below spells this out in detail.

Example I.38 (Dual representation)**.**
Let $\varrho\colon G \to \mathrm{Aut}(V)$ be a representation of a group G on a vector space V. Consider the dual vector space

$$V^* = \mathrm{Hom}(V, \mathbb{K}).$$

By assumption, V carries a representation of G. We also consider $\mathbb{K}$ as the trivial representation of G: we just have $\varrho_{\mathbb{K}}(g) = \mathrm{id}_{\mathbb{K}}$ for all $g \in G$ (see Example I.22). Therefore Definition I.37 makes also $V^* = \mathrm{Hom}(V, \mathbb{K})$ a representation by introducing a certain

$$\varrho'\colon G \to \mathrm{Aut}(V^*).$$

According to that definition, for $g \in G$ and $\phi \in V^*$, the dual element $\varrho'(g)\phi$ is

$$\varrho'(g)\phi = \varrho_{\mathbb{K}}(g) \circ \phi \circ \varrho(g^{-1}) = \phi \circ \varrho(g^{-1}),$$

where we took into account the triviality of $\varrho_{\mathbb{K}}$.

To describe the dual representation by a concrete formula, let us denote the values of dual elements $\phi \in V^*$ on vectors $v \in V$ by $\langle \phi, v \rangle \in \mathbb{K}$. Then the above becomes

$$\langle \varrho'(g)\phi, v \rangle = \langle \phi, \varrho(g^{-1})v \rangle, \tag{I.6}$$

or in module notation

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

[13] As before — each of the three dots hides a different underlying representation.

[14] And once more — three dots, three different representations.

There is also a way to understand spaces of linear maps via duals and tensor products: according to Exercise A.8, we can identify the vector spaces

$$W \otimes V^* \cong \mathrm{Hom}(V, W)$$

if both V and W are finite-dimensional. But if V and W are representations of G, then both sides of the above identification can be made representations of G as well. A priori, these two representations are defined differently, but you can check that they in fact coincide.

Exercise I.39 (The relation between representations $\mathrm{Hom}(V, W)$ and $W \otimes V^*$).
Suppose that V and W are finite-dimensional representations of G. Show that the linear isomorphism

$$W \otimes V^* \cong \mathrm{Hom}(V, W)$$

is also an equivalence of representations.

2.5. The subspace of invariants

There is yet one more important operation by which we can construct new representations out of old ones, but it is of a slightly different flavor compared to the ones in Section 2.4. Given any representation V of G, we can form the subspace of vectors which transform trivially — the **subspace of invariants**[15] in the original representation.

Definition I.40 (Subspace of invariants).
Let $\varrho \colon G \to \mathrm{Aut}(V)$ be a representation of a group G in a vector space V. Then the subspace $V^G \subset V$ defined by

$$V^G := \Big\{ v \in V \;\Big|\; \varrho(g)v = v \quad \forall g \in G \Big\} \tag{I.7}$$

is called the **subspace of invariants** in the representation V.

In the module notation, Definition I.40 reads

$$v \in V^G \quad \Longleftrightarrow \quad g.v = v \quad \forall g \in G \tag{I.8}$$

Thus indeed a vector $v \in V^G$ in the subspace of invariants (an invariant vector) is not changed in any way by the action of any group element. It follows that the one-dimensional subspace $\mathbb{K}v \subset V$ spanned by any non-zero invariant vector $v \in V^G \setminus \{0\}$ is a subrepresentation in V, which is a trivial representation in the sense of Example I.22. Also the subspace $V^G \subset V$ as a whole is a subrepresentation which is a trivial representation, but its dimension could be anything from zero (if no non-zero invariant vectors exist in V) to something very large (depending on the representation V).

Exercise I.41 (Invariants in the defining representation of a symmetric group).
Consider the symmetric group $\mathfrak{S}_n$ on n letters and its defining n-dimensional representation $\mathbb{K}^n$ (see Example I.26). What is the subspace of invariants $(\mathbb{K}^n)^{\mathfrak{S}_n} \subset \mathbb{K}^n$?

[15] To avoid any confusion from similarities in terminology, let us point out that the subspace of invariants (Definition I.40) can be easily seen to be an invariant subspace (Definition I.29), but not every invariant subspace is a subspace of invariants.

The subspace of invariants inside the space $\mathrm{Hom}(V,W)$ of linear maps between representations V and W (Definition I.37) is of particular interest due to the following observation.

Proposition I.42 (Intertwining maps are the invariants in the space of linear maps)**.**
Let $\varrho_V : G \to \mathrm{Aut}(V)$ and $\varrho_W : G \to \mathrm{Aut}(W)$ be two representations of G. Then we have

$$\mathrm{Hom}_G(V,W) = \mathrm{Hom}(V,W)^G. \tag{I.9}$$

Proof. Recall that the intertwining maps at least form a vector subspace $\mathrm{Hom}_G(V,W) \subset \mathrm{Hom}(V,W)$ in the space of all linear maps. We prove the equality (I.9) by proving inclusions in both directions.

Let us first show $\mathrm{Hom}_G(V,W) \subset \mathrm{Hom}(V,W)^G$. So assume that $T \in \mathrm{Hom}_G(V,W)$ is an intertwining map. Then calculate (in the module notation), for any $g \in G$ and $v \in V$

$$(g.T)(v) = g.T(g^{-1}.v) \overset{\star}{=} T\big(g.(g^{-1}v)\big) = T(e.v) = T(v),$$

where at $\star$ we used the intertwining property of T. This shows that $g.T = T$, that is $T \in \mathrm{Hom}(V,W)^G$, so we have proven the first inclusion.

It remains to show that $\mathrm{Hom}_G(V,W) \supset \mathrm{Hom}(V,W)^G$. So assume that $T \in \mathrm{Hom}(V,W)^G$, i.e, $T : V \to W$ is a linear map such that $g.T = T$ for all $g \in G$. By definition, $g.T$ is the linear map $V \to W$, whose value at $v' \in V$ is given by

$$(g.T)(v') = g.T(g^{-1}.v').$$

Therefore the required equality $g.T = T$ explicitly reads

$$g.T(g^{-1}.v') = T(v')$$

for any $v' \in V$. Now for any given $v \in V$ we can consider $v' = g.v \in V$, and from the above we get

$$g.T(v) = T(g.v).$$

This shows that T is intertwining, and proves the other inclusion. □

2.6. Irreducibile representations

Recall that the subrepresentations (Definition I.31) of a given representation correspond to invariant subspaces (Definition I.29).

Example I.43 (Two obvious subrepresentations)**.**
Let $\varrho : G \to \mathrm{Aut}(V)$ be a representation of a group G on a vector space V. The subspaces

$$\{0\} \subset V \qquad \text{and} \qquad V \subset V$$

are obviously invariant. The corresponding subrepresentations are not particularly enlightening.

If a representation has no other subrepresentations except these obvious ones, then it can not be broken down to simpler pieces in any meaningful way. Therefore we define:

Definition I.44 (Irreducible representation)**.**
A representation $\varrho : G \to \mathrm{Aut}(V)$ of a group G on a vector space $V \neq \{0\}$ is called **irreducible** if its only invariant subspaces are $\{0\} \subset V$ and $V \subset V$.

There are, of course, situations in which a representation can be broken down to simpler pieces.

Example I.45 (Subrepresentations in a direct sum)**.**
Let $\varrho_1 \colon G \to \mathrm{Aut}(V_1)$ and $\varrho_2 \colon G \to \mathrm{Aut}(V_2)$ be two representations of a group G on vector spaces V_1 and V_2. Consider the direct sum (Definition I.35)

$$V_1 \oplus V_2 = \Big\{ (v_1, v_2) \;\Big|\; v_1 \in V_1,\, v_2 \in V_2 \Big\}.$$

The subspaces

$$\begin{aligned} \big\{ (v_1, 0) \;\big|\; v_1 \in V_1 \big\} \;&\subset\; V_1 \oplus V_2 \\ \big\{ (0, v_2) \;\big|\; v_2 \in V_2 \big\} \;&\subset\; V_1 \oplus V_2 \end{aligned}$$

are invariant. The corresponding subrepresentations of $V_1 \oplus V_2$ are isomorphic to V_1 and V_2, respectively.

2.7. Complementary subrepresentations

The previous example turns out to be the general case[16]: any time we find a subrepresentation, there is also another complementary subrepresentation such that the whole is the direct sum of the two parts.

To make this precise, recall first the corresponding idea for vector subspaces. If V is a vector space and $W \subset V$ is a subspace, then another subspace $W' \subset V$ is said to be **complementary** to W if any vector $v \in V$ can be uniquely written as

$$v = w + w' \qquad \text{with } w \in W,\ w' \in W'.$$

The existence of such decompositions says that W and W' together span V, while the uniqueness amounts to $W \cap W' = \{0\}$. In such a situation, the map

$$\begin{aligned} W \oplus W' &\to V \\ (w, w') &\mapsto w + w' \end{aligned}$$

is a linear isomorphism, so we write

$$V = W \oplus W',$$

and say that the vector space V is the direct sum of the subspaces W and W'. We can then also define a projection $\pi \colon V \to W$ associated to the direct sum decomposition $V = W \oplus W'$, by defining $\pi(w + w') = w$ when $w \in W$ and $w' \in W'$. Clearly this projection satisfies $\pi|_W = \mathrm{id}_W$, $\mathrm{Im}(\pi) = W$, and $\mathrm{Ker}(\pi) = W'$. Note, however, that given the original subspace $W \subset V$, the choice of a complementary subspace W' (and the corresponding projection) is far from being unique!

Now often the same thing can be done with representations, and the choice of the complementary subrepresentation is in fact unique!

We assume that G is a finite group, and $\varrho \colon G \to \mathrm{Aut}(V)$ is a representation of G in a vector space V over a field $\mathbb{K}$ such that:

$$\text{The characteristic of } \mathbb{K} \text{ does not divide the order of } G. \qquad (\mathrm{char}(\mathbb{K}) \nmid |G|)$$

16. . . . for representations of finite groups in vector spaces over a field whose characteristic does not divide the order of the group. . .

The assumption on the characteristic of the field is needed so that $|G| \in \mathbb{K}$ is a non-zero element, and it has an inverse $\frac{1}{|G|} \in \mathbb{K}$. In particular fields of characteristic zero such as $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ always satisfy this assumption.

Proposition I.46 (Complementary subrepresentation).
Let G be a finite group, and $\varrho \colon G \to \mathrm{Aut}(V)$ a representation of G in a vector space V over a field $\mathbb{K}$ such that $(\mathrm{char}(\mathbb{K}) \nmid |G|)$ holds.

Then for any given subrepresentation $W \subset V$, there exists another subrepresentation $U \subset V$ such that $V \cong W \oplus U$ as representations of G.

Proof. First choose any complementary vector subspace W' for W, that is $W' \subset V$ such that $V = W \oplus W'$ as a vector space. Let $\pi' \colon V \to W$ be the canonical projection corresponding to this direct sum, that is

$$\pi'(w + w') \;=\; w \qquad \text{when } w \in W,\ w' \in W'.$$

Define

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

Observe that $\mathrm{Im}\,(\pi) \subset W$, since $\mathrm{Im}\,(\pi') \subset W$ and W is an invariant subspace. Observe also that $\pi|_W = \mathrm{id}_W$, since W is an invariant subspace and $\pi'|_W = \mathrm{id}_W$. Together these observations imply that π is a projection from V to W. If we set $U = \mathrm{Ker}\,(\pi)$, then at least $V = W \oplus U$ as a vector space. To show that U is a subrepresentation, it suffices to show that π is an intertwining map. This is checked by doing the change of summation variable $\tilde{g} = hg$ in the following

$$\begin{aligned} h.\pi(v) \;=\; \frac{1}{|G|} \sum_{g \in G} hg.\pi'(g^{-1}.v) &= \frac{1}{|G|} \sum_{g \in G} hg.\pi'\big(g^{-1}h^{-1}h.v\big) \\ &= \frac{1}{|G|} \sum_{\tilde{g} \in G} \tilde{g}.\pi'\big(\tilde{g}^{-1}h.v\big) \;=\; \pi(h.v). \end{aligned}$$

We conclude that $U = \mathrm{Ker}\,(\pi) \subset V$ is a subrepresentation and thus $V = W \oplus U$ as a representation. □

Exercise I.47 (Complement of invariants in the defining representation of a symmetric group).
Consider the symmetric group $\mathfrak{S}_n$ on n letters and its defining n-dimensional representation $\mathbb{K}^n$ (see Example I.26) over a field $\mathbb{K}$ of characteristic zero. Recall that there is a non-zero subspace of invariants $(\mathbb{K}^n)^{\mathfrak{S}_n} \subset \mathbb{K}^n$ (see Exercise I.41). Find a complementary subrepresentation $U \subset \mathbb{K}^n$ to $(\mathbb{K}^n)^{\mathfrak{S}_n}$.

Exercise I.48 (Complementary subrepresentations on functions on the faces of a cube).
In Exercise I.30 you found some subrepresentations of the space $\mathbb{C}^F$ of functions on the set F of faces of a cube, seen as a representation of the group of orientation preserving symmetries of the cube (Exercise I.8). Find complementary subrepresentations to each of these.

2.8. Complete reducibility

After decomposing a representation into a direct sum of two complementary subrepresentations, one can continue and try to decompose each piece further. By induction on dimension one gets the following — still under the same assumption about the characteristic of the ground field.

Corollary I.49 (Complete reducibility).
Let G be a finite group, and $\varrho\colon G \to \mathrm{Aut}(V)$ a representation of G in a finite-dimensional vector space V over a field $\mathbb{K}$ such that $(\mathrm{char}(\mathbb{K}) \nmid |G|)$ *holds.*

Then, as representations, we have

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_n,$$

where each subrepresentation $V_j \subset V$, $j = 1, 2, \ldots, n$, is an irreducible representation of G.

2.9. Schur's lemmas

We now state three related results which say that there is not much freedom in constructing intertwining maps between irreducible representations. These are known as Schur's lemmas, and they are a fundamentally important basic tool in representation theory. The general version is the following.

Lemma I.50 (Schur's lemma).
If V and W are irreducible representations of a group G, and $T\colon V \to W$ is an intertwining map, then either $T = 0$ or T is an isomorphism.

Proof. If we have a nontrivial kernel $\mathrm{Ker}\,(T) \neq \{0\}$, then by irreducibility of V, the subrepresentation $\mathrm{Ker}\,(T) \subset V$ has to be the whole space, $\mathrm{Ker}\,(T) = V$. In this case we have $T = 0$.

If, on the other hand, $\mathrm{Ker}\,(T) = \{0\}$, then T is injective and by irreducibility of W, the non-zero subrepresentation $\mathrm{Im}\,(T) \subset W$ has to be the whole space, $\mathrm{Im}\,(T) = W$. Therefore T is also surjective, and thus an isomorphism. □

From the general statement of Lemma I.50 we know that no non-zero intertwining maps exist between non-isomorphic irreducible representations. The remaining question therefore is about the intertwining maps between isomorphic irreducibles — or equivalently from one irreducible to itself.

For the remaining two formulations of Schur's lemma, we make an assumption about the ground field $\mathbb{K}$ — we assume that the field $\mathbb{K}$ is **algebraically closed**:

$$\text{Every non-constant polynomial } p(x) \in \mathbb{K}[x] \text{ has a root.} \tag{AlgClos}$$

By far the most commonly used algebraically closed field is the field $\mathbb{K} = \mathbb{C}$ of complex numbers.[17] The formulations moreover assume the representations to be finite-dimensional. To undersdand the role of these assumptions, note the following implication. For a square matrix $A \in \mathbb{K}^{n\times n}$ over an algebraically closed field $\mathbb{K}$, the characteristic polynomial $p_A(x) := \det(x\,\mathbb{I} - A)$ of A has a root, and therefore A has at least one eigenvalue $\lambda \in \mathbb{K}$.

Lemma I.51 (Schur's lemma over albegraically closed fields).
Let G be a group, and $\varrho\colon G \to \mathrm{Aut}(V)$ an irreducible representation of G in a finite-dimensional vector space V over a field $\mathbb{K}$ such that (AlgClos) *holds.*

Then any intertwining map $T\colon V \to V$ is necessarily of the form $T = \lambda\,\mathrm{id}_V$ for some $\lambda \in \mathbb{K}$.

[17]The Fundamental theorem of algebra states that every non-constant polynomial with complex coefficients has a root in the complex plane. A proof of this fact is usually provided in courses of complex analysis or algebra.

Proof. Since $\mathbb{K}$ is algebraically closed, there exists an eigenvalue $\lambda \in \mathbb{K}$ of T. This implies that $\mathrm{Ker}(T - \lambda\,\mathrm{id}_V) \neq \{0\}$, and by irreducibility of V, the subrepresentation $\mathrm{Ker}(T - \lambda\,\mathrm{id}_V) \subset V$ has to be the whole space, $\mathrm{Ker}(T - \lambda\,\mathrm{id}_V) = V$. In other words, we have $T - \lambda\,\mathrm{id}_V = 0$, and the assertion follows. □

Corollary I.52 (Schur's lemma for dimension of intertwining maps).
Let G be a group, and $\varrho_V : G \to \mathrm{Aut}(V)$ and $\varrho_W : G \to \mathrm{Aut}(W)$ two irreducible representations of G in finite-dimensional vector spaces V and W over a field $\mathbb{K}$ such that (AlgClos) *holds.*

Then the dimension of the space of intertwining maps between these irreducible representations is given by

$$\dim\Big(\mathrm{Hom}_G(V, W)\Big) = \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{if } V \not\cong W. \end{cases}$$

Proof. The case $V \not\cong W$ is a direct consequence of Lemma I.50, and the case $V \cong W$ follows from Lemma I.51. □

Exercise I.53 (Irreducible representations of abelian groups).
Let G be an abelian (=commutative) group. Show that any irreducible representation of G is one-dimensional. Conclude that (isomorphism classes of) irreducible representations can be identified with group homomorphisms $G \to \mathbb{C}^*$.

Exercise I.54 (Irreducible representations of finite cyclic groups).
Let $C_n \cong \mathbb{Z}/n\mathbb{Z}$ be the cyclic group of order n, i.e. the group with one generator c and relation $c^n = e$. Find all irreducible representations of C_n.

3. Character theory for representations of finite groups

Throughout this Section 3, we make the following assumptions:

- The ground field is $\mathbb{C}$.
- The group G is finite.
- We consider only finite-dimensional representations of G.

In particular the assumptions ($\mathrm{char}(\mathbb{K}) \nmid |G|$) and (AlgClos) about the ground field that were occasionally used, are valid. The above assumptions are also sufficient for complete reduciblity (Corollary I.49), and all formulations of Schur's lemmas (Lemma I.50, Lemma I.51, and Corollary I.52).

A very practical and powerful tool for representation theory in this setup is characters of representations, which will be the topic of this section.

3.1. Characters: definition and first examples

Definition I.55 (Character).
The **character** of a representation $\varrho : G \to \mathrm{Aut}(V)$ is the function $\chi_V : G \to \mathbb{C}$ defined by

$$\chi_V(g) \;=\; \mathrm{tr}\big(\varrho(g)\big) \qquad \text{for } g \in G.$$

Observe that for the neutral element $e \in G$ one necessarily has $\varrho(e) = \mathrm{id}_V$, and since $\mathrm{tr}(\mathrm{id}_V) = \dim(V)$, this implies

$$\chi_V(e) \;=\; \dim(V)\,. \tag{I.10}$$

Therefore the character contains the information about the dimension of the representation, and can indeed be seen as a rather natural generalization of just the dimension.

Observe also that if two group elements $g_1, g_2 \in G$ are conjugates, $g_2 = h g_1 h^{-1}$ for some $h \in G$, then the cyclicity of trace[18] gives

$$\begin{aligned}\chi_V(g_2) \;=\; \mathrm{tr}\big(\varrho(g_2)\big) &= \mathrm{tr}\left(\varrho(h)\,\varrho(g_1)\,\varrho(h)^{-1}\right)\\ &= \mathrm{tr}\left(\varrho(g_1)\,\varrho(h)^{-1}\,\varrho(h)\right) = \mathrm{tr}\big(\varrho(g_1)\big) \;=\; \chi_V(g_1).\end{aligned}$$

Therefore the value of a character is constant on each conjugacy class of G. We call any function $G \to \mathbb{C}$ which is constant on each conjugacy class a **class function**, and a major goal for this section is to prove that the characters of the irreducible representations of G form an orthonormal basis of the space of class functions, when this space is equipped with a suitable inner product.

Example I.56 (Character of the defining representation of D_3).
Consider the dihedral group D_3, and its two-dimensional representation introduced in Example I.28. We can obviously view it also as a representation defined over $\mathbb{C}$, so that the vector space is $V = \mathbb{C}^2$, and the representation

$$\varrho\colon D_3 \to \mathrm{Aut}(V) \cong \mathrm{GL}_2(\mathbb{C})$$

[18] The cyclicity of trace is the property $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for any $A, B \in \mathbb{C}^{n\times n}$.

is determined by the values on the generators $r, m \in D_3$.

$$\varrho(r) = \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}, \qquad \varrho(m) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The group D_3 consists of six elements,

$$D_3 = \{e, r, r^2, m, mr, mr^2\},$$

and it is straightforward to check that the conjugacy classes in D_3 are

$$\{e\}, \qquad \{r, r^2\}, \qquad \{m, mr, mr^2\}.$$

Since the character values are the same within each conjugacy class, it suffices to calculate $\chi_V(g)$ for just one g in each of these. From the explicit matrices given above, we find

$$\begin{aligned} \chi_V(e) &= \mathrm{tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 2, \\ \chi_V(r) &= \mathrm{tr}\left(\begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}\right) = -1, \\ \chi_V(m) &= \mathrm{tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right) = 0. \end{aligned}$$

Let us furthermore observe the following.

Lemma I.57 (Diagonalizability of representation matrices)**.**
Let G be a finite group of order $n := |G|$, and let $\varrho \colon G \to \mathrm{Aut}(V)$ be a representation of G. Then for any $g \in G$, the linear map $\varrho(g) \colon V \to V$ is diagonalizable, and all its eigenvalues λ are roots of unity of order n, i.e., they satisfy $\lambda^n = 1$.

Proof. Fix $g \in G$. The order m of g divides the order n of the group G. From $g^m = e$ we get

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

Therefore the minimal polynomial of $\varrho(g)$ divides the polynomial $x^m - 1$. The roots of $x^m - 1$ are simple, so $\varrho(g)$ is diagonalizable (it has no Jordan blocks of size greater than one). Also all the eigenvalues λ of $\varrho(g)$ then satisfy $\lambda^m = 1$, and since $m \mid n$, this implies in particular $\lambda^n = 1$. □

In view of Lemma I.57, we see that the character value $\chi_V(g)$ is also the sum of eigenvalues of $\varrho(g) \colon V \to V$, counted with multiplicity.

Example I.58 (Character of the defining representation of D_3 revisited)**.**
In Example I.56 we calculated the character of the defining representation of D_3. We can also see that the rotation r is represented by the matrix

$$\varrho(r) = \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix},$$

whose (complex) eigenvalues are $e^{\mathbb{i}2\pi/3}$ and $e^{-\mathbb{i}2\pi/3}$. We thus recover

$$\chi_V(r) = e^{\mathbb{i}2\pi/3} + e^{-\mathbb{i}2\pi/3} = 2\cos(2\pi/3) = -1.$$

Similarly the reflection $m \in D_3$ is represented by a matrix with two eigenvalues, $+1$ and -1, so $\varrho_V(m) = +1 - 1 = 0$.

All of the eigenvalues here were, of course, sixth roots of unity, in accordance with $|D_3| = 6$.

The characters of permutation representations are generally described in the following.

Example I.59 (Character of a permutation representation).
Let $\alpha\colon G \to \mathfrak{S}(X)$ be an action of a group G on a set X, denoted by $g \mapsto \alpha_g$. By Exercise I.25, the vector space V with basis $(u_x)_{x\in X}$ has a **permutation representation** such that

$$\varrho(g)u_x = u_{\alpha_g(x)} \qquad \text{for all } g \in G,\ x \in X.$$

The matrix of $\varrho(g)$ in this basis has only zeroes and ones as its entries, and in particular the diagonal entry corresponding to the index $x \in X$ equals one if $\alpha_g(x) = x$ and zero otherwise. The trace $\mathrm{tr}(\varrho(g))$ therefore becomes the number of those $x \in X$ which are fixed by the action of $g \in G$, i.e., we have

$$\chi_V(g) = \#\left\{x \in X \;\middle|\; \alpha_g(x) = x\right\}.$$

Example I.60 (Character of the regular representation).
The regular representation $\mathbb{C}[G]$ of a group G is a particular case of a permutation representation, see Example I.27. It is associated with the action $\alpha_g(h) = gh$ of the group on itself by left multiplication. For a group element $g \neq e$, there exists no $h \in G$ such that $gh = h$ (right multiply by h^{-1} to see this), so the action of such g has no fixed points. For $g = e$, we obviously have $eh = h$ for all $h \in G$, so all elements are fixed by the action of e. The formula of the previous example therefore gives in particular

$$\chi_{\mathbb{C}[G]}(g) = \begin{cases} |G| & \text{if } g = e \\ 0 & \text{if } g \neq e. \end{cases} \tag{I.11}$$

3.2. Characters of duals, tensor products, and direct sums

Given two representations V and W of G, we have seen how to make the dual V^* a representation, how to make direct sum $V \oplus W$ a representation, and how to make the tensor product $V \otimes W$ a representation (Example I.38, Definition I.35, and Definition I.36, respectively).

Let us now see how these operations affect characters.

Theorem I.61 (How characters are affected by operations).
Let V, W be representations of G. Then for any $g \in G$ we have

$$\begin{aligned} \text{(i)}\ \chi_{V^*}(g) &= \overline{\chi_V(g)} \\ \text{(ii)}\ \chi_{V\oplus W}(g) &= \chi_V(g) + \chi_W(g) \\ \text{(iii)}\ \chi_{V\otimes W}(g) &= \chi_V(g)\,\chi_W(g). \end{aligned}$$

Proof of (i): Fix $g \in G$. By Lemma I.57, we can choose a basis $(v_i)_{i=1}^n$ of eigenvectors of $\varrho_V(g)\colon V \to V$. Denote the corresponding eigenvalues by $(\lambda_i)_{i=1}^n$ so that

$$\varrho_V(g)v_i = \lambda_i\, v_i \qquad \text{for } i = 1, \ldots, n.$$

These eigenvalues can be used to express the character value as

$$\chi_V(g) = \sum_{i=1}^n \lambda_i.$$

We must relate this character value to a value of the character of the dual representation V^*. Let $(\phi_j)_{j=1}^n$ be the dual basis to $(v_i)_{i=1}^n$, i.e., the basis of V^* defined by

$$\langle \phi_j, v_i\rangle = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \qquad \text{for } i, j = 1, \ldots, n.$$

By the definition of the dual representation, the dual element $g.\phi_j \in V^*$ satisfies

$$\langle \varrho_{V^*}(g)\phi_j, v_i\rangle = \langle \phi_j, \underbrace{\varrho_V(g^{-1})v_i}_{=\lambda_i^{-1}v_i}\rangle = \lambda_i^{-1}\,\langle \phi_j, v_i\rangle = \lambda_i^{-1}\,\delta_{i,j} = \lambda_j^{-1}\,\delta_{i,j},$$

so we see that $\varrho_{V^*}(g)\phi_j = \lambda_j^{-1}\phi_j$. But since $|\lambda_j| = 1$ (the eigenvalues are roots of unity), we have $\lambda_j^{-1} = \overline{\lambda}_j$. These are therefore the eigenvalues of $\varrho_{V^*}(g)\colon V^* \to V^*$, and we can express the character value as

$$\chi_{V^*}(g) = \sum_{j=1}^{n} \overline{\lambda}_j = \overline{\chi_V(g)}.$$

Proof of (ii): We leave the details as an exercise to the reader.

Proof of (iii): Fix $g \in G$. Choose a basis $(v_i)_{i=1}^n$ of eigenvectors of $\varrho_V(g)\colon V \to V$, and a basis $(w_j)_{j=1}^m$ of eigenvectors of $\varrho_W(g)\colon W \to W$. Denote the corresponding eigenvalues of the former by $(\lambda_i)_{i=1}^n$ and the latter by $(\mu_j)_{j=1}^m$, so that

$$\begin{aligned} \varrho_V(g)v_i &= \lambda_i\, v_i && \text{for } i = 1, \ldots, n \\ \varrho_W(g)w_j &= \mu_j\, w_j && \text{for } j = 1, \ldots, m. \end{aligned}$$

Now the collection $v_i \otimes w_j$ indexed by pairs (i, j) is a basis of the tensor product $V \otimes W$. Directly from the definition of tensor product representation, we find that

$$\varrho(g)(v_i \otimes w_j) = \big(\varrho_V(g)v_i\big) \otimes \big(\varrho_W(g)w_j\big) = \big(\lambda_i v_i\big) \otimes \big(\mu_j w_j\big) = \lambda_i\mu_j\ (v_i \otimes w_j).$$

Thus we have found the eigenvalues of $\varrho(g)\colon V \otimes W \to V \otimes W$, and we can use them to write the character value as

$$\chi_{V\otimes W}(g) \;=\; \sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_i\mu_j \;=\; \Big(\sum_{i=1}^{n} \lambda_i\Big)\Big(\sum_{j=1}^{m} \mu_j\Big) \;=\; \chi_V(g)\,\chi_W(g).$$

□

The following example application will be important soon.

Example I.62 (Character of the representation of linear maps).
Let V and W be two representations of G, and consider the representation $\mathrm{Hom}(V, W)$ of linear maps from V to W as in Definition I.37. In Exercise I.39 it was shown that this representation is isomorphic to

$$\mathrm{Hom}(V, W) \cong W \otimes V^*.$$

As an application of Theorem I.61, we therefore get that the character of this representation is given by

$$\chi_{\mathrm{Hom}(V,W)}(g) = \chi_W(g)\,\overline{\chi_V(g)}.$$

3.3. Dimension of the subspace of invariants

For V a representation of G, recall from Definition I.40 that

$$V^G \;=\; \big\{ v \in V \;\big|\; g.v = v\ \forall g \in G \big\}$$

is called the subspace of invariants in V. We next describe an easy way to find this subpace by a projection. This simple idea turns out to have some profound consequences in character theory.

Proposition I.63 (Projection to the subspace of invariants).
Let $\varrho : G \to \mathrm{Aut}(V)$ be a representation of G. Define a linear map π on V by the formula

$$\pi(v) \;=\; \frac{1}{|G|} \sum_{g \in G} \varrho(g) v \qquad \text{for } v \in V. \tag{I.12}$$

Then this map π is a projection $V \to V^G$.

Proof. If $v \in V^G$, then $g.v = v$ for all g, so clearly we have $\pi(v) = v$. This shows that $\pi|_{V^G} = \mathrm{id}_{V^G}$. On the other hand, for any $h \in G$ and $v \in V$, use the change of variables $\tilde{g} = hg$ to calculate

$$h.\pi(v) \;=\; \frac{1}{|G|} \sum_{g \in G} hg.v \;=\; \frac{1}{|G|} \sum_{\tilde{g} \in G} \tilde{g}.v \;=\; \pi(v),$$

from which we conclude $\pi(v) \in V^G$. This shows that $\mathrm{Im}\,(\pi) \subset V^G$. Together these two observations show that $\pi \colon V \to V^G$ is a projection. □

Thus we have an explicitly defined projection to the subspace of invariants. In particular we can find the dimension of this subspace with the help of the projection.

Proposition I.64 (Dimension of the subspace of invariants).
Let $\varrho : G \to \mathrm{Aut}(V)$ be a representation of G. Then we have

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

Proof. Let $\pi \colon V \to V^G$ be the map defined by (I.12). Since π is a projection onto the subspace of invariants V^G, its trace is the dimension of this subspace, $\mathrm{tr}(\pi) = \dim(V^G)$. On the other hand, π is explicitly a linear combination of the maps $\varrho(g) \colon V \to V$, so its trace is a linear combination of character values

$$\mathrm{tr}(\pi) \;=\; \frac{1}{|G|} \sum_{g \in G} \chi_V(g).$$

The equality of these two expressions for $\mathrm{tr}(\pi)$ proves the assertion. □

Example I.65 (Dimension of invariants in the space of linear maps).
Let V and W be two representations of G, and consider the representation $\mathrm{Hom}(V, W)$ in Definition I.37. The character of this representation was found in Example I.62 to be

$$\chi_{\mathrm{Hom}(V,W)}(g) = \overline{\chi_V(g)}\, \chi_W(g).$$

By application of Proposition I.64 to this case, we find

$$\dim\Big(\mathrm{Hom}(V, W)^G\Big) = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_V(g)}\, \chi_W(g). \tag{I.13}$$

Recall that functions $G \to \mathbb{C}$ which are constant on each conjugacy class of G are called **class functions**. It is natural[19] to define and inner product on the space of all functions $G \to \mathbb{C}$ by

$$(\phi, \psi) := \frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)}\, \psi(g) \qquad \text{for } \phi, \psi \colon G \to \mathbb{C}. \tag{I.14}$$

[19]For this inner product, the delta-functions at single elements of the group are orthogonal, and the constant function 1 has unit norm. (Moreover, the inner product has a natural invariance under "translations".)

Note that our convention is that the inner product is linear in the second argument and conjugate-linear in the first argument.[20]

Corollary I.66 (Dimension of the space of intertwining maps)**.**
Let V and W be two representations of G. Then the dimension of the space $\mathrm{Hom}_G(V, W)$ of intertwining maps is the inner product of characters

$$\dim\Big(\mathrm{Hom}_G(V, W)\Big) = (\chi_V, \chi_W). \tag{I.15}$$

Proof. Recall from Proposition I.42 that $\mathrm{Hom}_G(V, W) = \mathrm{Hom}(V, W)^G$. The dimension of the right hand side was found in Example I.65. But the formula (I.13) for its dimension is just the inner product (I.14) of χ_V and χ_W. □

3.4. Irreducible characters

Corollary I.66 says something very interesting about the characters of irreducible representations.

Theorem I.67 (Orthonormality of characters of irreducible representations)**.**

(i) *If V and W are irreducible representations of G, then*

$$(\chi_V, \chi_W) \;=\; \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{otherwise.} \end{cases}$$

In particular, the character χ_V of an irreducible representation V is sufficient to determine V up to isomorphism.

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

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

Remark I.68. In fact there is an equality in (iii), the number of irreducible representations of a finite group is precisely the number of its conjugacy classes. This will be proven later.

Proof of Theorem I.67. By Schur's lemma, for irreducible representations V and W, we have

$$\dim\Big(\mathrm{Hom}_G(V, W)\Big) = \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{otherwise.} \end{cases}$$

Assertion (i) therefore follows from Corollary I.66. The linear independence assertion (ii) also follows at once, since orthonormal elements are linearly independent. The dimension of the space of all class functions $G \to \mathbb{C}$ (functions which are constant on each conjugacy class) is the number of conjugacy classes of the group G, so this is an upper bound on the number of linearly independent class functions. Since characters are class functions, assertion (iii) follows from (ii). □

In view of Theorem I.67, the information about the characters of the finite list of (mutually non-isomorphic) irreducible representations of a group G can be naturally collected in a table, whose rows are indexed by these irreducible representations, columns are indexed by the conjugacy classes, and the entries give the value of

[20] This is the convention used by virtually all physicists, and we also prefer that the functional $\psi \mapsto (\phi, \psi)$ defined by ϕ is linear (rather than conjugate-linear).

the character of the representation on the conjugacy class. Such a summary of the irreducible characters is known as the **character table** of the group G.

Example I.69 (Character table of $\mathfrak{S}_3$)**.**
Consider the symmetric group $\mathfrak{S}_3$ on three letters. The conjugacy classes are

$$\underbrace{\{e\}}_{\text{neutral elem.}} \quad , \quad \underbrace{\{(12),(13),(23)\}}_{\text{transpositions}}, \quad \underbrace{\{(123),(132)\}}_{\text{three-cycles}}.$$

We know two one-dimensional representations of $\mathfrak{S}_3$ from Examples I.22 and I.23: the trivial representation $U = \mathbb{C}$ with $\varrho_U(\sigma) = \mathrm{id}_U$ and the alternating representation $U' = \mathbb{C}$ with $\varrho_{U'}(\sigma) = \mathrm{sgn}(\sigma)\,\mathrm{id}_{U'}$. The characters χ_U and $\chi_{U'}$ are of course simply

$$\chi_U(\sigma) = 1 \qquad \text{and} \qquad \chi_{U'}(\sigma) = \mathrm{sgn}(\sigma) \qquad \text{for } \sigma \in \mathfrak{S}_3.$$

By observing that $\mathfrak{S}_3$ is in fact isomorphic to the dihedral group D_3 (every permutation of the vertices A, B, C of an equilateral triangle can be achieved by some symmetry $g \in D_3$ of the triangle), we get from Example I.28 a two-dimensional representation $V = \mathbb{C}^2$ of $\mathfrak{S}_3$. The isomorphism $D_3 \cong \mathfrak{S}_3$ maps the rotation $r \in D_3$ to a three-cycle, and the reflection $m \in D_3$ to a transposition, so the calculations in Example I.56 readily give the character χ_V of V as well. One can also show[21] that V is an irreducible representation of $\mathfrak{S}_3$.

Above we have three irreducible representations U, U', V of $\mathfrak{S}_3$. Since the number of conjugacy classes of $\mathfrak{S}_3$ is also three, there can not be any other irreducible representations besides these. The character values are listed in the following table for one representative in each conjugacy class:

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

The assertions of Theorem I.67 can be seen in this example by direct inspection. The order of the group is $|\mathfrak{S}_3| = 3! = 6$. The conjugacy classes of e, (12), and (123) contain respectively 1, 3, and 2 elements, so for example the pairwise orthogonality of irreducible characters amounts to

$$(\chi_U, \chi_{U'}) = \frac{1}{6}\Big(1 \times 1 \times 1 + 3 \times 1 \times (-1) + 2 \times 1 \times 1\Big) = \frac{1-3+2}{6} = 0$$
$$(\chi_U, \chi_V) = \frac{1}{6}\Big(1 \times 1 \times 2 + 3 \times 1 \times 0 + 2 \times 1 \times (-1)\Big) = \frac{2+0-2}{6} = 0$$
$$(\chi_{U'}, \chi_V) = \frac{1}{6}\Big(1 \times 1 \times 2 + 3 \times (-1) \times 0 + 2 \times 1 \times (-1)\Big) = \frac{2-0-2}{6} = 0$$

and the unit norms property of these characters amounts to

$$(\chi_U, \chi_U) = \frac{1}{6}\Big(1 \times 1^2 + 3 \times 1^2 + 2 \times 1^2\Big) = \frac{1+3+2}{6} = 1$$
$$(\chi_{U'}, \chi_{U'}) = \frac{1}{6}\Big(1 \times 1^2 + 3 \times (-1)^2 + 2 \times 1^2\Big) = \frac{1+3+2}{6} = 1$$
$$(\chi_V, \chi_V) = \frac{1}{6}\Big(1 \times 2^2 + 3 \times 0^2 + 2 \times (-1)^2\Big) = \frac{4+0+2}{6} = 1.$$

Exercise I.70 (Character table of the dihedral group D_4)**.**
Let D_4 be the dihedral group of order 8, generated by r, m subject to relations $r^4 = e$, $m^2 = e$ and $mrm = r^{-1}$.

(a) Show that the conjugacy classes of D_4 are

$$\{e\}, \quad \{r, r^3\}, \quad \{r^2\}, \quad \{m, mr^2\}, \quad \{mr, mr^3\}.$$

[21] The irreducibility of V can either be checked directly, or follows from the calculation of the norm $(\chi_V, \chi_V) = 1$ combined with Theorem I.72(iv) below.

(b) Compute the character of the 2-dimensional representation of D_4 defined by

$$r \mapsto \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \text{and} \qquad m \mapsto \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

and use the result to conclude that this representation is irreducible.

(c) Find four non-isomorphic one-dimensional complex representations of D_4.

(d) Check that the characters of the five irreducible representations given in parts (b) and (c) form an orthonormal basis of the set of class functions on D_4. Write down the character table of D_4.

(e) Let V be the 2-dimensional representation as in part (b). Calculate the character of the tensor product representation $V \otimes V$, and use this to infer what kinds of irreducible subrepresentations $V \otimes V$ has.

In the following exercise we use the fact that the number of irreducible representations equals the number of conjugacy classes.

Exercise I.71 (Another orthogonality relation for irreducible characters)**.**

Let G be a finite group. Denote the set of its conjugacy classes by $G^\sharp$, and the set of its irreducible complex representations by $\widehat{G}$. Using the orthogonality relation for irreducible characters χ_α, $\alpha \in \widehat{G}$,

$$\frac{1}{|G|} \sum_{C \in G^\sharp} |C| \, \overline{\chi_\alpha(C)} \chi_\beta(C) = \delta_{\alpha,\beta} \quad \text{for } \alpha, \beta \in \widehat{G},$$

prove the following: for any $C, D \in G^\sharp$ we have

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

Hint: Use the irreducible characters to define linear maps between the vector spaces $\mathbb{C}^{G^\sharp}$ and $\mathbb{C}^{\widehat{G}}$. Interpret the orthogonality relation as telling how to invert such a map.

3.5. Multiplicities of irreducibles in a representation

Fix a finite group G. Let $(W_\alpha)_{\alpha=1}^s$ denote the (finite) collection of (isomorphism classes of) its irreducible representations. We know that the number s of members of this collection is at most the number of conjugacy classes (and later will see that there is, in fact, an equality).

Let V be a representation of G. By completer reducibility, Corollary I.49, one can write V as a direct sum of irreducible representations. If we order these direct summands suitably, we get

$$V \;\cong\; \underbrace{W_1 \oplus \cdots \oplus W_1}_{m_1 \text{ copies}} \oplus \underbrace{W_2 \oplus \cdots \oplus W_2}_{m_2 \text{ copies}} \oplus \;\cdots\; \oplus \underbrace{W_s \oplus \cdots \oplus W_s}_{m_s \text{ copies}}.$$

It is convenient to abbreviate this as

$$V \;\cong\; \bigoplus_{\alpha=1}^{s} m_\alpha W_\alpha.$$

The number m_α of times that W_α appears as a direct summand here is called the **multiplicity** of W_α in V.

Theorem I.72 (Multiplicities via characters)**.**

In the setup above, we have the following:

(i) *The multiplicities are given by* $m_\alpha = (\chi_{W_\alpha}, \chi_V)$. *Explicitly, this reads*

$$m_\alpha \;=\; \frac{1}{|G|}\sum_{g\in G}\overline{\chi_{W_\alpha}(g)}\,\chi_V(g).$$

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

(iii) *We have* $(\chi_V, \chi_V) = \sum_{\alpha=1}^{s} m_\alpha^2$. *Explicitly, this can be written as*

$$\frac{1}{|G|}\sum_{g\in G}|\chi_V(g)|^2 \;=\; \sum_{\alpha=1}^{s} m_\alpha^2.$$

(iv) *The representation* V *is irreducible if and only if* $(\chi_V, \chi_V) = 1$. *Explicitly, this condition characterizing irreducibility amounts to*

$$\frac{1}{|G|}\sum_{g\in G}|\chi_V(g)|^2 \;=\; 1.$$

Proof. By Proposition I.61, the character of V is given by $\chi_V(g) = \sum_\alpha m_\alpha\,\chi_{W_\alpha}(g)$. Now (i) is obtained by taking in this equality the inner product with χ_{W_α} and using the orthonormality of irreducible characters from Theorem I.67(i). In particular we obtain the (anticipated) fact that in complete reducibility the direct sum decomposition is unique up to permutation of the irreducible summands, and (ii) follows. Also (iii) is a consequence of $\chi_V = \sum_\alpha m_\alpha\,\chi_{W_\alpha}$ and orthonormality. Finally, (iv) is obvious in view of (iii). □

Example I.73 (Multiplicities in the regular representation)**.**
The regular representation $\mathbb{C}[G]$ of a group G was defined in Example I.27, and its character

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

was found in Example I.60. Let $(W_\alpha)_{\alpha=1}^{s}$ denote the collection of irreducible representations of G, and consider the multiplicities $(m_\alpha)_{\alpha=1}^{s}$ of these irreducible representations in the regular representation,

$$\mathbb{C}[G] \;\cong\; \bigoplus_{\alpha=1}^{s} m_\alpha W_\alpha.$$

We can calculate these multiplicities by Theorem I.72(i) and the formula for the character $\chi_{\mathbb{C}[G]}$, and we get

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

In other words, each irreducible representation of G appears in the regular representation $\mathbb{C}[G]$, and its multiplicity equals its dimension, $m_\alpha = \dim(W_\alpha)$.

Corollary I.74 (Sum of squares of dimensions of irreducibles).

For a finite group G*, we have*

$$\sum_\alpha \dim(W_\alpha)^2 = |G|, \tag{I.16}$$

where the sum on the left hand side is over the (isomorphism classes of) irreducible representations of G.

Proof. Consider the dimensions of both sides of the formula $\mathbb{C}[G] \cong \bigoplus_\alpha m_\alpha W_\alpha$, and take into account the conclusion $m_\alpha = \dim(W_\alpha)$ of Example I.73. □

Example I.75 (Dimensions of irreducible representations of $\mathfrak{S}_4$)**.**
Equation (I.16) can give useful and nontrivial information. Consider for example the group $\mathfrak{S}_4$, whose order is $|\mathfrak{S}_4| = 4! = 24$. Since there are five conjugacy classes (identity, transposition, two disjoint transpositions, three-cycle, four-cycle), we know that the number s of irreducible representations of $\mathfrak{S}_4$ is at most five, $s \leq 5$. We have seen the trivial and alternating representations of $\mathfrak{S}_4$. Let us denote these by W_1 and W_2, respectively, and the rest of the irreducible representations by W_α, $\alpha = 3, \ldots, s$. Let us also denote the dimensions briefly $d_\alpha := \dim(W_\alpha)$, and observe that the trivial and alternating representations are one-dimensional, $d_1 = 1$, $d_2 = 1$. The sum of squares formula (I.16) now says

$$1^2 + 1^2 + \sum_{\alpha=3}^{s} d_\alpha^2 = 24,$$

which of course implies

$$\sum_{\alpha=3}^{s} d_\alpha^2 = 22.$$

One thing we learn immediately is that there must exist other irreducible representations besides the trivial W_1 and the alternating W_2, since the sum above is at least non-zero. For the number of irreducibles, this means $s \geq 3$. But 22 is not a prefect square, so there must in fact be more than one other irreducible, i.e., $s \geq 4$. A little more thinking shows that 22 is not even the sum of two squares, so in fact $s \geq 5$, and therefore we must have $s = 5$ and

$$d_3^2 + d_4^2 + d_5^2 = 22.$$

A little examination shows that the only way to write 22 as a sum of three squares is $22 = 2^2 + 3^2 + 3^2$, so we see that the three remaining irreducible representations have dimensions $2, 3, 3$. Up to relabeling the irreducibles, we have proven that the dimensions of the irreducible representations of $\mathfrak{S}_4$ are

$$\dim(W_1) = 1, \quad \dim(W_2) = 1, \quad \dim(W_3) = 2, \quad \dim(W_4) = 3, \quad \dim(W_5) = 3.$$

Exercise I.76 (Character table of $\mathfrak{S}_4$)**.**
The purpose of this exercise is to reconstruct the full character table of $\mathfrak{S}_4$. We already know the trivial and alternating representations, W_1 and W_2, and we know the dimensions of the remaining irreducible representations W_3, W_4, W_5 from Example I.75.

(a) Guided by Exercise I.47, find a three-dimensional subrepresentation inside the four-dimensional defining representation of $\mathfrak{S}_4$, calculate its character, and conclude that it is irreducible. Denote this irreducible representation by W_4.
(b) Calculate the character of $W_2 \otimes W_4$ and conclude that it is irreducible. Denote this irreducible representation by W_5.

It remains to find only one more irreducible representation, W_3 of dimension $\dim(W_3) = 2$. Consider the following two alternative strategies:

(c) Use orthonormality of the irreducible characters to fill in the missing row χ_{W_3} of the character table (orthonormality only determines the missing row up to a complex scalar multiple of absolute value 1, but this factor is determined by the observation that $\chi_{W_3}(e) = \dim(W_3)$ is real and positive).
(c') Calculate the character of $W_4 \otimes W_4$. Find the multiplicities of W_1, W_2, W_4, W_5 in it. Show that $W_4 \otimes W_4$ contains also a non-zero subrepresentation in the complement of the direct sum of copies of W_1, W_2, W_4, W_5. Calculate the character of this non-zero subrepresentation and conclude that it is irreducible. Denote this irreducible representation by W_3.

Write down the full character table of $\mathfrak{S}_4$.

Let G be a finite group. Its **group algebra** $\mathbb{C}[G]$ is the vector space with basis $(u_g)_{g \in G}$ equipped with the product $u_g\, u_h = u_{gh}$ (extended bilinearly).

The following exercise characterizes the center of the group algebra $\mathbb{C}[G]$, and finishes the proof that the number of irreducible representations equals the number of conjugacy classes.

Exercise I.77. Recall that the center of an algebra A is the set $Z \subset A$ of elements that commute with the whole algebra, i.e.

$$Z = \{z \in A \mid za = az \ \ \forall a \in A\}.$$

In the following, consider $A = \mathbb{C}[G]$, the group algebra of a finite group G.

(a) Show that the element

$$a = \sum_{g \in G} \alpha(g)\, u_g \ \in A$$

is in the center of the group algebra if and only if $\alpha(g) = \alpha(hgh^{-1})$ for all $g, h \in G$.

(b) Suppose that $\alpha : G \to \mathbb{C}$ is a function which is constant on each conjugacy class, and suppose furthermore that α is orthogonal (with respect to the inner product $(\psi, \phi) = |G|^{-1} \sum_{g \in G} \overline{\psi(g)}\phi(g)$) to the characters of all irreducible representations of G. Show that for any representation $\varrho : G \to \mathrm{Aut}(V)$ the map $\sum_g \alpha(g)\varrho(g) : V \to V$ is the zero map. Conclude that α has to be zero.

(c) Using (b) and the results from the lectures, show that the number of irreducible representations of the group G is equal to the number of conjugacy classes of G

Part II

Lie groups and their Lie algebras

1. Matrix Lie groups

1.1. Continuous symmetries

We now turn to continuous symmetries. Like any other symmetries, they should be described by groups — but now by groups with a topology so that one can speak of continuity. Moreover, the most useful types of continuous symmetries are such that one can perform calculus with them. This requires not only a topology, but also the structure of a smooth manifold. A set G which is simultaneously a group and a smooth manifold, in such a way that the two structures are compatible with each other[1], is called a **Lie group**. One could approach the theory starting from this abstract definition, but it requires some background in differential geometry, and it takes some time to establish some of the basic properties. It turns out that most Lie groups of interest[2] are in fact groups of suitable matrices, so we choose to limit our attention to these. In order to avoid conflict (albeit mild) with the general and established mathematical terminology, we will use the term **matrix Lie group** for these slightly less general objects.

In summary, besides their group structure, the key features of Lie groups are their differentiable structure enabling calculus, and their topological structure. The most essential topological features concern the Lie group globally. Especially the topological notions of connectedness[3], simply-connectedness[4], and compactness turn out to be extremely consequential for Lie groups. The differentiable structure, on the other hand, concerns how the Lie groups appear locally, or in fact infinitesimally near the neutral element. We will see how to neatly capture the infinitesimal structure of a Lie group by its **Lie algebra**. This largely reduces the study of Lie groups to the much more straightforward linear algebraic study of their Lie algebras.

Lie groups also have a (topological / differential geometric) notion of dimension. In most cases this is quite intuitive: the dimension is the "number of independent directions to which one can continuously move within the group". As we avoid referring to differential geometry of smooth manifolds explicitly, the easiest precise definition of the dimension will be in terms of the Lie algebra (which is a vector space and therefore has a familiar notion of dimension). We do, however, encourage the reader to think about the dimension of Lie groups intuitively already before its precise definition.

1.2. Matrix Lie groups: definition and examples

We now give some examples of continuous groups. We will revisit many of these examples in more detail later.

The most fundamental of all such examples is the following.

[1]Compatibility of the group structure and smooth structure requires that the group multiplication $\mathrm{G} \times \mathrm{G} \to \mathrm{G}$, $(g, h) \mapsto gh$, and the inverse $\mathrm{G} \to \mathrm{G}$, $g \mapsto g^{-1}$, are smooth functions.

[2]For example, all compact Lie groups are literally matrix groups or obtained from matrix groups by topological coverings.

[3]If a Lie group is not connected, the fruitful approach is to study the connected components.

[4]If a connected Lie group is not simply-connected, the fruitful approach is to study its universal cover — a simply connected space of which the original Lie group is a quotient.

Example II.1 (The general linear group $\mathrm{GL}_n(\mathbb{R})$)**.**
The set of $n \times n$ invertible real matrices

$$\mathrm{GL}_n(\mathbb{R}) = \Big\{ g \in \mathbb{R}^{n\times n} \;\Big|\; \det(g) \neq 0 \Big\}$$

forms a group, called the **general linear group**. This group is naturally viewed as an open subset

$$\mathrm{GL}_n(\mathbb{R}) \subset \mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}$$

of the n^2-dimensional real vector space $\mathbb{R}^{n\times n}$: indeed it is the preimage

$$\mathrm{GL}_n(\mathbb{R}) = \det{}^{-1}\big(\mathbb{R} \setminus \{0\}\big)$$

of the open subset $\mathbb{R} \setminus \{0\} \subset \mathbb{R}$ under the continuous function

$$\det \colon \mathbb{R}^{n\times n} \to \mathbb{R}$$

(the determinant is a polynomial in the matrix entries, therefore in particular continuous). Consequently, $\mathrm{GL}_n(\mathbb{R})$ should clearly be thought of as a "continuous group" on which "differential calculus makes sense".[5] It is intuitively clear that the dimension of $\mathrm{GL}_n(\mathbb{R})$ should be n^2: the general linear group being an open subset in $\mathbb{R}^{n\times n}$, there are n^2 independent directions to which we can vary any given matrix (by at least a small amount) while remaining within the group. For a smooth function $f \colon \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}$ defined on this group, the partial derivatives in these directions are exactly the derivatives with respect to the matrix entries,

$$\frac{\partial}{\partial g_{ij}} f(g), \qquad i, j = 1, \ldots, n.$$

In general, one defines a **Lie subgroup** of a Lie group G to be a subgroup $\mathrm{G}' \subset \mathrm{G}$ which is also a closed subset. It is possible to show that such a closed subgroup G' inherits a manifold structure from the Lie group G, and thus itself becomes a Lie group. We will not prove this fact, but the ideas of the proof are largely contained in Von Neumann's theorem about closed subgroups of $\mathrm{GL}_n(\mathbb{R})$, which we will prove later. This fact does, nevertheless, motivate the following definition.

Definition II.2 (Matrix Lie group)**.**
A **matrix Lie group** is a subgroup

$$\mathrm{G} \subset \mathrm{GL}_n(\mathbb{R}) \qquad (\text{for some } n \in \mathbb{N})$$

which is also a closed subset.

We are immediately ready for a number of examples.

Example II.3 (The special linear group $\mathrm{SL}_n(\mathbb{R})$)**.**
The set of $n \times n$ real matrices with determinant one

$$\mathrm{SL}_n(\mathbb{R}) = \Big\{ g \in \mathbb{R}^{n\times n} \;\Big|\; \det(g) = 1 \Big\}$$

is a subgroup of $\mathrm{GL}_n(\mathbb{R})$ called the **special linear group**. The subgroup property is easy to check directly, but we will establish it below through a different observation.

[5]Note that the operations of multiplication and inverse in the general linear group are smooth maps (i.e., infinitely differentiable, C^∞). Indeed the product of two group elements $g = (g_{ij})_{i,j=1}^n$, $h = (h_{ij})_{i,j=1}^n \in \mathrm{GL}_n(\mathbb{R})$ is the matrix gh whose entries $(gh)_{ij} = \sum_{k=1}^n g_{ik} h_{kj}$ are polynomials (hence smooth) in the entries of g and h. Likewise, the inverse g^{-1} of g has entries which are rational functions (hence smooth) in the entries of g. Thus the group multiplication and inverse can be studied by means of differential calculus.

In order to see that the special linear group is a matrix Lie group, we should in addition verify that it is $\mathrm{SL}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$ is a closed subset. For this, note again that the determinant $\det\colon \mathbb{R}^{n\times n} \to \mathbb{R}$ is a continuous function, so $\mathrm{SL}_n(\mathbb{R})$ is indeed closed as the preimage

$$\mathrm{SL}_n(\mathbb{R}) = \det{}^{-1}(\{1\})$$

of the closed subset $\{1\} \subset \mathbb{R}$. We can also view the determinant as defined only on the subset $\mathrm{GL}_n(\mathbb{R}) \subset \mathbb{R}^{n\times n}$, where it becomes a group homomorphism $\det\colon \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$. The above observation therefore also implies that $\mathrm{SL}_n(\mathbb{R}) = \det{}^{-1}(\{1\}) = \mathrm{Ker}(\det)$ is a (normal) subgroup of $\mathrm{GL}_n(\mathbb{R})$.

The special linear group is a continuous group, although we must vary matrix entries carefully to ensure that the determinant doesn't change: the condition of determinant equals one reduces the dimension by one, and we will see that $\mathrm{SL}_n(\mathbb{R})$ is (n^2-1)-dimensional.

Example II.4 (The group $\mathrm{B}_n(\mathbb{R})$ of upper triangular matrices)**.**
The set of $n \times n$ invertible upper triangular real matrices

$$\mathrm{B}_n(\mathbb{R}) = \left\{ g \in \mathbb{R}^{n\times n} \;\middle|\; \forall i > j : g_{ij} = 0,\ \forall i : g_{ii} \neq 0 \right\}$$

is a subgroup of $\mathrm{GL}_n(\mathbb{R})$. Indeed, the inverse of an invertible upper triangular matrix is upper triangular, and the product of two upper triangular matrices is upper triangular.

In order to see it is a matrix Lie group, we should verify closedness. An easy way is to notice that if a sequence $(g^{(m)})_{m\in\mathbb{N}}$ in $\mathrm{B}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$ converges to a limit g in $\mathrm{GL}_n(\mathbb{R})$, then the limit g must be upper triangular (for $i > j$ indeed $g_{ij} = \lim_{m\to\infty} g_{ij}^{(m)} = 0$) and have non-zero diagonal entries (otherwise the determinant $\det(g)$ vanishes, so the limit g was not in the general linear group to start with).

The dimension of $\mathrm{B}_n(\mathbb{R})$ turns out to be the number of upper triangular entries, $\frac{n^2+n}{2}$, unsurprisingly since the $\mathrm{B}_n(\mathbb{R})$ naturally forms an open subset in the space where these entires are considered the coordinates.

Example II.5 (The group $\mathrm{N}_n(\mathbb{R})$ of unipotent upper triangular matrices)**.**
The set of $n \times n$ invertible upper triangular unipotent real matrices (i.e., upper triangular with ones on the diagonal)

$$\mathrm{N}_n(\mathbb{R}) = \left\{ g \in \mathbb{R}^{n\times n} \;\middle|\; \forall i > j : g_{ij} = 0,\ \forall i : g_{ii} = 1 \right\}$$

is a subgroup of $\mathrm{B}_n(\mathbb{R})$ (and therefore also of $\mathrm{GL}_n(\mathbb{R})$), and a closed subset. The checks are similar to the previous example.

The group $\mathrm{N}_n(\mathbb{R})$ is obviously parametrized by its strictly upper triangular entries, so its dimension should clearly be $\frac{n^2-n}{2}$.

The following example is perhaps more interesting. We consider Lie groups to be real manifolds, and have defined matrix Lie groups as closed subgroups of the real general linear group, but the general linear group of complex matrices can also be seen as a matrix Lie group!

Example II.6 (Complex general linear group)**.**
The set of $n \times n$ invertible complex matrices

$$\mathrm{GL}_n(\mathbb{C}) = \left\{ g \in \mathbb{C}^{n\times n} \;\middle|\; \det(g) \neq 0 \right\}$$

also forms a group, which we call the **complex general linear group**. A complex $n \times n$ matrix $g \in \mathbb{C}^{n\times n}$ can be viewed as a real $2n \times 2n$ matrix, through replacing each complex entry $g_{ij} = x + \mathbb{i}y \in \mathbb{C}$ by the real 2×2 block

$$\begin{bmatrix} x & -y \\ y & x \end{bmatrix}.$$

Since the multiplication of such blocks behaves just as multiplication of complex numbers,

$$\begin{bmatrix} x & -y \\ y & x \end{bmatrix} \begin{bmatrix} x' & -y' \\ y' & x' \end{bmatrix} = \begin{bmatrix} xx' - yy' & -xy' - yx' \\ xy' + yx' & xx' - yy' \end{bmatrix}$$

in accordance with $(x + \mathbb{i}y)(x' + \mathbb{i}y') = xx' - yy' + \mathbb{i}(xy' + yx')$, this embedding of $\mathbb{C}^{n\times n}$ into $\mathbb{R}^{2n\times n}$ respects products of matrices, and in particular also invertibility. Thus we can interpret $\mathrm{GL}_n(\mathbb{C}) \subset \mathrm{GL}_{2n}(\mathbb{R})$ as a subgroup. The subgroup is closed, because if we have a convergent sequence of matrices which consists of 2×2 blocks of the above type, then the limit also consists of such 2×2 blocks.

In summary, also $\mathrm{GL}_n(\mathbb{C})$ is a matrix Lie group. We should expect its dimension to be $2n^2$, because $\mathrm{GL}_n(\mathbb{C}) \subset \mathbb{C}^{n\times n}$ is an open subset, and the real and imaginary parts of the n^2 entries can be independently varied (by a small amount, at least).

We finish with four examples, which are extremely important in a number of different applications.

Example II.7 (The orthogonal group O_n)**.**
We denote by O_n the set of $n \times n$ orthogonal matrices,

$$\mathrm{O}_n = \left\{ g \in \mathbb{R}^{n\times n} \;\middle|\; g^\top g = \mathbb{I}_n \right\},$$

and we call it the **orthogonal group**. It is straightforward to check that it is a subgroup of $\mathrm{GL}_n(\mathbb{R})$. Indeed, for example if we have $g, h \in \mathrm{O}_n$, then the matrix multiplication gh gives an $n \times n$ matrix, whose transpose is $(gh)^\top = h^\top g^\top$, and therefore we have $(gh)^\top g\, h = h^\top g^\top g\, h = h^\top \mathbb{I}\, h = h^\top h = \mathbb{I}$. Also if $g \in \mathrm{O}_n$, then g is invertible and $g^{-1} = g^\top$, so the inverse satisfies $(g^{-1})^\top g^{-1} = (g^{-1})^\top g^\top = (gg^{-1})^\top = \mathbb{I}^\top = \mathbb{I}$ and we see that also $g^{-1} \in \mathrm{O}_n$.

In order to verify that this is a matrix Lie group, we need to show that it is closed. Observe that the condition $g^\top g = \mathbb{I}$ can be written in the form

$$\sum_{k=1}^{n} g_{ki} g_{kj} = \delta_{ij} \qquad \text{for all } i, j = 1, \ldots, n.$$

If we denote by $f_{ij} \colon \mathbb{R}^{n\times n} \to \mathbb{R}$ the function $f_{ij}(g) = \sum_{k=1}^{n} g_{ki} g_{kj}$ which is clearly continuous, then the orthogonal group is the intersection of preimages of closed subsets (singletons $\{0\} \subset \mathbb{R}$ or $\{1\} \subset \mathbb{R}$)

$$\mathrm{O}_n = \bigcap_{i,j=1}^{n} f_{ij}^{-1}\big(\{\delta_{ij}\}\big),$$

and as such is closed. Therefore the orthogonal group indeed is a matrix Lie group.

We encourage the reader to think about the dimension of the matrix Lie group O_n even before we work it out explicitly later.

Example II.8 (The special orthogonal group SO_n)**.**
We denote by SO_n the set of $n \times n$ orthogonal matrices with determinant one,

$$\mathrm{SO}_n = \left\{ g \in \mathbb{R}^{n\times n} \;\middle|\; g^\top g = g\, g^\top = \mathbb{I}_n,\ \det(g) = 1 \right\},$$

and we call it the **special orthogonal group**. It is again easy to check that it is a subgroup of $\mathrm{GL}_n(\mathbb{R})$, and that it is a closed subset. Therefore also the special orthogonal group is a matrix Lie group.

The orthogonal group O_n can be seen to consist of two parts: elements of determinant $+1$ form the subgroup $\mathrm{SO}_n \subset \mathrm{O}_n$, and elements of determinant -1 form a subset which is topologically the same (but not a group). In particular we should expect the dimensions of the matrix Lie groups SO_n and O_n to be the same, although one is a proper subset of the other.

Example II.9 (The unitary group U_n).
We denote by U_n the set of $n \times n$ unitary matrices,

$$\mathrm{U}_n = \left\{ g \in \mathbb{C}^{n\times n} \;\middle|\; g^\dagger g = \mathbb{I}_n \right\},$$

where $g^\dagger$ denotes the conjugate transpose of $g \in \mathbb{C}^{n\times n}$,

$$(g^\dagger)_{ij} = \overline{g_{ji}}.$$

We call U_n the **unitary group**. It is straightforward to check[6] that $\mathrm{U}_n \subset \mathrm{GL}_n(\mathbb{C})$ is a subgroup and that it is closed. Now recalling further that $\mathrm{GL}_n(\mathbb{C}) \subset \mathrm{GL}_{2n}(\mathbb{R})$, we see that the unitary group is a matrix Lie group.

We again invite the reader to think about the dimension of the group U_n already.

Example II.10 (The special unitary group SU_n).
We denote by SU_n the set of $n \times n$ unitary matrices with determinant one,

$$\mathrm{SU}_n = \left\{ g \in \mathbb{C}^{n\times n} \;\middle|\; g^\dagger g = \mathbb{I}_n,\ \det(g) = 1 \right\},$$

and we call it the **special unitary group**. As before, we have $\mathrm{SU}_n \subset \mathrm{GL}_n(\mathbb{C}) \subset \mathrm{GL}_{2n}(\mathbb{R})$ are closed subgroups, so the special unitary group is a matrix Lie group.

Once more, the reader is advised to think about the dimension of the group SU_n.

We leave it as an exercise to work out similarly the symplectic group, i.e., the group of linear transformations which preserves a symplectic form (just like the orthogonal group is the group of linear transformations which preserve an inner product).

Exercise II.11 (Symplectic group).
Let $J \in \mathbb{R}^{2n\times 2n}$ be an antisymmetric matrix such that $J^2 = -\mathbb{I}$. Define a bilinear form $\boldsymbol{\omega}$ on $\mathbb{R}^{2n}$ by $\boldsymbol{\omega}(v, w) = v^\top J w$. Let Sp_{2n} be the set of those $g \in \mathbb{R}^{2n\times 2n}$ which preserve the form $\boldsymbol{\omega}$ in the sense that for all $v, w \in \mathbb{R}^{2n}$ we have $\boldsymbol{\omega}(v, w) = \boldsymbol{\omega}(gv, gw)$.

Prove that the definition above is equivalent to

$$\mathrm{Sp}_{2n} = \left\{ g \in \mathbb{R}^{2n\times 2n} \;\middle|\; g^\top J g = J \right\},$$

and show that $\mathrm{Sp}_{2n} \subset \mathrm{GL}_{2n}(\mathbb{R})$ is a closed subgroup.

We call the matrix Lie group Sp_{2n} the **symplectic group**.

We will also encounter further examples later.

1.3. Topological considerations

Let us first address the notions of compactness and connectedness in the context of Lie groups, and especially matrix Lie groups.

Connectedness

There are two a priori different connectedness notions that we use. For a matrix Lie group (more generally a Lie group), we in practice care mostly about the following.

Definition II.12 (Path-connectedness).
Let $\mathrm{G} \subset \mathrm{GL}_n(\mathbb{R})$ be a matrix Lie group. We say that G is **path-connected**,

[6]Compare with the corresponding properties for the orthogonal group.

if for any $g, h \in \mathrm{G}$ there exists a continuous function $\gamma\colon [0,1] \to \mathrm{G}$ (called a **path**) such that $\gamma(0) = g$ and $\gamma(1) = h$.

We furthermore say that G is **smoothly path-connected**, if the path $\gamma\colon [0,1] \to \mathrm{G}$ above can be chosen so that it is piecewise smooth as a map from $[0,1]$ to $\mathbb{R}^{n\times n}$ (note that $\mathrm{G} \subset \mathrm{GL}_n(\mathbb{R}) \subset \mathbb{R}^{n\times n}$).

In general topology, on the other hand, the usual notion of connectedness is the following — easiest defined through its negation: disconnectedness.

Definition II.13 (Connectedness)**.**
A topological space $\mathfrak{X}$ is **disconnected**, if there exist two open subsets $U, V \subset \mathfrak{X}$ which are non-empty $U, V \neq \emptyset$, disjoint $U \cap V = \emptyset$, and which cover the whole space $U \cup V = \mathfrak{X}$.

A topological space $\mathfrak{X}$ is **connected** if it is not disconnected, i.e., if no two disjoint open non-empty subsets cover the whole space.

Path-connectedness is a priori a stronger notion than connectedness.

Lemma II.14 (Path-connectedness implies connectedness)**.**
If G *is path-connected, then it is also connected.*

Proof. Suppose, by contrapositive, that G is path-connected but not connected. By disconnectedness, we can find two disjoint non-empty open subsets $U, V \subset \mathrm{G}$ which cover G. Since both are non-empty, we can pick points $g \in U$ and $h \in V$, and by path-connectedness there exists a continuous $\gamma\colon [0,1] \to \mathrm{G}$ such that $\gamma(0) = g$ and $\gamma(1) = h$. Then the preimages $\gamma^{-1}(U) \subset [0,1]$ and $\gamma^{-1}(V) \subset [0,1]$ are open, and they are by construction also disjoint and cover the interval $[0,1]$. By our choice of end points of the path γ, we have $0 \in \gamma^{-1}(U)$ and $1 \in \gamma^{-1}(V)$.

Let now $s_0 := \sup \gamma^{-1}(U)$. From the above observations we see that $0 < s_0 < 1$. By construction on the one hand $\gamma(t) \in V$ for all $t > s_0$, and on the other hand there exists a sequence $(t_n)_{n\in\mathbb{N}}$ such that $t_n \uparrow s_0$ as $n \to \infty$ and $\gamma(t_n) \in U$ for all n. But since U and V cover G, we must have either $\gamma(s_0) \in U$ or $\gamma(s_0) \in V$. The former possibility contradicts $\gamma(t) \in V$ for all $t > s_0$, since $\gamma^{-1}(U)$ is open. The latter possibility contradicts the existence of the sequence (t_n) above, since $\gamma^{-1}(V)$ is open. We have derived a contradiction, which proves that a disconnected G can not be path-connected. □

Although in view of Lemma II.14, path-connectedness a priori appears stronger than connectedness, for manifolds the two notions are equivalent.

Fact II.15 (Equivalence of connectedness and path-connectedness for manifolds)**.**
A smooth manifold is connected if and only if it is smoothly path-connected.

Idea of proof of Fact II.15. In view of Lemma II.14, it is sufficient to show that a connected manifold is path-connected. Suppose, by contrapositive, that M is a connected manifold which is not path-connected. Therefore there exists two points $p, q \in M$, which can not be joined by any continuous path γ in M. Consider the set $U \subset M$ of those points p', which can be connected by a path to the point p. By working in local coordinate charts, it is easy to see that U is open. Consider also the complement $V = M \setminus U$, consisting of those points which can not be connected to p by a path. Again by working in local coordinate charts, it is easy to see that V is open. Now U and V cover M and are disjoint by construction, they were shown to be open, and they are non-empty since $p \in U$ and $q \in V$. This contradicts connectedness of M, and proves the assertion. □

In view of this fact, in the context of Lie groups, we will henceforth usually use the term connectedness to mean the logically equivalent but more practical notion of smoothly path-connectedness.

Example II.16 (Some disconnected Lie groups: $\mathrm{GL}_n(\mathbb{R})$, O_n, $\mathrm{B}_n(\mathbb{R})$).
Recall from Example II.1 that the general linear group $\mathrm{GL}_n(\mathbb{R}) \subset \mathbb{R}^{n\times n}$ is the subset of those matrices, whose determinant is non-vanishing,

$$\mathrm{GL}_n(\mathbb{R}) = \det{}^{-1}\big(\mathbb{R} \setminus \{0\}\big).$$

We can write $\mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, +\infty)$, a disjoint union of two open subsets, and correspondingly

$$\mathrm{GL}_n(\mathbb{R}) = \det{}^{-1}\big((0, +\infty)\big) \;\cup\; \det{}^{-1}\big((-\infty, 0)\big).$$

The subsets $\det{}^{-1}((0, +\infty)), \det{}^{-1}((-\infty, 0)) \subset \mathrm{GL}_n(\mathbb{R})$ are non-empty, open, and disjoint. Therefore $\mathrm{GL}_n(\mathbb{R})$ is by definition disconnected. It then follows from Lemma II.14 that $\mathrm{GL}_n(\mathbb{R})$ is neither path-connected nor smoothly path-connected.

Exactly similarly one can show that, e.g, O_n is disconnected: the orthogonal matrices of positive determinant and negative determinant are two disjoint non-empty open subsets that cover all of O_n. Since the determinant of an orthogonal matrix is in fact ± 1, the former of these two is just the special orthogonal group, $\mathrm{SO}_n \subset \mathrm{O}_n$. In Exercise II.17 below, you show that SO_n is connected, so we can conclude that $\mathrm{SO}_n \subset \mathrm{O}_n$ is the connected component which contains the neutral element $\mathbb{I} \in \mathrm{O}_n$.

The group $\mathrm{B}_n(\mathbb{R})$ is also not connected: the i:th diagonal matrix entry defines a continuous function $g \mapsto g_{ii} \in \mathbb{R}$, and the subsets where this (non-zero) entry is positive or negative are two disjoint open subsets which cover the whole space. The signs of all n diagonal entries combined in fact split $\mathrm{B}_n(\mathbb{R})$ to 2^n different components.

Exercise II.17 (Connectedness of the special orthogonal group).
Show that the set $\mathrm{SO}_n = \big\{g \in \mathbb{R}^{n\times n} \,\big|\, g^\top g = \mathbb{I},\ \det(g) = 1\big\}$ is path-connected, i.e., that for any $g \in \mathrm{SO}_n$, there exists a continuous path $\gamma \colon [0, 1] \to \mathrm{SO}_n$ such that $\gamma(0) = g$ and $\gamma(1) = \mathbb{I}$.

Hint: One concrete proof strategy is to implement the following idea. Recalling that elements of SO_n can be identified with ordered orthonormal bases $v_1, \ldots, v_n$ in $\mathbb{R}^n$, it is sufficient to continuously deform a given orthonormal basis to the standard basis. To do this, it is convenient to first show that the first basis vector v_1 can be continuously rotated to the first standard basis vector. Then inductively in the dimension n, one can show that in the $n-1$-dimensional orthogonal complement to the first standard basis vector, one can deform the remaining basis vectors.

Example II.18 (Connectedness of $\mathrm{N}_n(\mathbb{R})$).
Let us check that the group $\mathrm{N}_n(\mathbb{R})$ of upper triangular matrices with ones on the diagonal is (smoothly) path-connected. So, let $g, h \in \mathrm{N}_n(\mathbb{R})$ be two elements of this group. Then for $t \in [0, 1]$, the matrix obtained as the convex combination

$$(1 - t)\, g + t\, h \in \mathbb{R}^{n\times n}$$

is also upper triangular with ones on the diagonal. The formula $\gamma(t) = (1 - t)g + th$ thus defines a (smooth) path $\gamma \colon [0, 1] \to \mathrm{N}_n(\mathbb{R})$ with $\gamma(0) = g$ and $\gamma(1) = h$.

Example II.19 (Complex general linear group is connected).
Unlike the real general linear group, the complex general linear group $\mathrm{GL}_n(\mathbb{C})$ turns out to be connected. Indeed, to show path-connectedness, let $g, h \in \mathrm{GL}_n(\mathbb{C})$ be two elements. Consider the complex linear combinations of these two matrices, specifically combinations of the form $(1 - z)g + zh \in \mathbb{C}^{n\times n}$, with $z \in \mathbb{C}$. For $z = 0$ we recover $g \in \mathrm{GL}_n(\mathbb{C})$, and for $z = 1$ we recover $h \in \mathrm{GL}_n(\mathbb{C})$. For an arbitrary z, such a linear combinations need not be

invertible, but the determinant

$$\det\Big((1-z)g + zh\Big)$$

is a polynomial in z, and thus has only finitely many zeroes $z_1, \ldots, z_m \in \mathbb{C}$. Therefore it is possible to choose a (smooth) path $\zeta \colon [0,1] \to \mathbb{C}$ in the complex plane from $\zeta(0) = 0$ to $\zeta(1) = 1$, which avoids all these zeroes. Then $\gamma(t) := (1 - \zeta(t))\, g + \zeta(t)\, h$ defines a (smooth) path $\gamma \colon [0,1] \to \mathrm{GL}_n(\mathbb{C})$ such that $\gamma(0) = g$ and $\gamma(1) = h$. This proves that $\mathrm{GL}_n(\mathbb{C})$ is (smoothly) path-connected.

Exercise II.20 (Connected Lie groups generated by any neighborhood of the neutral element)**.**
Let G be a connected Lie group, and $U \subset \mathrm{G}$ an open neighborhood of the neutral element $e \in \mathrm{G}$. Show that the elements in U generate the entire group G (i.e., that the smallest subgroup of G containing U is the entire group).
Hint: Definition II.13 of connectedness is useful here.

Compactness

There are also in principle two compactness notions that we use.

Definition II.21 (Sequential compactness)**.**
A topological space $\mathfrak{X}$ is said to be **sequentially compact** if for any sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathfrak{X}$, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ which converges, i.e., the limit $\lim_{k\to\infty} x_{n_k} \in \mathfrak{X}$ exists.

Definition II.22 (Compactness)**.**
A topological space $\mathfrak{X}$ is said to be **compact** if for every collection $(U_\alpha)_{\alpha\in A}$ of open sets $U_\alpha \subset \mathfrak{X}$ which cover the whole space, $\bigcup_{\alpha\in A} U_\alpha = \mathfrak{X}$, there exists a finite subcollection $(U_{\alpha_j})_{j=1}^m$ which also covers the whole space, $\bigcup_{j=1}^m U_{\alpha_j} = \mathfrak{X}$.

For metric spaces and metrizable topological spaces, the above notions are equivalent.

Lemma II.23 (The two notions of compactness are equivalent for metric spaces)**.**
A metric space is compact if and only if it is sequentially compact.

Manifolds are metrizable topological spaces, so in the context of Lie groups we can refer to either of these equivalent properties as compactness. Even more concretely, matrix Lie groups

$$\mathrm{G} \subset \mathrm{GL}_n(\mathbb{R}) \subset \mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}$$

are subsets of the Euclidean space $\mathbb{R}^d$ with $d = n^2$, and compactness in this situation has the following familiar characterization.

Proposition II.24 (Heine-Borel theorem)**.**
For a subset $S \subset \mathbb{R}^d$, the following are equivalent:

(i) *S is compact.*
(ii) *S is sequentially compact.*
(iii) *S is closed and bounded.*

Example II.25 (Examples of non-compact Lie groups)**.**

The general linear group $\mathrm{GL}_n(\mathbb{R})$ is clearly not compact: it is neither closed nor bounded as a subset of $\mathbb{R}^{n\times n}$.

If $n \geq 2$, then the special linear group $\mathrm{SL}_n(\mathbb{R})$ is not bounded, since it contains in particular the diagonal matrices with diagonal entries $k, k^{-1}, 1, \ldots, 1$ for any $k \neq 0$. In particular the special linear group $\mathrm{SL}_n(\mathbb{R})$ is not compact if $n \geq 2$.

Example II.26 (Examples of compact Lie groups)**.**

The groups O_n, SO_n, U_n, SU_n are compact. We have indeed already seen that they are closed subsets of $\mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}$ or $\mathbb{C}^{n\times n} \cong \mathbb{R}^{2n^2}$. Their boundedness follows rather easily from the defining conditions.

2. Lie algebras of matrix Lie groups

2.1. Commutators

The notion of a commutator makes sense on any associative algebra. For our purposes, the most important case is the algebra $\mathrm{End}(V)$ of linear self-maps of a vector space V, so we give the following definition.

Definition II.27 (Commutator)**.**
For $X, Y \in \mathrm{End}(V)$, we define the **commutator** of X and Y as[7]

$$[X, Y] := X \circ Y - Y \circ X \ \in \mathrm{End}(V)\,. \tag{II.1}$$

We leave it as an exercise to prove the following two key properties of commutators.

Exercise II.28 (Antisymmetry and Jacobi identity for commutators)**.**
Show that the commutators satisfy the **antisymmetry**

$$[X, Y] = -[Y, X] \tag{II.2}$$

and **Jacobi identity**

$$\big[[X, Y], Z\big] + \big[[Y, Z], X\big] + \big[[Z, X], Y\big] = 0. \tag{II.3}$$

for all $X, Y, Z \in \mathrm{End}(V)$.

Lie algebras, to be defined in Section 2.4, are not associative algebras. Rather, the two formulas (II.2) and (II.3) established above for commutators will be taken as axioms for Lie brackets when we give the general abstract definition of a Lie algebra. In other words, commutators are important examples of Lie brackets, but an abstract Lie bracket is not necessarily a commutator.

2.2. Matrix exponentials and Lie's formulas

Let V be a finite-dimensional vector space over $\mathbb{k} = \mathbb{R}$ or $\mathbb{k} = \mathbb{C}$. If $d = \dim(V)$, then by choosing a basis for V we can identify

$$V \cong \mathbb{R}^d \qquad \text{or} \qquad V \cong \mathbb{C}^d,$$

respectively. We let $\mathrm{End}(V)$ denote the set of all linear maps of V to itself, which can be identified with the set of all $d \times d$ square matrices

$$\mathrm{End}(V) \cong \mathbb{R}^{d\times d} \qquad \text{or} \qquad \mathrm{End}(V) \cong \mathbb{C}^{d\times d}.$$

Therefore $\mathrm{End}(V)$ is itself a finite-dimensional vector space, $\dim(\mathrm{End}(V)) = d^2$. The set

$$\mathrm{Aut}(V) \subset \mathrm{End}(V)$$

of linear bijections of V to itself forms a group, the automorphism group of V. We can identify it with the group of invertible $d \times d$ matrices,

$$\mathrm{Aut}(V) \cong \mathrm{GL}_d(\mathbb{R}) \qquad \text{or} \qquad \mathrm{Aut}(V) \cong \mathrm{GL}_d(\mathbb{C}).$$

[7]On the right hand side the product is that of $\mathrm{End}(V)$, i.e., the composition of linear maps. The careful (but obviously too cumbersome) formula would therefore be $[X, Y] = X \circ Y - Y \circ X$.

All norms on a given finite-dimensional vector space are equivalent (i.e., they determine the same topology). To verify statements about limits in V and in $\mathrm{End}(V)$ we can therefore in principle use any norm we like. Once a norm $\|\cdot\|$ on V has been chosen, a convenient choice of norm on $\mathrm{End}(V)$ is the operator norm

$$\|X\|_{\mathrm{op}} := \sup\Big\{\|Xv\| \;\Big|\; v \in V,\ \|v\| = 1\Big\} \qquad \text{for } X \in \mathrm{End}(V),$$

because it has the submultiplicativity property

$$\|X \circ Y\|_{\mathrm{op}} \;\leq\; \|X\|_{\mathrm{op}}\,\|Y\|_{\mathrm{op}} \qquad \text{for } X, Y \in \mathrm{End}(V).$$

Below we assume that a basis choice has been made, so that an identification of linear maps with $d \times d$-matrices is fixed. Since moreover real matrices can be viewed as a special case of complex matrices, it suffices to present the following general statements only for the algebra $\mathbb{C}^{n\times n}$ of complex matrices, equipped with the operator norm $\|\cdot\|_{\mathrm{op}}$ inherited from the standard norm $\|\cdot\|$ on $\mathbb{C}^d$.

The first important limit statement for us is the following familiar fact.

Lemma II.29 (Matrix exponential)**.**
For any $X \in \mathbb{C}^{d\times d}$, the series

$$\mathrm{Exp}(X) := \sum_{n=0}^{\infty} \frac{1}{n!} X^n \tag{II.4}$$

is absolutely convergent in the norm $\|\cdot\|_{\mathrm{op}}$.

We leave the following two important properties of the matrix exponential as exercises.

Exercise II.30 (One parameter group property of matrix exponentials)**.**
Let $X \in \mathbb{C}^{n\times n}$, and consider $\mathrm{Exp}(tX)$ for $t \in \mathbb{R}$. Show that we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Exp}(tX) = X\,\mathrm{Exp}(tX) = \mathrm{Exp}(tX)\; X. \tag{II.5}$$

Conclude also that for any $t, s \in \mathbb{R}$, we have

$$\mathrm{Exp}((t+s)X) = \mathrm{Exp}(tX)\;\mathrm{Exp}(sX) \tag{II.6}$$

and in particular

$$\big(\mathrm{Exp}(tX)\,\big)^{-1} = \mathrm{Exp}(-tX)\,. \tag{II.7}$$

Exercise II.31 (Derivative of determinant)**.**
The determinant of an $n \times n$ complex matrix $M = (M_{ij})_{i,j=1}^{n}$ defines a polynomial function

$$\det \colon \mathbb{C}^{n\times n} \to \mathbb{C}$$

in the matrix entries. Show that the partial derivatives of this function evaluated at the unit matrix $\mathbb{I}$ are

$$\frac{\partial}{\partial M_{ij}} \det(M)\Big|_{M=\mathbb{I}} = \delta_{i,j}.$$

Use this to show that for any $X \in \mathbb{C}^{n\times n}$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \det\big(\mathrm{Exp}(tX)\,\big)\Big|_{t=0} = \mathrm{tr}(X). \tag{II.8}$$

Exercise II.32 (Determinant of matrix exponential)**.**
Show that for any $X \in \mathbb{C}^{n\times n}$ we have

$$\det\big(\mathrm{Exp}(X)\big) = e^{\mathrm{tr}(X)}. \tag{II.9}$$

Hint: Let $M(t) := \mathrm{Exp}(tX)$ for all $t \in \mathbb{R}$, and consider $D(t) := \det\big(M(t)\big)$ and $T(t) := e^{t\,\mathrm{tr}(X)}$. To prove the equality $D(t) = T(t)$ for all $t \in \mathbb{R}$, observe that the equality holds for $t = 0$, and establish the differential equations $T'(t) = \mathrm{tr}(X)\,T(t)$ and $D'(t) = \mathrm{tr}(X)\,D(t)$. The proof of the latter differential equation becomes simpler if you use the previous exercise after noting that the one parameter group property of exponentials, $M(t+s) = M(t)M(s)$, implies also $D(t+s) = D(t)D(s)$.

One can define also matrix logarithms for matrices sufficiently close to the unit matrix $\mathbb{I}$.

Lemma II.33 (Matrix logarithm)**.**
For any $A \in \mathbb{C}^{d\times d}$ such that $\|A - \mathbb{I}\|_{\mathrm{op}} < 1$, the series

$$\mathrm{Log}(A) := -\sum_{m=1}^{\infty} \frac{(-1)^m}{m}(A-\mathbb{I})^m \tag{II.10}$$

is absolutely convergent in the norm $\|\cdot\|_{\mathrm{op}}$. For such A we have

$$\mathrm{Exp}(\mathrm{Log}(A)) = A.$$

The natural looking converse, the equality between $\mathrm{Log}(\mathrm{Exp}(X))$ and X, holds for $\|X\|_{\mathrm{op}}$ sufficiently small, but does not hold in general. As an example of its failure, note that for $X = 2\pi\mathbb{i}\,\mathbb{I}$ we clearly have $\mathrm{Exp}(X) = \mathbb{I}$, which implies that the logarithm of the exponential is $\mathrm{Log}(\mathrm{Exp}(X)) = \mathrm{Log}(\mathbb{I}) = 0 \neq X$.

The following Lie's formulas will soon become crucial.

Lemma II.34 (Lie's formulas)**.**
For any $X, Y \in \mathbb{C}^{d\times d}$, we have

$$\big(\mathrm{Exp}(X/n)\,\mathrm{Exp}(Y/n)\big)^n \underset{n\to\infty}{\longrightarrow} \mathrm{Exp}(X+Y) \tag{II.11}$$

and

$$\big(\mathrm{Exp}(X/n)\,\mathrm{Exp}(Y/n)\,\mathrm{Exp}(-X/n)\,\mathrm{Exp}(-Y/n)\big)^{n^2} \underset{n\to\infty}{\longrightarrow} \mathrm{Exp}([X,Y])\,. \tag{II.12}$$

2.3. The Lie algebra of a matrix Lie group

For simplicity of notation, we will from here on denote matrix exponentials also simply by $\mathrm{Exp}(X) = e^X$.

We first define the Lie algebra of a matrix Lie group.

Definition II.35 (The Lie algebra of a matrix Lie group)**.**
For any subgroup $G \subset \mathrm{GL}_d(\mathbb{R})$ which is closed, we define the **Lie algebra of** G to be

$$\mathcal{L}(\mathrm{G}) := \Big\{X \in \mathbb{R}^{d\times d} \;\Big|\; e^{tX} \in \mathrm{G} \quad \forall t \in \mathbb{R}\Big\}\,. \tag{II.13}$$

We will soon check that Lie algebras of matrix Lie groups are vector spaces, and that they are stable under taking commutators. The commutators will be taken to define the Lie bracket in the Lie algebra of a matrix Lie group. Besides establishing these properties, the following important theorem also shows that the matrix exponential lets us use a small neighborhood of the origin in the Lie algebra (which is a vector space) as coordinates for a small neighborhood of the neutral element in the group (which is a manifold).

Theorem II.36 (Von Neumann's theorem)**.**
Let $G \subset \mathrm{GL}_d(\mathbb{R})$ be a closed subgroup. Then $\mathcal{L}(\mathrm{G}) \subset \mathbb{R}^{d\times d}$ is a vector subspace, which has the property

$$\text{if } X, Y \in \mathcal{L}(\mathrm{G}), \text{ then also } [X, Y] \in \mathcal{L}(\mathrm{G}). \tag{II.14}$$

Moreover, $\mathrm{Exp}(\mathcal{L}(\mathrm{G})) \subset \mathrm{G}$ *is a neighborhood of the neutral element* $\mathbb{I} \in \mathrm{G}$*, and if $U \subset \mathcal{L}(\mathrm{G})$ is a sufficiently small open neighborhood of 0 in $\mathcal{L}(\mathrm{G})$, then* $\mathrm{Exp}(U)$ *is an open neighborhood of* $\mathbb{I}$ *in* G *and*

$$\mathrm{Exp}\colon U \to \mathrm{Exp}(U)$$

is a homeomorphism.

Proof. Let us start by showing that $\mathcal{L}(\mathrm{G}) \subset \mathbb{R}^{d\times d}$ is a vector subspace. Suppose that $X \in \mathcal{L}(\mathrm{G})$. Then for any $c \in \mathbb{R}$ and any $t \in \mathbb{R}$ we have

$$e^{t(cX)} = e^{(tc)X} \in \mathrm{G},$$

by virtue of $X \in \mathcal{L}(\mathrm{G})$ and $tc \in \mathbb{R}$. Therefore we have $cX \in \mathcal{L}(\mathrm{G})$, so we have shown that $\mathcal{L}(\mathrm{G})$ is at least stable under scalar multiplication. To show that it is stable under forming sums of vectors, consider $X, Y \in \mathcal{L}(\mathrm{G})$. Then for any $t \in \mathbb{R}$ we have by Lie's formula (II.11)

$$e^{t(X+Y)} = e^{tX+tY} = \lim_{n\to\infty} \left(e^{\frac{tX}{n}} e^{\frac{tY}{n}} \right)^n .$$

By the assumptions $X, Y \in \mathcal{L}(\mathrm{G})$, we have $e^{tX/n} \in \mathrm{G}$ and $e^{tY/n} \in \mathrm{G}$, and therefore also $\left(e^{tX/n} e^{tY/n}\right)^n \in \mathrm{G}$. Since G is closed, also the limit of these as $n \to \infty$ remains in G, so we conclude $e^{t(X+Y)} \in \mathrm{G}$. This shows that $X + Y \in \mathcal{L}(\mathrm{G})$, and finishes the proof that $\mathcal{L}(\mathrm{G})$ is a vector subspace of $\mathbb{R}^{d\times d}$.

Let us then show that $\mathcal{L}(\mathrm{G})$ is stable under forming commutators. So assume again that $X, Y \in \mathcal{L}(\mathrm{G})$. For any $t \in \mathbb{R}$, we have by Lie's formula (II.12)

$$e^{t[X,Y]} = e^{[tX,Y]} = \lim_{n\to\infty} \left(e^{\frac{tX}{n}} e^{\frac{Y}{n}} e^{-\frac{tX}{n}} e^{-\frac{Y}{n}} \right)^{n^2} .$$

By the assumptions $X, Y \in \mathcal{L}(\mathrm{G})$, we again have $e^{\pm tX/n} \in \mathrm{G}$ and $e^{\pm Y/n} \in \mathrm{G}$, and therefore the expression inside the limit above lies in G. Since G is closed, also the limit remains in it, and we get that $e^{t[X,Y]} \in \mathrm{G}$. This shows that $[X, Y] \in \mathcal{L}(\mathrm{G})$, and finishes the proof that $\mathcal{L}(\mathrm{G})$ is stable under forming commutators.

For the rest of the statements, observe the following. We have shown that $\mathcal{L}(\mathrm{G}) \subset \mathbb{R}^{d\times d}$ is a vector subspace, so let us choose a complementary vector subspace $L' \subset \mathbb{R}^{d\times d}$ so that $\mathbb{R}^{d\times d} = \mathcal{L}(\mathrm{G}) \oplus L'$ (vector space direct sum). Consider the function $\mathbb{R}^{d\times d} \to \mathrm{GL}_d(\mathbb{R})$ defined by $X \oplus X' \mapsto e^X e^{X'}$ for $X \in \mathcal{L}(\mathrm{G})$ and $X' \in L'$. From the power series defining the matrix exponentials, is easy to see that the differential at 0 of this function is the identity, so by the inverse function theorem this map gives a homeomorphism between a neighborhood of $0 \in \mathbb{R}^{d\times d}$ and a neighborhood of $e^0 = \mathbb{I} \in \mathrm{GL}_d(\mathbb{R})$.

We now show that $\mathrm{Exp}(\mathcal{L}(\mathrm{G}))$ is a neighborhood of $\mathbb{I} \in \mathrm{G}$. If it were not, then we could find a sequence $(g_n)_{n\in\mathbb{N}}$ in G with $g_n \to \mathbb{I}$ as $n \to \infty$, but $g_n \notin \mathrm{Exp}(\mathcal{L}(\mathrm{G}))$. For large enough n we can write $g_n = e^{X_n} e^{X'_n}$ uniquely with $X_n \in \mathcal{L}(\mathrm{G})$ and $X'_n \in L'$, and we have $X_n \to 0$ and $X'_n \to 0$ as $n \to \infty$. Since both g_n and e^{X_n} are in G, we get that $e^{X'_n} \in \mathrm{G}$ as well. By assumption $g_n \notin \mathrm{Exp}(\mathcal{L}(\mathrm{G}))$, we must have $X'_n \neq 0$. Then consider the elements $\frac{X'_n}{\|X'_n\|} \in L'$ for $n \in \mathbb{N}$. They lie on the unit sphere of L' which is compact, so by passing to a subsequence if necessary we can assume that $\frac{X'_n}{\|X'_n\|} \to X' \in L'$ as $n \to \infty$. Let $t > 0$. Since $\|X'_n\| \to 0$ as $n \to \infty$, we can choose integers m_n such that $m_n \|X'_n\| \to t$ as $n \to \infty$. Now we have $e^{m_n X'_n} = (e^{X'_n})^{m_n} \in \mathrm{G}$, and on the other hand $e^{m_n X'_n} \to e^{tX'}$ as $n \to \infty$, so by closedness of G we get that $e^{tX'} \in \mathrm{G}$. We showed this for any $t > 0$, but we obtain it also for any $t < 0$ by taking inverses. By definition this means that $X' \in \mathcal{L}(\mathrm{G})$, which contradicts $X' \in L'$, $\|X'\| = 1$. Therefore $\mathrm{Exp}(\mathcal{L}(\mathrm{G}))$ is indeed a neighborhood of $\mathbb{I} \in \mathrm{G}$.

Finally, since $X \oplus X' \mapsto e^X e^{X'}$ is a homeomorphism of a neighborhood of $0 \in \mathbb{R}^{d\times d}$ to a neighborhood of $\mathbb{I} \in \mathrm{GL}(V)$, by restricting to $\mathcal{L}(\mathrm{G})$ we obtain a homeomorphism from a neighborhood of $0 \in \mathcal{L}(\mathrm{G})$ to a neighborhood of $\mathbb{I} \in \mathrm{G}$, which is given by $X \mapsto e^X$. This finishes the proof. □

Since the Lie algebra is a vector space, it in particular has a dimension. This enables the following definition.

Definition II.37 (Dimension of a matrix Lie group)**.**
We define the **dimension** of a closed subgroup $G \subset \mathrm{GL}_d(\mathbb{R})$ to be the dimension $\dim\big(\mathcal{L}(\mathrm{G})\big)$ of the vector space $\mathcal{L}(\mathrm{G}) \subset \mathbb{R}^{d\times d}$,

We continue with some examples. In all of the following examples we find an explicit description of the Lie algebra of a specific group, which makes it manifest that the Lie algebra is indeed a vector space. To directly check from the explicit description that the Lie algebra is stable under commutators is usually also very easy, although not always totally immediate.

The first example is almost tautological, but it confirms in particular that the earlier intuitive notion of dimension matches the one defined above.

Example II.38 (The Lie algebra $\mathfrak{gl}_n(\mathbb{R})$ of the general linear group)**.**
The set $\mathrm{GL}_n(\mathbb{R})$ of all invertible $n \times n$ real matrices is a closed subgroup of itself! We denote the Lie algebra of this by

$$\mathfrak{gl}_n(\mathbb{R}) = \mathcal{L}\big(\mathrm{GL}_n(\mathbb{R})\big).$$

By definition, it consists of all matrices $X \in \mathbb{R}^{n\times n}$ such that $\exp(X)$ is invertible, but it follows from (II.7) that this is always the case, so

$$\mathfrak{gl}_n(\mathbb{R}) = \mathbb{R}^{n\times n}. \tag{II.15}$$

The vector space property and stability under commutators would of course be obvious for $\mathfrak{gl}_n(\mathbb{R}) = \mathbb{R}^{n\times n}$ even before Von Neumann's theorem. Now that we have the precise notion of the dimension of a matrix Lie group in Definition II.37, and we have explicitly found the Lie algebra of $\mathrm{GL}_n(\mathbb{R})$, we get that the dimension of $\mathrm{GL}_n(\mathbb{R})$ is

$$\dim\Big(\mathfrak{gl}_n(\mathbb{R})\Big) = \dim\big(\mathbb{R}^{n\times n}\big) = n^2,$$

as anticipated in Example II.1.

The next example concerns already a proper subgroup.

Example II.39 (The Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ of the special linear group)**.**
Recall from Example II.3 that the set $\mathrm{SL}_n(\mathbb{R})$ of matrices of determinant one forms a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. We denote the Lie algebra of this by

$$\mathfrak{sl}_n(\mathbb{R}) = \mathcal{L}\big(\mathrm{SL}_n(\mathbb{R})\big).$$

To get an explicit description of this Lie algebra, observe the following. If $X \in \mathbb{R}^{n\times n}$ is traceless, $\mathrm{tr}(X) = 0$, then by (II.9) we have $\det\big(e^{tX}\big) = 1$ for all $t \in \mathbb{R}$, and consequently $X \in \mathfrak{sl}_n(\mathbb{R})$. Conversely, suppose that $X \in \mathfrak{sl}_n(\mathbb{R})$, i.e., that $\det\big(e^{tX}\big) = 1$ for all $t \in \mathbb{R}$. Using (II.8), we can express the trace of X as the derivative at $t = 0$ of the determinant, and we get

$$\mathrm{tr}(X) \overset{\text{(II.8)}}{=} \frac{\mathrm{d}}{\mathrm{d}t} \det\Big(e^{tX}\Big)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} 1\big|_{t=0} = 0.$$

We have thus shown that the Lie algebra of the special linear group is precisely the set of traceless matrices

$$\mathfrak{sl}_n(\mathbb{R}) = \Big\{ X \in \mathbb{R}^{n\times n} \;\Big|\; \mathrm{tr}(X) = 0 \Big\}. \tag{II.16}$$

The traceless matrices obviously form a vector space, and the trace of a commutator is always zero, $\mathrm{tr}([X,Y]) = \mathrm{tr}(XY) - \mathrm{tr}(YX) = 0$ (by cyclicity of trace), so the set of traceless matrices is also stable under taking commutators. The trace is a surjective linear map $\mathrm{tr}\colon \mathbb{R}^{n\times n} \to \mathbb{R}$, so by rank-nullity theorem we find that the set of traceless matrices has dimension

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

The dimension of the the matrix Lie group $\mathrm{SL}_n(\mathbb{R})$ is thus exactly as anticipated in Example II.3.

We leave the following two examples as exercises.

Exercise II.40 (The Lie algebra $\mathfrak{b}_n(\mathbb{R})$ of the group $\mathrm{B}_n(\mathbb{R})$)**.**
Recall from Example II.4 that the set $\mathrm{B}_n(\mathbb{R})$ of invertible upper triangular real matrices is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. Show that its Lie algebra

$$\mathfrak{b}_n(\mathbb{R}) := \mathcal{L}\big(\mathrm{B}_n(\mathbb{R})\big)$$

is the set of all upper triangular matrices, and conclude that the dimension of the the matrix Lie group $\mathrm{B}_n(\mathbb{R})$ is given, as expected, by

$$\dim\big(\mathfrak{b}_n(\mathbb{R})\big) = \frac{n^2+n}{2}.$$

Show also directly that the commutator of two upper triangular matrices is upper triangular.

Exercise II.41 (The Lie algebra $\mathfrak{n}_n(\mathbb{R})$ of the group $\mathrm{N}_n(\mathbb{R})$)**.**
Recall from Example II.5 that the set $\mathrm{N}_n(\mathbb{R})$ of unipotent upper triangular real matrices is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. Show that its Lie algebra

$$\mathfrak{n}_n(\mathbb{R}) := \mathcal{L}\big(\mathrm{N}_n(\mathbb{R})\big)$$

is the set of all strictly upper triangular matrices, and conclude that the dimension of the the matrix Lie group $\mathrm{N}_n(\mathbb{R})$ is given, as expected, by

$$\dim\big(\mathfrak{n}_n(\mathbb{R})\big) = \frac{n^2-n}{2}.$$

Show also directly that the commutator of two strictly upper triangular matrices is strictly upper triangular.

The next example could be considered as the first really interesting one. For clarity, we isolate the key calculations for the example in the following lemma. The matrix Lie groups O_n and SO_n appearing in the statement were defined in Examples II.7 and II.8, respectively.

Lemma II.42 (Characterizing the Lie algebra of the orthogonal group)**.**
The following three conditions for a matrix $X \in \mathbb{R}^{n\times n}$ are all equivalent:

(a) $e^{tX} \in \mathrm{SO}_n$ *for all* $t \in \mathbb{R}$
(b) $e^{tX} \in \mathrm{O}_n$ *for all* $t \in \mathbb{R}$
(c) X *is antisymmetric,* $X^\top = -X$.

Proof. The implication (a) $\Rightarrow$ (b) is trivial since $\mathrm{SO}_n \subset \mathrm{O}_n$, so it suffices to prove that (b) $\Rightarrow$ (c) and (c) $\Rightarrow$ (a).

Suppose therefore that $e^{tX} \in \mathrm{O}_n$ for all $t \in \mathbb{R}$. Recalling that $M \in \mathrm{O}_n$ amounts to the condition $M^\top M = \mathbb{I}$, and observing that

$$\left(e^{tX}\right)^\top = e^{tX^\top},$$

we find that the assumption (b) can be written as

$$\mathbb{I} = e^{tX^\top} e^{tX} \qquad \text{for all } t \in \mathbb{R}.$$

If we take the derivative of this condition with respect to t (note that the left hand side is constant matrix $\mathbb{I}$, so its derivative is the zero matrix $\mathbb{O}$), we find

$$\begin{aligned}\mathbb{O} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tX^\top} e^{tX}\right) \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}t}e^{tX^\top}\right) e^{tX} + e^{tX^\top} \left(\frac{\mathrm{d}}{\mathrm{d}t}e^{tX}\right) \\ &= X^\top e^{tX^\top} e^{tX} + e^{tX^\top} e^{tX} X \\ &= X^\top + X.\end{aligned}$$

This shows antisymmetry $X^\top = -X$, so we have proven (b) $\Rightarrow$ (c).

Suppose now that X is antisymmetric, $X^\top = -X$. Then for any $t \in \mathbb{R}$ we have

$$\left(e^{tX}\right)^\top = e^{tX^\top} = e^{-tX},$$

which shows that $e^{tX} \in \mathrm{O}_n$. On the other hand all the diagonal entries of an antisymmetric matrix vanish, and therefore the trace is zero as well, $\mathrm{tr}(X) = 0$. This implies by (II.9) that $\det(e^{tX}) = e^{\mathrm{tr}(X)} = e^0 = 1$, so in fact we have $e^{tX} \in \mathrm{SO}_n$, and the last implication (c) $\Rightarrow$ (a) is also proven. □

With this lemma proven, the next example is easily handled.

Example II.43 (The Lie algebra $\mathfrak{so}_n$ of the orthogonal group).
Recall from Examples II.7 and II.8 that the set O_n of orthogonal matrices and the set SO_n of orthogonal matrices of determinant one form closed subgroups of $\mathrm{GL}_n(\mathbb{R})$. We claim that the Lie algebras of both are the same, and let us therefore define

$$\mathfrak{so}_n = \mathcal{L}\big(\mathrm{SO}_n\big).$$

Indeed by definition, property (a) of Lemma II.42 characterizes elements X of the Lie algebra $\mathcal{L}\big(\mathrm{SO}_n\big)$ the special orthogonal group, and property (b) characterizes elements X of the Lie algebra $\mathcal{L}\big(\mathrm{O}_n\big)$ of the orthogonal group. Since both of them are equivalent to property (c), we find that the Lie algebras are the same, and concretely described as the set of antisymmetric matrices

$$\mathfrak{so}_n = \left\{ X \in \mathbb{R}^{n\times n} \;\middle|\; X^\top = -X \right\}. \tag{II.17}$$

The set of antisymmetric matrices is obviously a vector space of dimension

$$\dim\big(\mathfrak{so}_n\big) = \frac{n^2 - n}{2},$$

which is the number of strictly upper triangular entries (these can be chosen independently for an antisymmetric matrix, and they determine the whole matrix). By definition this is then also the dimension of the matrix Lie groups O_n and SO_n.

It can also be checked directly that the commutator of two antisymmetric matrices is antisymmetric. Indeed if X and Y are antisymmetric, i.e., $X^\top = -X$ and $Y^\top = -Y$, then we see that the matrix product XY satisfies

$$(XY)^\top = Y^\top X^\top = (-Y)(-X) = YX,$$

and therefore the antisymmetry of the commutator follows from

$$\begin{aligned}[X,Y]^\top = \big(XY - YX\big)^\top &= (XY)^\top - (YX)^\top \\ &= YX - XY \\ &= -[X,Y].\end{aligned}$$

Exercise II.44 (Lie algebras of the unitary and special unitary groups)**.**
Show that the Lie algebra $\mathfrak{u}_n := \mathcal{L}(\mathrm{U}_n)$ of the unitary group U_n (see Example II.9) is given by

$$\mathfrak{u}_n = \Big\{ X \in \mathbb{C}^{n\times n} \;\Big|\; X^\dagger = -X \Big\} \tag{II.18}$$

and that the Lie algebra $\mathfrak{su}_n := \mathcal{L}(\mathrm{SU}_n)$ of the special unitary group SU_n (see Example II.10) is given by

$$\mathfrak{su}_n = \Big\{ X \in \mathbb{C}^{n\times n} \;\Big|\; X^\dagger = -X,\ \mathrm{tr}(X) = 0 \Big\}. \tag{II.19}$$

Calculate also the dimensions $\dim(\mathfrak{u}_n)$ and $\dim(\mathfrak{su}_n)$.

Exercise II.45 (Symplectic Lie algebra)**.**
Let $J \in \mathbb{R}^{2n\times 2n}$ be an antisymmetric matrix such that $J^2 = -\mathbb{I}$. Recall the definition of the symplectic group

$$\mathrm{Sp}_{2n} = \big\{ g \in \mathbb{R}^{2n\times 2n} \,\big|\, g^\top J g = J \big\}$$

from Exercise II.11. Show that the Lie algebra $\mathfrak{sp}_{2n} := \mathcal{L}(\mathrm{Sp}_{2n})$ of this matrix Lie group is

$$\mathfrak{sp}_{2n} = \Big\{ X \in \mathbb{R}^{2n\times 2n} \;\Big|\; X^\top J + JX = 0 \Big\}.$$

Calculate also the dimension $\dim(\mathfrak{sp}_{2n}(\mathbb{R}))$.

2.4. The abstract notion of a Lie algebra

In Section 2.3 we defined the Lie algebras associated to matrix Lie groups. There is a more general abstract definition of a Lie algebra.

Let $\Bbbk$ be a field.

Definition II.46 (Lie algebra)**.**
A **Lie algebra** over $\Bbbk$ is a $\Bbbk$-vector space $\mathfrak{g}$ equipped with a bilinear operation $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$,

$$(X,Y) \mapsto [X,Y] \in \mathfrak{g} \qquad \text{for } X,Y \in \mathfrak{g},$$

which satisfies **antisymmetry**

$$[X,Y] = -[Y,X] \tag{II.20}$$

and **Jacobi identity**

$$\big[[X,Y],Z\big] + \big[[Y,Z],X\big] + \big[[Z,X],Y\big] = 0. \tag{II.21}$$

The operation $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, is called the **Lie bracket**.

Example II.47 (Lie algebras of matrix Lie groups are Lie algebras)**.**
According to Theorem II.36 and Exercise II.28, any Lie algebra $\mathcal{L}(\mathrm{G})$ of a matrix Lie group $\mathrm{G} \subset \mathrm{GL}_n(\mathbb{R})$ is a special case of the above definition. For $\mathfrak{g} = \mathcal{L}(\mathrm{G})$ the field is $\Bbbk = \mathbb{R}$ real numbers, and the Lie bracket is given by the commutator in $\mathrm{End}\big(\mathbb{R}^d\big) \cong \mathbb{R}^{d\times d}$.

Similarly, if one were to study Lie groups which are complex manifolds rather than (the much more common) real manifolds, one would correspondingly obtain Lie algebras over $\Bbbk = \mathbb{C}$ associated to such complex Lie groups. Also one can associate Lie algebras to algebraic groups defined over other fields, and thus one arrives at even more general choices of $\Bbbk$. But these are not the only reasons for allowing generality in the definition — it turns out also that the fruitful way to study a Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ is often to form its compexification $\mathfrak{g}_{\mathbb{C}} := \mathfrak{g} \otimes \mathbb{C}$, which is a Lie algebra over $\mathbb{C}$. We will return to this later.

Let us give a few more examples.

Example II.48 (Abelian Lie algebras)**.**
Let $\mathfrak{h}$ be any vector space over $\Bbbk$. If we equip it with the zero bilinear map $[\,\cdot,\ \cdot\,] \colon \mathfrak{h} \times \mathfrak{h} \to \mathfrak{h}$,

$$[X, Y] = 0 \qquad \text{for all } X, Y \in \mathfrak{h},$$

then antisymmetry and Jacobi identity obviously hold. Thus $\mathfrak{h}$ is a Lie algebra! We call a Lie algebra whose Lie bracket is zero either **abelian** or **commutative**.

Needless to say, abelian Lie algebras are not the most exciting examples of Lie algebras in their own right. But — it will turn out to be very fruitful to analyze maximally large abelian Lie subalgebras[8] $\mathfrak{h} \subset \mathfrak{g}$ within more interesting Lie algebras $\mathfrak{g}$.

Example II.49 (The Lie algebra of all linear maps)**.**
Let V be a vector space over $\Bbbk$. The set $\mathfrak{g} = \mathrm{End}_{\Bbbk}(V)$ of $\Bbbk$-linear maps $V \to V$ is a vector space over $\Bbbk$. The commutator $[X, Y] = X \circ Y - Y \circ X$ in $\mathrm{End}_{\Bbbk}(V)$ also satisfies antisymmetry and Jacobi identity (as in Exercise II.28). Therefore $\mathfrak{g} = \mathrm{End}_{\Bbbk}(V)$ is a Lie algebra over $\Bbbk$.

Example II.50 (The Lie algebra of all derivations of an associative algebra)**.**
Let A be an associative algebra over a field $\Bbbk$.[9] A map $D \colon A \to A$ is called a **derivation** if it is $\Bbbk$-linear and satisfies the **Leibnitz rule**

$$D(ab) = D(a)\, b + a\, D(b) \qquad \text{for all } a, b \in A.$$

The set $\mathrm{Der}(A) \subset \mathrm{End}_{\Bbbk}(A)$ of all derivations on A is clearly a vector space. Note that if $D_1, D_2 \in \mathrm{Der}(A)$, then a direct calculation with the commutator $[D_1, D_2] := D_1 \circ D_2 - D_2 \circ D_1$ gives

$$\begin{aligned}
[D_1, D_2](ab) &= D_1\Big(D_2(a)\, b + a\, D_2(b)\Big) - D_2\Big(D_1(a)\, b + a\, D_1(b)\Big) \\
&= \quad D_1\big(D_2(a)\big)\, b + D_2(a)\, D_1(b) + D_1(a)\, D_2(b) + a\, D_1\big(D_2(b)\big) \\
&\quad\; - D_2\big(D_1(a)\big)\, b - D_1(a)\, D_2(b) - D_2(a)\, D_1(b) - a\, D_2\big(D_1(b)\big) \\
&= D_1\big(D_2(a)\big)\, b + a\, D_1\big(D_2(b)\big) - D_2\big(D_1(a)\big)\, b - a\, D_2\big(D_1(b)\big) \\
&= \big([D_1, D_2](a)\big)\; b \;+\; a\; \big([D_1, D_2](b)\big),
\end{aligned}$$

[8] The reader can probably already guess the meaning of a Lie subalgebra, although we have postponed formally introducing this notion until Definition III.54 below.

[9] This means... [TODO: write def.]

which shows that also $[D_1, D_2] \in \mathrm{Der}(A)$. Since the commutator in $\mathrm{End}_{\Bbbk}(A)$ satisfies anti-symmetry and Jacobi identity, we get that $\mathfrak{g} = \mathrm{Der}(A)$ is a Lie algebra over $\Bbbk$.

Remark II.51 (Vector fields on a manifold).
An important particular case of the type covered in Example II.50 is the space of all vector fields on a smooth manifold M. Vector fields act as derivations of the algebra $C^\infty(M)$ of smooth functions on M, and they form a Lie algebra. In fact according to the usual abstract definition, the Lie algebra of a Lie group G is the set of all vector fields on the Lie group which are invariant under left multiplication in the group. To show that these fall into the general definition, one only has to verify that the commutator of the derivations by two left-invariant vector fields is again the derivation by a left-invariant vector field.

Having defined the structure called Lie algebra, we next define — as usual — homomorphisms of Lie algebras as maps which preserve this structure.

Definition II.52 (Lie algebra homomorphisms and isomorphisms).
Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be two Lie algebras over $\Bbbk$, with Lie brackets denoted by $[\cdot,\cdot]_{\mathfrak{g}_1}$ and $[\cdot,\cdot]_{\mathfrak{g}_2}$, respectively. A map

$$\phi \colon \mathfrak{g}_1 \to \mathfrak{g}_2$$

is called a **Lie algebra homomorphism** if it is $\Bbbk$-linear and satisfies

$$\phi\big([X,Y]_{\mathfrak{g}_1}\big) = \big[\phi(X), \phi(Y)\big]_{\mathfrak{g}_2} \qquad \text{for all } X, Y \in \mathfrak{g}_1.$$

A **Lie algebra isomorphism** is a bijective Lie algebra homomorphism. If there exists a Lie algebra isomorphism between two Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, then we call the two Lie algebras **isomorphic**, and denote $\mathfrak{g}_1 \cong \mathfrak{g}_2$.

In terms of homomorphisms, we can already define what a representation of a Lie algebra is.

Definition II.53 (Representation of a Lie algebra).
Let $\mathfrak{g}$ be a Lie algebra over $\Bbbk$, and let V be a vector space over $\Bbbk$. A **representation** of $\mathfrak{g}$ on V is a Lie algebra homomorphism

$$\vartheta \colon \mathfrak{g} \to \mathrm{End}_{\Bbbk}(V),$$

where the Lie algebra structure on $\mathrm{End}_{\Bbbk}(V)$ is as in Example II.49.

Remark II.54 (Representations defined over larger fields).
In fact the field $\Bbbk$ over which the Lie algebra $\mathfrak{g}$ is defined does not have to be the same as the field $\mathbb{K}$ over which the vector space V is defined, as long as $\mathbb{K}$ is an extension of $\Bbbk$. The most important such case arises when we consider complex representations of real Lie algebras, i.e., $\Bbbk = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$.

Namely, if $\Bbbk \subset \mathbb{K}$, then all $\mathbb{K}$-linear maps are also $\Bbbk$-linear

$$\mathrm{End}_{\mathbb{K}}(V) \subset \mathrm{End}_{\Bbbk}(V),$$

and the $\mathbb{K}$-vector space $\mathrm{End}_{\mathbb{K}}(V)$ of $\mathbb{K}$-linear maps $V \to V$ can be viewed in particular as a vector space over the smaller field $\Bbbk$.[10] Commutators in $\mathrm{End}_{\mathbb{K}}(V)$ of course coincide with

[10] To view a vector space W over a larger field $\mathbb{K}$ as a vector space over a smaller field $\Bbbk \subset \mathbb{K}$, one simply "forgets" that scalar multiplication is also permitted with scalars outside $\Bbbk$. This of course typically increases dimension: we have $\dim_{\Bbbk}(W) \geq \dim_{\mathbb{K}}(W)$, or more precisely in fact $\dim_{\Bbbk}(W) = \dim_{\mathbb{K}}(W) \dim_{\Bbbk}(\mathbb{K})$. For the most concrete case of $\Bbbk = \mathbb{R} \subset \mathbb{C} = \mathbb{K}$ we have the familiar doubling of dimensions $\dim_{\mathbb{R}}(W) = 2 \dim_{\mathbb{C}}(W)$.

commutators in $\mathrm{End}_{\Bbbk}(V)$, so there is no ambiguity about the choice of the Lie bracket. Thus $\mathrm{End}_{\mathbb{K}}(V) \subset \mathrm{End}_{\Bbbk}(V)$ is a Lie subalgebra, when both are viewed as Lie algebras over $\Bbbk$.

In summary, when $\Bbbk \subset \mathbb{K}$ is a field extension, then a $\mathbb{K}$-representation of a $\Bbbk$-Lie algebra is a $\mathbb{K}$-vector space V equipped with a homomorphism of $\Bbbk$-Lie algebras

$$\vartheta \colon \mathfrak{g} \to \mathrm{End}_{\mathbb{K}}(V) \subset \mathrm{End}_{\Bbbk}(V).$$

We will come to examples later — we first turn to the relationship between homomorphisms of Lie groups and their Lie algebras.

2.5. Homomorphisms of Lie groups and Lie algebras

A homomorphism of Lie groups should respect both the group structure and the structure of a smooth manifold. One therefore defines a Lie group homomorphism as a group homomorphism which is a smooth function. Smoothness in fact follows already by assuming merely continuity. As we are focusing on the particular case of matrix Lie groups, the approproate precise definition is the following.

Definition II.55 (Lie group homomorphisms and isomorphisms)**.**
Let G_1 and G_2 be two matrix Lie groups. A map

$$f \colon \mathrm{G}_1 \to \mathrm{G}_2$$

is called a **Lie group homomorphism** if it is a group homomorphism and a continuous function.

A **Lie group isomorphism** is a group homomorphism, which is also a homeomorphism.

The notion of a representation is a special case. We assume that the underlying space V is a finite dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$. Then the automorphism group $\mathrm{Aut}(V)$ of V can be identified with $\mathrm{GL}_d(\mathbb{R})$ or $\mathrm{GL}_d(\mathbb{C})$, and in particular is itself a matrix Lie group.

Definition II.56 (Representation of a Lie group)**.**
Let G be a be a Lie group, and V be a finite-dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$. A **representation** of G on V is a Lie group homomorphism

$$\rho \colon \mathrm{G} \to \mathrm{Aut}(V)\,.$$

Example II.57 (Trivial representation of a Lie group)**.**
As in the case of finite groups, the real one dimensional vector space $\mathbb{K} = \mathbb{R}$ or the complex one dimensional vector space $\mathbb{K} = \mathbb{C}$ carries a representation of any (matrix) Lie group G through setting each $g \in \mathrm{G}$ to correspond to the identity linear map

$$\rho(g) = \mathrm{id}_{\mathbb{K}} \in \mathrm{Aut}(\mathbb{K}) \cong \mathrm{GL}_1(\mathbb{K}).$$

The homomorphism property and continuity are both obvious in this case.

Example II.58 (Defining representation of a matrix Lie group)**.**
Since a matrix Lie group is by definition a closed subgroup $\mathrm{G} \subset \mathrm{GL}_n(\mathbb{R})$ for some $n \in \mathbb{Z}_{>0}$, the inclusion

$$\iota \colon \mathrm{G} \hookrightarrow \mathrm{GL}_n(\mathbb{R}) \cong \mathrm{Aut}(\mathbb{R}^n)$$

is a continuous group homomorphism. This obvious homomorphism equips the vector space $\mathbb{R}^n$ with the structure of a representation of the Lie group G, and the corresponding representation is known as the **defining representation** of G.

As a word of warning, however, note that the above depends on the way we realize G as a closed subgroup of a general linear group — we could for instance have interpreted an isomorphic copy of G as a subgroup of $\mathrm{GL}_{n'}(\mathbb{R})$ for any $n' > n$, and the notion of a defining representation would have changed.

Moreover, some matrix Lie groups such as $\mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{C})$, U_n, and SU_n are more naturally subgroups of the group of invertible complex matrices $\mathrm{GL}_n(\mathbb{C}) \subset \mathrm{GL}_{2n}(\mathbb{R})$. In this case the defining representation (usually) refers to the complex vector space $\mathbb{C}^n$ equipped with the structure of a representation coming from an inclusion of G into $\mathrm{GL}_n(\mathbb{C})$.

Suppose that G_1 and G_2 are two matrix Lie groups, and $\mathfrak{g}_1 = \mathcal{L}(\mathrm{G}_1)$ and $\mathfrak{g}_2 = \mathcal{L}(\mathrm{G}_2)$ their Lie algebras, respectively. The question we will address next is: To what extent is there a relationship between Lie group homomorhisms

$$f\colon \mathrm{G}_1 \to \mathrm{G}_2$$

and Lie algebra homomorhisms

$$\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2?$$

In particular, to which extent do representations of a (matrix) Lie group G correspond to representations of its Lie algebra $\mathfrak{g}$?

Our answer to this question can be summarized in three statements.

0°) Any Lie group homomorphism $f\colon \mathrm{G}_1 \to \mathrm{G}_2$ gives rise to a Lie algebra homomorphism $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$, by its differential at the neutral element.
1°) If G_1 is connected, then the Lie group homomorphism $f\colon \mathrm{G}_1 \to \mathrm{G}_2$ is uniquely determined by its differential at the neutral element, a Lie algebra homomorphism $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$.
2°) If G_1 is connected and simply-connected, then for any Lie algebra homomorphism $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ there exists a unique Lie group homomorphism $f\colon \mathrm{G}_1 \to \mathrm{G}_2$ whose differential at the neutral element ϕ is.

In this context, assertions 1° and 2° are sometimes called the first and second principles, respectively. Assertion 0° is more basic, so let us call it the zeroth principle. The technicalities in the zeroth principle actually require a longer proof, since it is in this part that we have to promote the continuity assumption to differentiability.

The zeroth principle

As we avoid explicitly using differential geometry, we define the necessary derivatives below directly using the Lie algebra and the matrix exponential.[11]

Consider a continuous group homomorphism

$$f\colon \mathrm{G}_1 \to \mathrm{G}_2.$$

We define the differential $(\mathrm{d}f)|_{\mathbf{e}_1}$ of the function f at the neutral element $\mathbf{e}_1 \in \mathrm{G}_1$ as the following matrix valued function on the Lie algebra $\mathfrak{g}_1$ of G_1

$$(\mathrm{d}f)|_{\mathbf{e}_1}(X_1) := \frac{\mathrm{d}}{\mathrm{d}t} f\big(e^{tX_1}\big)\Big|_{t=0} \qquad \text{for } X_1 \in \mathfrak{g}_1, \tag{II.22}$$

[11] Our definition of course coincides with the differential geometric one, only it is less general.

provided the derivative here exists. Below we prove the existence, and show what the group homomorphism explicitly does to the one-parameter subgroups defined by exponentials.

Theorem II.59 (Differential of homomorphism between matrix Lie groups).
For a continuous group homomorphism

$$f\colon \mathrm{G}_1 \to \mathrm{G}_2,$$

between matrix Lie groups G_1 *and* G_2 *with respective Lie algebras* $\mathfrak{g}_1$ *and* $\mathfrak{g}_2$*, the differential of* f *defined by* (II.22) *exists and is a Lie algebra homomorphism*

$$\phi := (\mathrm{d}f)|_{\mathbf{e}_1} \colon \mathfrak{g}_1 \to \mathfrak{g}_2.$$

Moreover, for any $X_1 \in \mathfrak{g}_1$ *and* $t \in \mathbb{R}$ *we have*

$$f\big(\mathrm{Exp}(tX_1)\big) = \mathrm{Exp}\big(t\,\phi(X_1)\big). \tag{II.23}$$

Remark II.60 (Continuous group homomorphisms between Lie groups are smooth).
In the general theory of Lie groups one can also show that if a group homomorphism between two Lie groups is continuous then it is in fact smooth, and therefore a Lie group homomorphism (i.e. both a group homomorphism and an infinitely differentiable map). The proof is essentially the same, if in the place of matrix exponentials one uses exponential maps defined via flows of left-invariant vector fields. The idea is to first establish the unsurprising equality (II.23), and then the rest follows easily.

Proof of Theorem II.59. By Theorem II.36 we can choose a small enough neighborhood $U_1 \subset \mathfrak{g}_1$ of the origin $0 \in \mathfrak{g}_1$ such that $\mathrm{Exp}\colon U_1 \to \mathrm{Exp}(U_1) \subset \mathrm{G}_1$ is a homeomorphism. Likewise, choose a small enough neighborhood $U_2 \subset \mathfrak{g}_2$ of the origin $0 \in \mathfrak{g}_2$ such that $\mathrm{Exp}\colon U_2 \to \mathrm{Exp}(U_2) \subset \mathrm{G}_2$ is a homeomorphism.

Given $X_1 \in \mathfrak{g}_1$ we can fix a small $t' > 0$ such that whenever $|t| \le t'$, we have $f\big(\mathrm{Exp}(tX_1)\big) \in \mathrm{Exp}(U_2)$ and also $tX_1 \in U_1$. Let us write $X_1' = t'X_1$. Consider now Lie algebra elements of the form $m2^{-n}X_1'$ with $n \in \mathbb{N}$ and $m \in \mathbb{Z}$ such that $|m| \le 2^n$. Since $f\big(\mathrm{Exp}(2^{-n}X_1')\big) \in \mathrm{Exp}(U_2)$ for any $n \in \mathbb{N}$, we can choose a unique $X_2^{(n)} \in U_2 \subset \mathfrak{g}_2$ such that

$$f\Big(\mathrm{Exp}\big(2^{-n}X_1'\big)\Big) = \mathrm{Exp}\Big(X_2^{(n)}\Big).$$

Then also for any m, by the group homomorphism property and properties of the exponential, we get

$$\begin{aligned} f\big(\mathrm{Exp}(m2^{-n}X_1')\big) &= f\Big(\mathrm{Exp}\big(2^{-n}X_1'\big)^m\Big) \\ &= \Big(f\big(\mathrm{Exp}(2^{-n}X_1')\big)\Big)^m = \Big(\mathrm{Exp}\Big(X_2^{(n)}\Big)\Big)^m = \mathrm{Exp}\Big(mX_2^{(n)}\Big). \end{aligned}$$

If we take a larger $n' \ge n$ and consider $m' = m2^{n'-n}$, then $m'2^{-n'} = m2^{-n}$ so by uniqueness of the choice of inverse exponentials we must have $m'X_2^{(n')} = mX_2^{(n)}$. In other words, we have obtained a well defined additive map from the set of dyadic numbers of the form $m2^{-n}$ to $\mathfrak{g}_2$. Since the closure of such dyadic numbers is the interval $[-1, 1]$, by continuity of f and the matrix exponentials and their inverses this map can be extended to a map $[-1, 1] \to \mathfrak{g}_2$, which has the form $t \mapsto tX_2'$ for some $X_2' \in \mathfrak{g}_2$ and satisfies

$$f\big(\mathrm{Exp}(tX_1')\big) = \mathrm{Exp}(tX_2').$$

By homomorphism property of f and properties of matrix exponentials, this continues to hold true for all $t \in \mathbb{R}$, and by simple redefinitions $X_1 = \frac{1}{t'}X_1'$ and $X_2 = \frac{1}{t'}X_2'$, it becomes

$$f\big(\mathrm{Exp}(tX_1)\big) = \mathrm{Exp}(tX_2) \qquad \text{for all } t \in \mathbb{R}.$$

In particular the derivative in (II.22) exists and is given by

$$(\mathrm{d}f)|_{\mathbf{e}_1}(X_1) := \frac{\mathrm{d}}{\mathrm{d}t} f\big(e^{tX_1}\big)\big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Exp}(tX_2)\,\big|_{t=0} = X_2 \in \mathfrak{g}_2.$$

We see that the values are in the Lie algebra $\mathfrak{g}_2$ of G_2, so at least the differential is a map

$$(\mathrm{d}f)|_{\mathbf{e}_1} : \mathfrak{g}_1 \to \mathfrak{g}_2.$$

We have established that $(\mathrm{d}f)|_{\mathbf{e}_1}(tX_1) = t\,(\mathrm{d}f)|_{\mathbf{e}_1}(X_1)$ for all $t \in \mathbb{R}$ and $X_1 \in \mathfrak{g}_1$, but to prove linearity we must still establish additivity. So let $X_1, Y_1 \in \mathfrak{g}_1$. Consider the map

$$\alpha\colon\ \mathbb{R}^2 \to \mathrm{G}_2 \qquad\qquad \alpha(s,t) = f\Big(\mathrm{Exp}(sX_1)\,\mathrm{Exp}(tY_1)\Big).$$

We already know that $\alpha(s,0) = \mathrm{Exp}(sX_2)$ and $\alpha(0,t) = \mathrm{Exp}(tY_2)$, where $X_2 = (\mathrm{d}f)|_{\mathbf{e}_1}(X_1)$ and $Y_2 = (\mathrm{d}f)|_{\mathbf{e}_1}(Y_1)$ are obtained from the construction above. By the homomorphism property of f, we obtain for any $(s,t) \in \mathbb{R}^2$

$$\begin{aligned}\alpha(s,t) &= f\Big(\mathrm{Exp}(sX_1)\,\mathrm{Exp}(tY_1)\Big)\\ &= f\Big(\mathrm{Exp}(sX_1)\Big)\ f\Big(\mathrm{Exp}(tY_1)\Big) = \mathrm{Exp}(sX_2)\ \mathrm{Exp}(tY_2)\,.\end{aligned}$$

This map $\alpha\colon \mathbb{R}^2 \to \mathrm{G}_2 \subset \mathrm{GL}(V_2)$ is infinitely differentiable, since matrix exponentials and matrix products are, and its differential at the origin gives in particular

$$\begin{aligned}\alpha\big(\frac{t}{n},\frac{t}{n}\big) &= \mathbb{I} + \frac{t}{n}X_2 + \frac{t}{n}Y_2 + o\Big(\frac{1}{n}\Big)\\ &= \mathrm{Exp}\bigg(\frac{tX_2+tY_2}{n} + o(n^{-1})\bigg)\,.\end{aligned}$$

From here one can show as before that

$$\begin{aligned}\Big(\alpha\big(\frac{t}{n},\frac{t}{n}\big)\Big)^n &= \bigg(\mathrm{Exp}\bigg(\frac{tX_2+tY_2}{n} + o(n^{-1})\bigg)\bigg)^n\\ &= \Big(\mathrm{Exp}(tX_2+tY_2+o(1))\Big) \longrightarrow \mathrm{Exp}\Big(t(X_2+Y_2)\Big) \qquad \text{as } n\to\infty.\end{aligned}$$

On the other hand by homomorphism property, continuity, and Lie's formula (II.11) we have

$$\begin{aligned}\Big(\alpha\big(\frac{t}{n},\frac{t}{n}\big)\Big)^n &= f\Big(\big(\mathrm{Exp}\big(n^{-1}tX_1\big)\,\mathrm{Exp}\big(n^{-1}tY_1\big)\,\big)^n\Big)\\ &\longrightarrow f\big(\mathrm{Exp}(tX_1+tY_1)\,\big) \qquad \text{as } n\to\infty.\end{aligned}$$

The equality of the two expressions $f\Big(\mathrm{Exp}\big(t(X_1+Y_1)\big)\Big) = \mathrm{Exp}\Big(t(X_2+Y_2)\Big)$ shows that

$$(\mathrm{d}f)|_{\mathbf{e}_1}(X_1+Y_1) = X_2+Y_2,$$

which finishes the proof of linearity of $(\mathrm{d}f)|_{\mathbf{e}_1} : \mathfrak{g}_1 \to \mathfrak{g}_2$.

Finally, we must show that this map $(\mathrm{d}f)|_{\mathbf{e}_1}$ respects Lie brackets. The proof is similar, referring to Lie's formula (II.12) this time, and we leave the details as an exercise. □

The first principle

The first principle says that a Lie group homomorphism is uniquely determined by the associated Lie algebra homomorphism, under a connectedness assumption.

Theorem II.61 (First principle of infinitesimals)**.**
If G_1 is connected, then a continuous group homomorphism $f\colon \mathrm{G}_1 \to \mathrm{G}_2$ is uniquely determined by the Lie algebra homomorphism $\phi = (\mathrm{d}f)|_{\mathbf{e}_1} : \mathfrak{g}_1 \to \mathfrak{g}_2$ given by its differential at the neutral element.

Proof. Choose again a small enough neighborhood $U_1 \subset \mathfrak{g}_1$ of $0 \in \mathfrak{g}_1$ such that $\mathrm{Exp}\colon U_1 \to \mathrm{Exp}(U_1)$ is a homeomorphism. Any group element in the neighborhood $\mathrm{Exp}(U_1) \subset \mathrm{G}_1$ of the neutral element can thus be written as e^{X_1} for some $X_1 \in U_1$. By Exercise II.20, the connected Lie group G_1 is generated by elements in the neighborhood $\mathrm{Exp}(U_1)$, i.e., any $g \in \mathrm{G}_1$ can be written as a finite product

$$g = e^{X_1^{(1)}} \cdots e^{X_1^{(m)}}$$

with some $X_1^{(1)}, \ldots, X_1^{(m)} \in U_1$. On the other hand the zeroth principle implies that we have

$$f\big(e^{X_1^{(j)}}\big) = e^{\phi(X_1^{(j)})} \qquad \text{for } j = 1, \ldots, m.$$

The homomorphism property of f then yields the expression

$$f(g) = f\big(e^{X_1^{(1)}} \cdots e^{X_1^{(m)}}\big) = f\big(e^{X_1^{(1)}}\big) \cdots f\big(e^{X_1^{(m)}}\big) = e^{\phi(X_1^{(1)})} \cdots e^{\phi(X_1^{(m)})}.$$

This expression uniquely recovers the values of f referring only to ϕ. □

An alternative approach to the first principle

To approach the second principle, it is useful to take another point of view to the first principle. Suppose that G_1 and G_2 are (matrix) Lie groups with respective Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, and

$$f \colon \mathrm{G}_1 \to \mathrm{G}_2$$

is a continuous group homomorphism (i.e., a Lie group homomorphism) and

$$\mathrm{d}f|_{\mathbf{e}_1} \colon \mathfrak{g}_1 \to \mathfrak{g}_2.$$

is its differential at the neutral element $\mathbf{e}_1 \in \mathrm{G}_1$. The fruitful general point of view to the alternative approach is that the Lie algebras are the tangent spaces of the Lie groups at their neutral elements, $\mathfrak{g}_1 = T_{\mathbf{e}_1}\mathrm{G}_1$ and $\mathfrak{g}_2 = T_{\mathbf{e}_2}\mathrm{G}_2$. We will also use the left multiplications

$$\begin{aligned} \mathsf{L}_g \colon \mathrm{G}_1 &\to \mathrm{G}_1 \\ h &\mapsto gh \end{aligned} \qquad\qquad \text{for } g \in \mathrm{G}_1$$

and similarly $\mathsf{L}_{g'} \colon \mathrm{G}_2 \to \mathrm{G}_2$ for $g' \in \mathrm{G}_2$. These are smooth maps of the Lie groups to themselves, and in fact diffeomorphisms of the Lie groups: the smooth map $\mathsf{L}_g \colon \mathrm{G}_1 \to \mathrm{G}_1$ has smooth inverse $\mathsf{L}_{g^{-1}} \colon \mathrm{G}_1 \to \mathrm{G}_1$ — the left multiplication with the inverse element. The differentials of the left multiplication maps are linear maps between the appropriate tangent spaces $\mathrm{d}\mathsf{L}_g|_h \colon T_h\mathrm{G}_1 \to T_{gh}\mathrm{G}_1$. Upon specializing to the neutral element $h = \mathbf{e}_1$ and identifying the tangent space $T_{\mathbf{e}_1}\mathrm{G}_1$ with the Lie algebra $\mathfrak{g}_1$, we obtain

$$\mathrm{d}\mathsf{L}_g|_{\mathbf{e}_1} \colon \mathfrak{g}_1 \to T_g\mathrm{G}_1$$

With a similar identification, the differential of $\mathsf{L}_{g^{-1}} \colon \mathrm{G}_1 \to \mathrm{G}_1$ at $g \in \mathrm{G}_1$ becomes

$$\mathrm{d}\mathsf{L}_{g^{-1}}|_g \colon T_g\mathrm{G}_1 \to \mathfrak{g}_1.$$

In the appropriate left-invariant sense, $\mathrm{d}\mathsf{L}_{g^{-1}}|_g$ allows to translate tangent vectors at a generic point $g \in \mathrm{G}_1$ in the group a vectors in the Lie algebra, and $\mathrm{d}\mathsf{L}_g|_{\mathbf{e}_1}$ allows to translate a Lie algebra element into a tangent vector at the generic point in the Lie group. These two linear maps are the inverses of each other; in particular

$$\big(\mathrm{d}\mathsf{L}_g|_{\mathbf{e}_1}\big) \circ \big(\mathrm{d}\mathsf{L}_{g^{-1}}|_g\big) \;=\; \mathrm{id}_{T_g\mathrm{G}_1},$$

follows by differentiating the composition $\mathsf{L}_g \circ \mathsf{L}_{g^{-1}} = \mathrm{id}_{\mathrm{G}_1}$ at g by the chain rule.

For discussing again the first principle, suppose that G_1 is connected, and note that by Fact II.15, then G_1 is in fact smoothly path-connected. Our goal was to reconstruct the Lie group homomorphism f from the associated Lie algebra homomorphism $\phi = \mathrm{d}f|_{\mathbf{e}_1} \colon \mathfrak{g}_1 \to \mathfrak{g}_2$. We of course know the value of the Lie group homomorphism at the neutral element, $f(\mathbf{e}_1) = \mathbf{e}_2 \in \mathrm{G}_2$; the question is how to recover the value $f(g)$ at a generic point $g \in \mathrm{G}_1$. Let us pick a path in G_1 from $\mathbf{e}_1$ to g, more precisely a smooth function[12]

$$\gamma\colon [a,b] \to \mathrm{G}_1 \qquad \text{such that} \qquad \gamma(a) = \mathbf{e}_1 \text{ and } \gamma(b) = g.$$

For $t \in [a,b]$ the derivative $\dot{\gamma}(t)$ is a vector in the tangent space $T_{\gamma(t)}\mathrm{G}_1$. By the above discussion, the appropriate way to map it to the Lie algebra is to consider[13]

$$X(t) = \left(\mathrm{d}\mathsf{L}_{\gamma(t)^{-1}}\big|_{\gamma(t)}\right)\big(\dot{\gamma}(t)\big) \in \mathfrak{g}_1,$$

Applying the inverse map, the derivative of the path is then obtained from the formula[14] $\dot{\gamma}(t) = \big(\mathrm{d}\mathsf{L}_{\gamma(t)}|_{\mathbf{e}_1}\big)(X(t))$ To a smooth path $\gamma\colon [a,b] \to \mathrm{G}_1$, we have thus associated a function

$$\begin{aligned} X\colon [a,b] &\to \mathfrak{g}_1, \\ t &\mapsto X(t), \end{aligned}$$

which keeps track of the direction and velocity of the path. It is appropriate to interpret $X\colon [a,b] \to \mathfrak{g}_1$ as the "steering" of the path γ. One can recover the path γ from its steering X by solving the first order differential equation $\dot{\gamma}(t) = \big(\mathrm{d}\mathsf{L}_{\gamma(t)}|_{\mathbf{e}_1}\big)(X(t))$ with the initial condition $\gamma(a) = \mathbf{e}_1$.[15]

Consider the smooth path $\tilde{\gamma} = (f \circ \gamma)\colon [a,b] \to \mathrm{G}_2$ in G_2. It starts at $\tilde{\gamma}(a) = f(\gamma(a)) = f(\mathbf{e}_1) = \mathbf{e}_2$ and ends at $\tilde{\gamma}(b) = f(\gamma(b)) = f(g)$, the value we seek to reconstruct. The steering of the path $\tilde{\gamma}$ is a function $\tilde{X}\colon [a,b] \to \mathfrak{g}_2$ defined by

$$\tilde{X}(t) = \left(\mathrm{d}\mathsf{L}_{\tilde{\gamma}(t)^{-1}}\big|_{\tilde{\gamma}(t)}\right)\big(\dot{\tilde{\gamma}}(t)\big) \in \mathfrak{g}_2.$$

Not too surprisingly, the Lie algebra homomorphism $\phi = \mathrm{d}f|_{\mathbf{e}_1}\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ relates the steering X of γ to the steering $\tilde{X}$ of $\tilde{\gamma}$.

Lemma II.62. *The steering of $\tilde{\gamma} = f \circ \gamma$ is given by $\tilde{X} = \phi \circ X$, i.e.,*

$$\tilde{X}(t) = \phi\big(X(t)\big) \qquad \text{for } t \in [a,b].$$

Proof. Let us write this simple proof in differential geometric language (where the formulas do not allow for conceptual confusion) and leave it as an optional exercise to the reader to give a more elementary proof in the matrix Lie group case.

The definition of the steering $\tilde{X}$ is via an application of the differential of the left multiplication to the derivative $\dot{\tilde{\gamma}}(t) = \frac{\mathrm{d}}{\mathrm{d}s}\tilde{\gamma}(t+s)|_{s=0}$ of the path $\tilde{\gamma}$, which is by definition the derivative of the composition of the path with the left multiplication. Using this and the

[12] Smoothness of γ instead of just piecewise smoothness can indeed be achieved, and simplifies the discussion here, although it is not essential.

[13] In a matrix Lie group setup, this simply reads

$$X(t) := \frac{\mathrm{d}}{\mathrm{d}s}\big(\gamma(t)^{-1}\gamma(t+s)\big)\big|_{s=0} = \gamma(t)^{-1}\frac{\mathrm{d}}{\mathrm{d}s}\big(\gamma(t+s)\big)\big|_{s=0}.$$

But a small argument would be needed here to show that this expression is indeed in the Lie algebra $\mathfrak{g}_1$. The key is again to use the local homeomorphism provided by the exponential. Note also that this formula (albeit correct) should appear suspicious at a conceptual level: it involves multiplying a Lie group element and a tangent vector; an operation which only makes some sense here because both happen to be square matrices (one from $G_1 \subset \mathrm{GL}_{d_1}(\mathbb{R})$ and the other from a vector space that resembles $\mathfrak{g}_1 \subset \mathbb{R}^{d_1 \times d_1}$).

[14] In matrix Lie group setting one simply has $\dot{\gamma}(t) = \gamma(t)X(t)$; another (correct but) conceptually suspicious formula.

[15] The usual existence and uniqueness results for solutions to differential equations apply, and a solution γ defined on the whole interval $[a,b]$ is obtained.

group homomorphism property of f, and a similar observation about the differential of f and a left multiplication also in G_1, we find

$$\begin{aligned}
\tilde{X}(t) &= \Big(\mathrm{d}\mathsf{L}_{\tilde{\gamma}(t)^{-1}}\big|_{\tilde{\gamma}(t)}\Big)\big(\dot{\tilde{\gamma}}(t)\big) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\Big(\mathsf{L}_{\tilde{\gamma}(t)^{-1}}\big(\tilde{\gamma}(t+s)\big)\Big)\Big|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\Big(f(\gamma(t))^{-1}\, f(\gamma(t+s))\Big)\Big|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\Big(f\big(\gamma(t)^{-1}\,\gamma(t+s)\big)\Big)\Big|_{s=0} \\
&= \big(\mathrm{d}f|_{\mathbf{e}_1}\big)\Big(\frac{\mathrm{d}}{\mathrm{d}s}\big(\gamma(t)^{-1}\,\gamma(t+s)\big)\big|_{s=0}\Big) \\
&= \big(\mathrm{d}f|_{\mathbf{e}_1}\big)\Big((\mathrm{d}\mathsf{L}_{\gamma(t)^{-1}}|_{\gamma(t)})(\dot{\gamma}(t))\Big) \\
&= \phi\big(X(t)\big).
\end{aligned}$$

□

We can therefore recover the path $\tilde{\gamma}$ by solving the first order differential equation

$$\dot{\tilde{\gamma}}(t) = \big(\mathrm{d}\mathsf{L}_{\tilde{\gamma}(t)}|_{\mathbf{e}_2}\big)\big(\phi(X(t))\big)$$

with the initial condition $\tilde{\gamma}(a) = \mathbf{e}_2$. This differential equation above only involves the Lie algebra homomorphism $\phi = \mathrm{d}f|_{\mathbf{e}_1} \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ and the known steering X of the chosen path γ from $\mathbf{e}_1$ to g. We thus re-establish the first principle: under the assumption of connectedness of G_1, we are able to recover the value $f(g) = f(\gamma(b)) = \tilde{\gamma}(b)$ of the Lie group homomorphism f using only its differential $\phi = \mathrm{d}f|_{\mathbf{e}_1}$ at the neutral element.

The second principle

The zeroth principle says that a Lie group homomorphism always gives rise to a Lie algebra homomorphism. The other two principles concern topological obstructions to having a converse statement. The first principle says that disconnectedness is the only possible obstruction to a Lie algebra homomorphism uniquely determining the Lie group homomorphism. The second principle says that lack of simply connectedness (i.e., a non-trivial fundamental group) is the only possible obstruction to lifting a Lie algebra homomorphism to a Lie group homomorphism. More precisely, it states the following.

Theorem II.63 (Second principle of infinitesimals)**.**
If G_1 is connected and simply connected, then for any Lie algebra homomorphism $\phi \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ there exists a Lie group homomorphism $f \colon \mathrm{G}_1 \to \mathrm{G}_2$ such that $\phi = (\mathrm{d}f)|_{\mathbf{e}_1}$.

The idea is rather simple in view of the alternative approach to the the first principle that we just presented. We seek to construct a function f on G_1, i.e., to define values $f(g)$ at arbitrary $g \in \mathrm{G}_1$, given only the Lie algebra homomorphism ϕ. Since G_1 is by assumption connected, we can choose a path $\gamma \colon [0,1] \to \mathrm{G}_1$ such that $\gamma(0) = \mathbf{e}_1$ and $\gamma(1) = g$. Then we calculate the steering $X \colon [0,1] \to \mathfrak{g}_1$ of γ, compose with the Lie algebra homomorphism ϕ to get a function $\tilde{X} = \phi \circ X \colon [0,1] \to \mathfrak{g}_2$, and solve a differential equation to define $\tilde{\gamma} \colon [0,1] \to \mathrm{G}_2$ using $\tilde{X}$ as steering. We would like to define $f(g) = \tilde{\gamma}(b)$ (and then of course check the required group homomorphism property, continuity, and desired differential at the neutral element). The key question is whether this gives rise to a well-defined function: if we had chosen a *different* path γ from $\mathbf{e}_1$ to g, would we have obtained the same proposal for the value $f(g)$? Assuming G_1 is simply connected will save the day: any two paths can be continuously (in fact smoothly) deformed into one another, and one can check that during the deformation, the proposed value remains unchanged.

We omit the details for now.

2.6. Adjoint representations

We now move to our first examples of Lie group homomorphisms and associated Lie algebra homomorphisms. The notion of adjoint representations provides examples of such homomorphisms for a general Lie group.

Suppose again that G is a matrix Lie group and let $\mathfrak{g}$ be its Lie algebra. Given any $g \in \mathrm{G}$, the conjugation by g

$$C_g\colon\ G \to G \qquad\qquad C_g(h) = ghg^{-1} \tag{II.24}$$

is clearly a group homomorphism, and moreover continuous.

Definition II.64 (Adjoint action of a Lie group on its Lie algebra)**.**
For any $g \in \mathrm{G}$, we denote $\mathrm{Ad}_g := (\mathrm{d}C_g)|_{\mathbf{e}}$ and call it the **adjoint action** of the element $g \in G$ on $\mathfrak{g}$. By the zeroth principle, the differential $(\mathrm{d}C_g)|_{\mathbf{e}}$ exists and defines a Lie algebra homomorphism

$$\mathrm{Ad}_g\colon \mathfrak{g} \to \mathfrak{g}.$$

Lemma II.65 (Explicit adjoint action for a matrix Lie group)**.**
For a matrix Lie group G *and an element* $g \in \mathrm{G}$*, the adjoint action is explicitly*

$$\mathrm{Ad}_g(X) = gXg^{-1} \qquad \text{for } X \in \mathfrak{g}. \tag{II.25}$$

Proof. By definition the adjoint action $\mathrm{Ad}_g := (\mathrm{d}C_g)|_{\mathbf{e}}$ is the differential of conjugation $C_g\colon \mathrm{G} \to \mathrm{G}$ at the neutral element, which is

$$\mathrm{Ad}_g(X) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} C_g(e^{tX}) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Big(g\, e^{tX}\, g^{-1}\Big) = gXg^{-1},$$

where the derivative was calculated using the series defining the exponential. □

Remark II.66 (Adjoint action is not conjugation)**.**
Although C_g and Ad_g in (II.24) and (II.25) look formally identical, they must not be confused! In (II.24), $h \in \mathrm{G}$ is a Lie group element, whereas in (II.25), $X \in \mathfrak{g}$ is a Lie algebra element. The same difference goes for $C_g(h) \in \mathrm{G}$ and $\mathrm{Ad}_g(X) \in \mathfrak{g}$. To reiterate, conjugation is an operation on the group, whereas adjoint action is an operation on the Lie algebra.

Remark II.67 (Adjoint action is indeed a Lie algebra homomorphism)**.**
As argued above, it follows from the zeroth principle that $\mathrm{Ad}_g\colon \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra homomorphism. In view of the explicit formula (II.25), this is also easy to see directly: for any $X, Y \in \mathfrak{g}$ the bracket of $\mathrm{Ad}_g(X)$ and $\mathrm{Ad}_g(Y)$ (defined as the commutator in $\mathfrak{g} = \mathcal{L}(\mathrm{G}) \subset$

$\mathbb{R}^{n\times n}$) reads

$$\begin{aligned}[\mathrm{Ad}_g(X), \mathrm{Ad}_g(Y)] &= \mathrm{Ad}_g(X)\mathrm{Ad}_g(Y) - \mathrm{Ad}_g(Y)\mathrm{Ad}_g(X) \\ &= gXg^{-1}gYg^{-1} - gYg^{-1}gXg^{-1} \\ &= g(XY - YX)g^{-1} \\ &= \mathrm{Ad}_g([X,Y]).\end{aligned}$$

Note that for $g_1, g_2 \in \mathrm{G}$, we have[16]

$$\mathrm{Ad}_{g_1 g_2} = \mathrm{Ad}_{g_1} \circ \mathrm{Ad}_{g_2},$$

and that $\mathrm{Ad}_{\mathbf{e}} = \mathrm{id}_{\mathfrak{g}}$. By thinking of g as the variable, the function $g \mapsto \mathrm{Ad}_g$ becomes a group homomorphism from G to the group $\mathrm{Aut}(\mathfrak{g})$ of invertible linear maps of $\mathfrak{g}$ to itself (the inverse of Ad_g is $\mathrm{Ad}_{g^{-1}}$), This homomorphism is also continuous, and therefore $\mathrm{Ad}_\bullet \colon \mathrm{G} \to \mathrm{Aut}(\mathfrak{g})$ is a Lie group homomorphism, and equips the vector space $\mathfrak{g}$ with the structure of a representation of the Lie group G.

Definition II.68 (Adjoint representation of a Lie group)**.**
The **adjoint representation of a Lie group** G is the representation

$$\mathrm{Ad}_\bullet \colon\ \mathrm{G} \to \mathrm{Aut}(\mathfrak{g}) \qquad\qquad g \mapsto \mathrm{Ad}_g = (\mathrm{d}C_g)|_{\mathbf{e}} \in \mathrm{Aut}(\mathfrak{g})$$

on the vector space $\mathfrak{g} = \mathcal{L}(\mathrm{G})$, the Lie algebra of G.

Furthermore, since $\mathrm{Ad}_\bullet \colon \mathrm{G} \to \mathrm{Aut}(\mathfrak{g})$ is a Lie group homomorphism, we can again look at the homomorphism of Lie algebras that it induces by the zeroth principle. The Lie algebra of G is of course $\mathfrak{g}$, and the Lie algebra of the Lie group

$$\mathrm{Aut}(\mathfrak{g}) \cong \mathrm{GL}_{\dim(\mathfrak{g})}(\mathbb{R})$$

is (as in Example II.38) the space of all linear maps from $\mathfrak{g}$ to itself,

$$\mathrm{End}(\mathfrak{g}) \cong \mathfrak{gl}_{\dim(\mathfrak{g})}(\mathbb{R}).$$

Therefore by the zeroth principle, the differential of $g \mapsto \mathrm{Ad}_g$ at the neutral element $\mathbf{e} \in \mathrm{G}$ is a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathrm{End}(\mathfrak{g})$. Since the homomorphism takes value in a Lie algebra of linear maps, this is a Lie algebra representation.

Definition II.69 (Adjoint representation of a Lie algebra of a Lie group)**.**
The **adjoint representation of the Lie algebra** $\mathfrak{g}$ of a Lie group G is the Lie algebra homomorphism $\mathrm{ad} := (\mathrm{d}\,\mathrm{Ad}_\bullet)|_{\mathbf{e}}$,

$$\mathrm{ad} \colon\ \mathfrak{g} \to \mathrm{End}(\mathfrak{g})\,.$$

We denote the value at $X \in \mathfrak{g}$ of the above homomorphism by $\mathrm{ad}_X \in \mathrm{End}(\mathfrak{g})$. This is itself a linear map $\mathrm{ad}_X \colon \mathfrak{g} \to \mathfrak{g}$. Then as usual, when this linear map is evaluated at $Z \in \mathfrak{g}$, the value is denoted by $\mathrm{ad}_X(Z) \in \mathfrak{g}$. For Lie algebras of matrix Lie groups, we easily obtain the following concrete formula.

[16]For matrix Lie groups the homomorphism property $\mathrm{Ad}_{g_1g_2} = \mathrm{Ad}_{g_1} \circ \mathrm{Ad}_{g_2}$ can be directly seen from $\mathrm{Ad}_g(X) = gXg^{-1}$. For general Lie groups, it is obtained by noticing that conjugation is a group action, $C_{g_1g_2} = C_{g_1} \circ C_{g_2}$, so the chain rule of differentiation (keeping also in mind that $C_{g_2}(\mathbf{e}) = \mathbf{e}$) implies

$$(\mathrm{d}C_{g_1g_2})|_{\mathbf{e}} = (\mathrm{d}C_{g_1})|_{\mathbf{e}} \circ (\mathrm{d}C_{g_2})|_{\mathbf{e}}.$$

Lemma II.70 (Explicit adjoint action for a matrix Lie algebra)**.**
The adjoint representation of its Lie algebra $\mathfrak{g}$ *of a matrix Lie group* G *is explicitly given by*

$$\mathrm{ad}_X(Z) = [X, Z] \qquad \text{for } X, Z \in \mathfrak{g}. \tag{II.26}$$

Proof. By definition, $\mathrm{ad} := (\mathrm{d}\,\mathrm{Ad}_\bullet)|_\mathsf{e}$ is the differential of the adjoint reprensetation $g \mapsto \mathrm{Ad}_g$ at the neutral element, which is

$$\mathrm{ad}_X(Z) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \mathrm{Ad}_{e^{tX}}(Z) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Big(e^{tX} Z e^{-tX}\Big) = XZ - ZX = [X, Z],$$

where the derivative was again calculated using the series defining the exponential. □

Remark II.71 (Adjoint representation is indeed a Lie algebra homomorphism)**.**
As argued above, it follows from the zeroth principle that $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra homomorphism. In view of the explicit formula (II.26) relating the adjoint action to the Lie bracket, this is also easy to see directly. Indeed, for any $X, Y \in \mathfrak{g}$ the bracket of ad_X and ad_Y (defined as the commutator in $\mathrm{End}(\mathfrak{g})$) is the linear map $\mathfrak{g} \to \mathfrak{g}$, whose value at $Z \in \mathfrak{g}$ is

$$\begin{aligned} \big[\mathrm{ad}_X, \mathrm{ad}_Y\big]_{\mathrm{End}(\mathfrak{g})}(Z) &= \mathrm{ad}_X\big(\mathrm{ad}_Y(Z)\big) - \mathrm{ad}_Y\big(\mathrm{ad}_X(Z)\big) \\ &= \big[X, [Y, Z]\big] - \big[Y, [X, Z]\big]. \end{aligned}$$

If we use antisymmetry of brackets once, then Jacobi identity, and then antisymmetry once more, the above becomes

$$\begin{aligned} \big[\mathrm{ad}_X, \mathrm{ad}_Y\big]_{\mathrm{End}(\mathfrak{g})}(Z) &= \big[X, [Y, Z]\big] - \big[Y, [X, Z]\big] \\ &= \big[X, [Y, Z]\big] + \big[Y, [Z, X]\big] \\ &= -\big[Z, [X, Y]\big] \\ &= \big[[X, Y], Z\big] \\ &= \mathrm{ad}_{[X,Y]}(Z). \end{aligned}$$

This explicitly shows the Lie algebra homomorphism property of ad.

3. The Lie groups SU_2 and SO_3 and their Lie algebras

3.1. The topology of the Lie group SU_2

Recall that the group of unitary 2×2 matrices is

$$\mathrm{SU}_2 = \left\{ g \in \mathbb{C}^{2\times 2} \;\Big|\; g^\dagger\, g = \mathbb{I}_2,\ \det(g) = 1 \right\}.$$

Note that this is a real Lie group, although the matrix entries of $g \in \mathrm{SU}_2$ are complex. Below we typically denote the entries of a 2×2 complex matrix

$$g = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}.$$

We start by describing the topology of SU_2. From Exercise II.44 we know that the dimension of its Lie algebra $\mathfrak{su}_2$ is three, so the matrix Lie group SU_2 is a three-dimensional manifold. We will in fact show that SU_2 is homeomorphic to the unit sphere in the four-dimensional Euclidean space.

Theorem II.72 (Topology of SU_2)**.**
The matrix Lie group SU_2 *is homeomorphic to the three-sphere*

$$\mathcal{S}^3 \;=\; \Big\{ (x_1,x_2,x_3,x_4) \in \mathbb{R}^4 \;\Big|\; x_1^2+x_2^2+x_3^2+x_4^2 = 1 \Big\}.$$

In particular, SU_2 *is compact, connected, and simply connected.*

Proof. Write a 2×2 complex matrix g as

$$g = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}, \qquad \text{where } \alpha,\beta,\gamma,\delta \in \mathbb{C}.$$

By definition we have $g \in SU_2$ if and only if $g^\dagger g = \mathbb{I}$ and $\det(g) = 1$. We will start by finding a characterization of these conditions directly in terms of the matrix entries $\alpha,\beta,\gamma,\delta$.

Calculate

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

In view of this calculation, the condition

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

amounts to a system of four equalities of complex numbers; one for each entry of the above 2×2 matrices. The equality of $(1,1)$-entries, e.g., reads $|\alpha|^2+|\gamma|^2 = 1$, and it amounts to requiring that $[\alpha\ \gamma]^\top$ is a unit vector in $\mathbb{C}^2$ with respect to the norm coming from the standard inner product of vectors in $\mathbb{C}^2$. Similarly, the equality of $(2,2)$-entries says that $[\beta\ \delta]^\top$ is a unit vector. The equality of $(1,2)$-entries and the equality of $(2,1)$-entries are in fact equivalent: these two equations are complex conjugates of each other. The former explicitly reads $\overline{\alpha}\beta+\overline{\gamma}\delta = 0$, which amounts to vanishing of the inner product of the vectors $[\alpha\ \gamma]^\top$ and $[\beta\ \delta]^\top$, i.e., to the requirement that these vectors are orthogonal. The orthogonal complement of the unit vector $[\alpha\ \gamma]^\top$ is one-dimensional, and spanned by the vector $[-\overline{\gamma}\ \overline{\alpha}]^\top$, so the requirement becomes $[\beta\ \delta]^\top = s\,[-\overline{\gamma}\ \overline{\alpha}]^\top$ for some $s \in \mathbb{C}$.

If $\beta = -s\overline{\gamma}$ and $\delta = s\overline{\alpha}$ and $|\alpha|^2+|\gamma|^2 = 1$ as above, then the determinant of g becomes $\det(g) = \alpha\delta - \beta\gamma = s|\alpha|^2 + s|\gamma|^2$. The condition $\det(g) = 1$ then requires $s = 1$, i.e., $\beta = -\overline{\gamma}$ and $\delta = \overline{\alpha}$.

We thus see that $g \in SU_2$ if and only if g is of the form

$$g = \begin{bmatrix} \alpha & -\overline{\gamma} \\ \gamma & \overline{\alpha} \end{bmatrix}$$

for $\alpha,\gamma \in \mathbb{C}$ such that $|\alpha|^2+|\gamma|^2 = 1$.

We can therefore define a function

$$\mathcal{S}^3 \longrightarrow SU_2 \qquad \text{by} \qquad (x_1,x_2,x_3,x_4) \mapsto \begin{bmatrix} x_1+\mathbb{i}\,x_2 & -x_3+\mathbb{i}\,x_4 \\ x_3+\mathbb{i}\,x_4 & x_1-\mathbb{i}x_2 \end{bmatrix},$$

since the property $x_1^2+x_2^2+x_3^2+x_4^2 = 1$ in $\mathcal{S}^3$ guarantees $|x_1+\mathbb{i}x_2|^2+|x_3+\mathbb{i}x_4|^2 = 1$. This function is clearly continuous: its components are first order polynomials. It is obviously invertible, with inverse

$$SU_2 \longrightarrow \mathcal{S}^3 \qquad \text{given by} \qquad \begin{bmatrix} \alpha & -\overline{\gamma} \\ \gamma & \overline{\alpha} \end{bmatrix} \mapsto \big(\mathfrak{Re}(\alpha),\, \mathfrak{Im}(\alpha),\, \mathfrak{Re}(\gamma),\, \mathfrak{Im}(\gamma)\big).$$

The inverse is also continuous: its components are projections (real and imaginary parts) of the matrix entries. We have thus explicitly exhibited a homeomorphism $SU_2 \to \mathcal{S}^3$ (in fact a diffeomorphism, i.e., smooth bijection with smooth inverse).

The unit sphere $\mathcal{S}^3 \subset \mathbb{R}^4$ is bounded and closed, thus compact. Therefore also SU_2 is compact. Also, $\mathcal{S}^3 \subset \mathbb{R}^4$ is connected and simply connected, $\pi_1(\mathcal{S}^3) = e$. Therefore SU_2 is connected and simply connected, too. $\square$

3.2. Isomorphism of the Lie algebras $\mathfrak{su}_2$ and $\mathfrak{so}_3$

We claim that the Lie algebras of the Lie groups SU_2 and SO_3 are isomorphic. To this end, let us describe both by giving explicit choices of basis for them, and by calculating the Lie brackets in these chosen bases.

A concrete description of the Lie algebra $\mathfrak{so}_3$

Let us start by concretely describing the Lie algebra $\mathfrak{so}_3$ of the Lie group SO_3 of rotations in the three-dimensional Euclidean space $\mathbb{R}^3$.

Example II.73 (Basis of the Lie algebra $\mathfrak{so}_3$)**.**
Example II.43 explicitly characterizes the Lie algebra $\mathfrak{so}_n = \mathcal{L}(\mathrm{SO}_n)$ of the special orthogonal group SO_n in n dimensions as the set of real antisymmetric $n \times n$-matrices. In the special case $n = 3$, this becomes

$$\mathfrak{so}_3 = \Big\{ X \in \mathbb{R}^{3\times 3} \;\Big|\; X^\top = -X \Big\}. \tag{II.27}$$

The dimension of this Lie algebra is $\dim(\mathfrak{so}_3) = \frac{n(n-1)}{2} = 3$ — one can freely choose the strictly upper triangular entries, and the rest of the matrix is then fully determined. One explicit basis thus consists of the three matrices[17]

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

To completely describe the Lie algebra, we will write down the Lie brackets among the above basis elements. By antisymmetry we of course have

$$[R_x, R_x] = 0, \qquad [R_y, R_y] = 0, \qquad [R_z, R_z] = 0,$$

and it suffices to find $[R_x, R_y] = -[R_y, R_x]$, $[R_y, R_z] = -[R_z, R_y]$, and $[R_z, R_x] = -[R_x, R_z]$. By explicit calculation we find, e.g.,

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

Similar calculations yield $[R_y, R_z] = R_x$, and $[R_z, R_x] = R_y$. In summary, the information about the Lie brackets of $\mathfrak{so}_3$ in the basis R_x, R_y, R_z is contained in the formulas

$$[R_x, R_y] = R_z, \qquad [R_y, R_z] = R_x, \qquad [R_z, R_x] = R_y. \tag{II.28}$$

Pauli spin matrices

To concretely describe the Lie algebra $\mathfrak{su}_2$ of SU_2, it is convenient and conventional to introduce the following **Pauli spin matrices** $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{C}^{2\times 2}$:

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_2 = \begin{bmatrix} 0 & -\mathbb{i} \\ \mathbb{i} & 0 \end{bmatrix}, \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{II.29}$$

[17] The explanation for the choices of labels and signs is that R_x, R_y, and R_z are the infinitesimal generators of positively oriented rotations along the x, y, and z-axes.

Let us record some key properties of these in the lemma below.

Lemma II.74 (Properties of the Pauli spin matrices).

(a) *The Pauli spin matrices are Hermitian: for each* $j \in \{1,2,3\}$ *we have*

$$\sigma_j^\dagger = \sigma_j.$$

(b) *The four matrices* $\mathbb{I}, \sigma_1, \sigma_2, \sigma_3 \in \mathbb{C}^{2\times 2}$ *form a basis of the real vector space of* 2×2 *Hermitian matrices.*

(c) *The Pauli spin matrices are involutive: for each* $j \in \{1,2,3\}$ *we have*

$$\sigma_j^2 = \mathbb{I}.$$

(d) *The products of any two different Pauli spin matrices are proportional to the third, with proportionality constants* $\pm\mathbb{i}$*; specifically we have*

$$\sigma_1\sigma_2 = \mathbb{i}\,\sigma_3 = -\sigma_2\sigma_1, \qquad \sigma_2\sigma_3 = \mathbb{i}\,\sigma_1 = -\sigma_3\sigma_2, \qquad \sigma_3\sigma_1 = \mathbb{i}\,\sigma_2 = -\sigma_1\sigma_3.$$

Proof. Each part is proven by a simple calculation. □

A concrete description of the Lie algebra $\mathfrak{su}_2$

The Pauli spin matrices are used below to obtain a convenient choice of basis for the Lie algebra $\mathfrak{su}_2$ of the Lie group SU_2.

Example II.75 (Basis of the Lie algebra $\mathfrak{su}_2$).

Exercise II.44, and equation (II.19) in particular, explicitly characterizes the Lie algebra $\mathfrak{su}_n = \mathcal{L}(SU_n)$ of the special unitary group SO_n in n dimensions as the set of traceless anti-Hermitian $n\times n$-matrices (with complex entires). In the special case $n = 2$, this becomes

$$\mathfrak{su}_2 = \left\{ X \in \mathbb{C}^{2\times 2} \;\middle|\; X^\dagger = -X \right\}. \tag{II.30}$$

The dimension of this real Lie algebra is $\dim(\mathfrak{su}_2) = n^2 - 1 = 3$ — one can freely choose the real and imaginary parts of the strictly upper triangular entry and the imaginary part of one of the diagonal entries, and the rest of the matrix is then fully determined. The Pauli spin matrices are a basis of traceless Hermitian 2×2-matrices, so the following imaginary multiples of them form one explicit basis of $\mathfrak{su}_2$:

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

To completely describe the Lie algebra, we will write down the Lie brackets among the above basis elements. By antisymmetry, it suffices to provide the following Lie brackets, whose calculation is straightforward by Lemma II.74(c):

$$\begin{aligned}
[S_x, S_y] &= S_x S_y - S_y S_x = \frac{-1}{4}(\sigma_1\sigma_2 - \sigma_2\sigma_1) = \frac{-\mathbb{i}}{2}\sigma_3 = S_z \\
[S_y, S_z] &= S_y S_z - S_z S_y = \frac{-1}{4}(\sigma_2\sigma_3 - \sigma_3\sigma_2) = \frac{-\mathbb{i}}{2}\sigma_1 = S_x \\
[S_z, S_x] &= S_z S_x - S_x S_z = \frac{-1}{4}(\sigma_3\sigma_1 - \sigma_1\sigma_3) = \frac{-\mathbb{i}}{2}\sigma_2 = S_y.
\end{aligned}$$

The Lie brackets of $\mathfrak{su}_2$ in the basis S_x, S_y, S_z thus have exactly the same form as the Lie brackets of $\mathfrak{so}_3$ in the basis R_x, R_y, R_z of Example II.73. This means that there exists a Lie algebra isomorphism

$$\mathfrak{su}_2 \cong \mathfrak{so}_3 \tag{II.31}$$

explicitly given by

$$c_x S_x + c_y S_y + c_z S_z \;\mapsto\; c_x R_x + c_y R_y + c_z R_z$$

for $c_x, c_y, c_z \in \mathbb{R}$.

3.3. The adjoint representation of SU_2

The isomorphism (II.31) may look coincidental, but we will next explain that the isomorphism actually arises from the adjoint representation of SU_2 on its three-dimensional Lie algebra $\mathfrak{su}_2$. The key observation is that this three dimensional Lie algebra $\mathfrak{su}_2$ has a natural inner product, allowing to identify it with the Euclidean space $\mathbb{R}^3$, and that the adjoint action of SU_2 leaves the inner products invariant, so the action is essentially by (special) orthogonal transformations.

The inner product on $\mathfrak{su}_2$ and an orthonormal basis

Invariant inner products on $\mathfrak{u}_n$ will be considered in detail and for general $n \in \mathbb{N}$ in Example III.62, and the subspaces $\mathfrak{su}_n \subset \mathfrak{u}_n$ of course inherit such inner products. From the general result it would follow as a particular case that the formula

$$(X, Y) = -2\,\mathrm{tr}(XY) \qquad \text{for } X, Y \in \mathfrak{su}_2 \tag{II.32}$$

defines an inner product, i.e., a positive-definite symmetric bilinear form on $\mathfrak{su}_2$. But in this particular case it is instructive to take a hands-on approach. Bilinearity of (II.32) is obvious and symmetricity follows from the cyclicity of the trace, $\mathrm{tr}(XY) = \mathrm{tr}(YX)$. Positive-definiteness follows by concretely checking with Lemma II.74 that the basis S_x, S_y, S_z is orthonormal as follows. We have, e.g., $S_x S_y = -\frac{1}{4}\,\sigma_1\sigma_2 = -\frac{\mathbb{i}}{4}\,\sigma_3$, so by tracelessness of the Pauli matrix σ_3 we find the orthogonality relation

$$(S_x, S_y) = -2\,\mathrm{tr}(S_x S_y) = \frac{1}{2}\,\mathrm{tr}(\sigma_3) = 0.$$

Similarly one gets $(S_y, S_z) = (S_z, S_y) = 0$, showing that the basis S_x, S_y, S_z is orthogonal. Also, we have $S_x^2 = -\frac{1}{4}\,\sigma_1^2 = -\frac{1}{4}\,\mathbb{I}$, so we find the normalization

$$(S_x, S_x) = -2\,\mathrm{tr}(S_x^2) = \frac{1}{2}\,\mathrm{tr}(\mathbb{I}) = 1.$$

Similarly one gets $(S_y, S_y) = (S_z, S_z) = 1$, allowing to conclude that the basis S_x, S_y, S_z is orthonormal, and that (II.32) is positive-definite.

Identification of $\mathfrak{su}_2$ with the Euclidean three-space

In view of the orthonormality calculation above, we may identify the three-dimensional inner product space $\mathfrak{su}_2$ with the Euclidean space $\mathbb{R}^3$, with $S_x, S_y, S_z \in \mathfrak{su}_2$ in the role of the standard basis. We may in particular consider the group of orthogonal transformations that preserve the inner product,

$$\mathrm{O}(\mathfrak{su}_2) \;=\; \Big\{ A \in \mathrm{Aut}(\mathfrak{su}_2) \;\Big|\; \forall X, Y \in \mathfrak{su}_2 : (AX, AY) = (X, Y) \Big\} \;\cong\; \mathrm{O}_3$$

and its subgroup of

$$\mathrm{SO}(\mathfrak{su}_2) \;=\; \big\{ A \in \mathrm{O}(\mathfrak{su}_2) \;\big|\; \det(A) = 1 \big\} \;\cong\; \mathrm{SO}_3.$$

The adjoint action

The adjoint representation

$$\mathrm{Ad}\colon \mathrm{SU}_2 \to \mathrm{Aut}(\mathfrak{su}_2) \tag{II.33}$$

is, according Lemma II.65, explicitly given by

$$\mathrm{Ad}_g(X) \; = gXg^{-1} \qquad \text{for } g \in \mathrm{SU}_2,\ X \in \mathfrak{su}_2.$$

For any $g \in \mathrm{SU}_2$, the automorphism $\mathrm{Ad}_g \in \mathrm{Aut}(\mathfrak{su}_2)$ preserves inner products: if $X, Y \in \mathfrak{su}_2$, then

$$\begin{aligned}
(\mathrm{Ad}_g(X), \mathrm{Ad}_g(Y)) \;&=\; -2\,\mathrm{tr}\big(\mathrm{Ad}_g(X)\,\mathrm{Ad}_g(Y)\big) \\
&=\; -2\,\mathrm{tr}\big(gXg^{-1}gYg^{-1}\big) \\
&=\; -2\,\mathrm{tr}\big(g^{-1}gXg^{-1}gY\big) \qquad\qquad \text{(by cyclicity of trace)} \\
&=\; -2\,\mathrm{tr}\big(XY\big) \\
&=\; (X, Y)\,.
\end{aligned}$$

In other words, we have $\mathrm{Ad}_g \in \mathrm{O}(\mathfrak{su}_2)$. So the adjoint representation (II.33) can be seen as a Lie group homomorphism

$$\mathrm{Ad}_\bullet\colon \mathrm{SU}_2 \to \mathrm{O}(\mathfrak{su}_2).$$

But by connectedness of SU_2 and continuity of $\mathrm{Ad}_\bullet\colon \mathrm{SU}_2 \to \mathrm{O}(\mathfrak{su}_2)$, the image must lie in the connected component $\mathrm{SO}(\mathfrak{su}_2)$ of $\mathrm{id}_{\mathfrak{su}_2} \in \mathrm{O}(\mathfrak{su}_2)$, so we can moreover view the adjoint representation (II.33) as a Lie group homomorphism

$$\mathrm{Ad}_\bullet\colon \mathrm{SU}_2 \to \mathrm{SO}(\mathfrak{su}_2).$$

Identifying $\mathrm{SO}(\mathfrak{su}_2) \cong \mathrm{SO}_3$, the adjoint representation thus gives a Lie group homomorphism

$$f\colon \mathrm{SU}_2 \to \mathrm{SO}_3. \tag{II.34}$$

The differential of this Lie group homomorphism at the neutral element $e \in \mathrm{SU}_2$ gives a homomorphism of the corresponding Lie algebras

$$\mathrm{d}f|_e\colon \mathfrak{su}_2 \to \mathfrak{so}_3.$$

An explicit calculation of

$$\mathrm{d}f|_e(S_x) = \frac{\mathrm{d}}{\mathrm{d}t} f(e^{tS_x})\Big|_{t=0}, \quad \mathrm{d}f|_e(S_y) = \frac{\mathrm{d}}{\mathrm{d}t} f(e^{tS_y})\Big|_{t=0}, \quad \mathrm{d}f|_e(S_z) = \frac{\mathrm{d}}{\mathrm{d}t} f(e^{tS_z})\Big|_{t=0},$$

shows that this Lie algebra homomorphism is in fact exactly the isomorphism (II.31). Or we may use Lemma II.70, which states

$$\big(\mathrm{dAd}_\bullet|_e(X)\big)(Y) \;=\; [X, Y],$$

and thus yields, for example,

$$\begin{aligned}
\big(\mathrm{dAd}_\bullet|_e(S_x)\big)(S_x) \;&= [S_x, S_x] \;=\; 0, \\
\big(\mathrm{dAd}_\bullet|_e(S_x)\big)(S_y) \;&= [S_x, S_y] \;=\; S_z, \\
\big(\mathrm{dAd}_\bullet|_e(S_x)\big)(S_z) \;&= [S_x, S_z] \;=\; -S_y,
\end{aligned}$$

showing that the matrix of the linear map $\mathrm{dAd}_\bullet|_e(S_x)\colon \mathfrak{su}_2 \to \mathfrak{su}_2$ in the basis S_x, S_y, S_z is

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = R_x$$

Similar calculations show that in the the matrices of $\mathrm{dAd}_\bullet|_e(S_y)\colon \mathfrak{su}_2 \to \mathfrak{su}_2$ and $\mathrm{dAd}_\bullet|_e(S_z)\colon \mathfrak{su}_2 \to \mathfrak{su}_2$ in the same basis are R_y and R_z.

3.4. The topology of the Lie group SO_3

Above we saw that the adjoint representation of SU_2 can be reinterpreted as a Lie group homomorphism $f\colon \mathrm{SU}_2 \to \mathrm{SO}_3$ by equipping the three-dimensional vector space $\mathfrak{su}_2$ with an inner product and thus identifying it with $\mathbb{R}^3$. We already understand the topology of SU_2 by Theorem II.72, and this mapping f will soon provide a concrete way to understand the topology of SO_3. We start with some concrete observations about SO_3, however.

Auxiliary results about rotations

The following lemma amounts to the intuitive statement that every element g of the "rotation group of the three-dimensional space", SO_3, indeed has an axis which the rotation which it fixes pointwise.

Lemma II.76 (The existence of an axis of rotation)**.**
Any $g \in \mathrm{SO}_3$ has an eigenvector with eigenvalue $+1$. In particular, there exists a unit vector $\vec{\xi} \in \mathbb{R}^3$ such that $g\,\vec{\xi} = \vec{\xi}$.

Proof. Fix $g \in \mathrm{SO}_3$. Note first that the only possible real eigenvalues of g are ± 1: if $\lambda \in \mathbb{R}$ is an eigenvalue and $\vec{v} \in \mathbb{R}^3$ is a corresponding eigenvector, then since $g^\top g = \mathbb{I}$, we find

$$\|\vec{v}\|^2 = \vec{v}^\top \vec{v} = \vec{v}^\top\, \mathbb{I}\, \vec{v} = \vec{v}^\top\, (g^\top g)\, \vec{v} = (g\vec{v})^\top g\vec{v} = \lambda^2\, \vec{v}^\top \vec{v} = \lambda^2\, \|\vec{v}\|^2.$$

In view of $\|\vec{v}\| \neq 0$, this implies $\lambda^2 = 1$ and therefore $\lambda \in \{-1, +1\}$.

Consider the characteristic polynomial $p(t) = \det(t\,\mathbb{I}_3 - g)$ of g. The leading coefficient is t^3, so we have $p(t) \to +\infty$ as $t \to +\infty$. On the other hand, $p(0) = \det(-g) = (-1)^3 \det(g) = -1$ since $\det(g) = 1$. By the intermediate value theorem, there exists a $\lambda \in (0, +\infty)$ such that $p(\lambda) = 0$, i.e., that $\lambda > 0$ is an eigenvalue of g. But the only possible real positive eigenvalue of g is $\lambda = +1$, so the assertion is proven. $\square$

We can in fact describe quite explicitly the rotations which have the axis $\mathbb{R}\vec{\xi} \subset \mathbb{R}^3$ as their axis of rotation. It is instructive to start thinking about the case $\vec{\xi} = \vec{e}_z = [0\ 0\ 1]^\top$ of rotations around the z-axis $\mathbb{R}\vec{e}_z \subset \mathbb{R}^3$. The rotation by angle $\theta \in \mathbb{R}$ is described by the matrix

$$\mathrm{Exp}\big(\theta R_z\big) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where R_z is as in Example II.73. More generally, to describe rotations around the axis $\mathbb{R}\vec{\xi} \subset \mathbb{R}^3$ spanned by a unit vector

$$\vec{\xi} = \begin{bmatrix} \xi_x \\ \xi_y \\ \xi_z \end{bmatrix},$$

let us form the Lie algebra element $R = \xi_x R_x + \xi_y R_y + \xi_z R_z \in \mathfrak{so}_3$, i.e., the following antisymmetric matrix

$$R = \begin{bmatrix} 0 & -\xi_z & \xi_y \\ \xi_z & 0 & -\xi_x \\ -\xi_y & \xi_x & 0 \end{bmatrix}.$$

Given another vector $\vec{\eta} = [\eta_x \; \eta_y \; \eta_z]^\top \in \mathbb{R}^3$, we can explicitly calculate

$$R\vec{\eta} = \begin{bmatrix} 0 & -\xi_z & \xi_y \\ \xi_z & 0 & -\xi_x \\ -\xi_y & \xi_x & 0 \end{bmatrix} \begin{bmatrix} \eta_x \\ \eta_y \\ \eta_z \end{bmatrix} = \begin{bmatrix} -\xi_z\eta_y + \xi_y\eta_z \\ \xi_z\eta_x - \xi_x\eta_z \\ -\xi_y\eta_x + \xi_x\eta_y \end{bmatrix} = \vec{\xi} \times \vec{\eta}.$$

In particular $R\vec{\xi} = \vec{\xi} \times \vec{\xi} = 0$. If $\vec{\xi} \in \mathbb{R}^3$ is a unit vector and $\vec{\eta} \in \mathbb{R}^3$ is a unit vector orthogonal to $\vec{\xi}$, then $\vec{\tau} = \vec{\xi} \times \vec{\eta} \in \mathbb{R}^3$ is the unique vector that makes the triple $(\vec{\xi}, \vec{\eta}, \vec{\tau})$ into an oriented (right-handed) orthonormal basis of $\mathbb{R}^3$. Now by the above calculation $R\vec{\eta} = \vec{\tau}$, and similarly $R\vec{\tau} = -\vec{\eta}$. Exponentiating R becomes easy in this basis: for $\theta \in \mathbb{R}$, we have

$$\begin{aligned} \operatorname{Exp}\big(\theta R\big)\, \vec{\xi} &= \vec{\xi} \\ \operatorname{Exp}\big(\theta R\big)\, \vec{\eta} &= \cos(\theta)\, \vec{\eta} + \sin(\theta)\, \vec{\tau} \\ \operatorname{Exp}\big(\theta R\big)\, \vec{\tau} &= -\sin(\theta)\, \vec{\eta} + \cos(\theta)\, \vec{\tau}. \end{aligned}$$

In other words, in a different oriented orthonormal basis, $\operatorname{Exp}(\theta R)$ would have the matrix representation just like the rotation around the z-axis.

Lemma II.77 (Surjectivity of the exponential map for SO_3)**.**
The exponential map $\operatorname{Exp}\colon \mathfrak{so}_3 \to SO_3$ *is surjective.*

Proof. Let $g \in SO_3$. According to Lemma II.76 there exists a unit vector $\vec{\xi} \in \mathbb{R}^3$ such that $g\,\vec{\xi} = \vec{\xi}$. Choose a unit vector $\vec{\eta} \in \mathbb{R}^3$ orthogonal to $\vec{\xi}$, and denote by $\vec{\tau} = \vec{\xi} \times \vec{\eta}$ the unique vector that makes the triple $(\vec{\xi}, \vec{\eta}, \vec{\tau})$ into an oriented orthonormal basis of $\mathbb{R}^3$. Note that by special orthogonality of g, the triple $(g\,\vec{\xi},\; g\,\vec{\eta},\; g\,\vec{\tau})$ is also an oriented orthonormal basis. In particular, $g\,\vec{\eta}$ is orthogonal to $g\,\vec{\xi} = \vec{\xi}$, so it lies in the plane spanned by $\vec{\eta}$ and $\vec{\tau}$. Since $g\,\vec{\eta}$ is moreover a unit vector, there exists a $\theta \in [0, 2\pi)$ such that $g\,\vec{\eta} = \cos(\theta)\vec{\eta} + \sin(\theta)\vec{\tau}$. We claim that $g = \operatorname{Exp}(\theta R)$, where $R = \xi_x R_x + \xi_y R_y + \xi_z R_z \in \mathfrak{so}_3$. Indeed we already have $\operatorname{Exp}(\theta R)\,\vec{\xi} = \vec{\xi} = g\,\vec{\xi}$ and $\operatorname{Exp}(\theta R)\,\vec{\eta} = \cos(\theta)\vec{\eta} + \sin(\theta)\vec{\tau} = g\,\vec{\eta}$. But this is sufficient, because both $(g\,\vec{\xi},\; g\,\vec{\eta},\; g\,\vec{\tau})$ and $(\operatorname{Exp}(\theta R)\,\vec{\xi},\; \operatorname{Exp}(\theta R)\,\vec{\eta},\; \operatorname{Exp}(\theta R)\,\vec{\tau})$ are oriented orthonormal bases, so the first two components uniquely determine the third, and therefore g and $\operatorname{Exp}(\theta R)$ agree on all three of the basis vectors $\vec{\xi}, \vec{\eta}, \vec{\tau}$. This explicitly shows that $g \in \operatorname{Exp}(\mathfrak{so}_3)$, proving surjectivity of the exponential map $\mathfrak{so}_3 \to SO_3$. $\square$

The double covering of the rotation group

We are now ready for a concrete description of the topology of SO_3.

Theorem II.78 (Topology of SO_3)**.**

The homomorphism $f\colon \mathrm{SU}_2 \to \mathrm{SO}_3$ of Lie groups in (II.34) *is a two-fold covering: f is surjective and its kernel is* $\mathrm{Ker}(f) = \{+\mathbb{I}, -\mathbb{I}\}$. *In particular f descends to the quotient and gives a group isomorphism*

$$\overline{f}\colon \mathrm{SU}_2/\{+\mathbb{I}, -\mathbb{I}\} \to \mathrm{SO}_3.$$

This $\overline{f}$ is moreover a homeomorphism. Since SU_2 *is homeomorphic to the three-shpere $\mathcal{S}^3$, the matrix Lie group* SO_3 *is homeomorphic to the projective three-sphere (three-sphere with antipodal points identified). In particular* SO_3 *is compact and connected, but it is not simply connected.*

Proof. Recall that by virtue of the connectedness of SU_2 and the first principle, Theorem II.63, $f\colon \mathrm{SU}_2 \to \mathrm{SO}_3$ can be characterized as the unique Lie group homomorphism whose differential $\mathrm{d}f|_{\mathbf{e}}\colon \mathfrak{su}_2 \to \mathfrak{so}_3$ equals the isomorphism of Lie algebras given by (II.31).

The surjectivity of f can now be proved using Lemma II.77. Let $g \in \mathrm{SO}_3$. By the surjectivity of the exponential map $\mathrm{Exp}\colon \mathfrak{so}_3 \to \mathrm{SO}_3$, we can write $g = e^X$ for some $X \in \mathfrak{so}_3$. Let $Y = (\mathrm{d}f|_{\mathbf{e}})^{-1}(X) \in \mathfrak{su}_2$. By the zeroth principle we have

$$f\big(e^Y\big) \;=\; e^{\mathrm{d}f|_{\mathbf{e}}(Y)} \;=\; e^X \;=\; g.$$

The surjectivity of $f\colon \mathrm{SU}_2 \to \mathrm{SO}_3$ follows.

Recall also that f is by construction the adjoint representation $\mathrm{Ad}_\bullet\colon \mathrm{SU}_2 \to \mathrm{Aut}(\mathfrak{su}2)$, except for an identification of $\mathfrak{su}_2$ with $\mathbb{R}^3$. The kernel $\mathrm{Ker}(f)$ is therefore exactly the kernel of the adjoint representation of SU_2. Therefore the assertion is that we have $\mathrm{Ad}_h = \mathrm{id}_{\mathfrak{su}_2}$ if and only if $h \in \{+\mathbb{I}, -\mathbb{I}\}$. For $h = \pm\mathbb{I}$ and $Y \in \mathfrak{su}_2$, we clearly have $\mathrm{Ad}_h(Y) = hYh^{-1} = Y$, so it remains to show that these multiples of the unit matrix are the only elements in the kernel. So suppose that $h \in \mathrm{Ker}(\mathrm{Ad}_\bullet)$, i.e., for all $y \in \mathfrak{su}_2$ we have

$$Y = \mathrm{Ad}_h(Y) = hYh^{-1}.$$

Plugging in $Y = S_x = -\frac{\mathbb{i}}{2}\sigma_1$, $Y = S_y = -\frac{\mathbb{i}}{2}\sigma_2$, and $Y = S_z = -\frac{\mathbb{i}}{2}\sigma_3$ above, we find that h commutes with each of the Pauli matrices: $h\sigma_j = \sigma_j h$ for $j \in 1, 2, 3$. Of course h also commutes with $\mathbb{I}$, and since any complex 2×2 matrix is a complex linear combination of $\mathbb{I}, \sigma_1, \sigma_2, \sigma_3$, it follows that h commutes with all 2×2 matrices. We now basically get to apply Schur's lemma. For example the defining representation $\mathbb{C}^2$ of $\mathrm{GL}_2(\mathbb{C})$ is irreducible (rather obviously), and since h commutes with every element of $\mathrm{GL}_2(\mathbb{C})$, it is an intertwining map from this irreducible representation to itself, therefore a scalar multiple of identity, $h = c\,\mathbb{I}$ for some $c \in \mathbb{C}$. But $h \in \mathrm{SU}_2$ implies $1 = \det(h) = \det(c\,\mathbb{I}) = c^2$, so $c \in \{\pm 1\}$. We have thus proven $\mathrm{Ker}(f) = \{+\mathbb{I}, -\mathbb{I}\}$.

Having seen that $\mathrm{Im}(f) = \mathrm{SO}_3$ and $\mathrm{Ker}(f) = \{+\mathbb{I}, -\mathbb{I}\}$, we get an isomorphism

$$\overline{f}\colon \mathrm{SU}_2/\{+\mathbb{I}, -\mathbb{I}\} \to \mathrm{SO}_3.$$

The quotient is by a discrete subgroup and the mapping f is locally a homeomorphism, so $\overline{f}$ is a homeomorphism. The projective three-sphere is compact, connected and not simply-connected, so the above homomorphism gives these properties for SO_3 as well. Alternatively, we directly see that SO_3 is compact since $\mathrm{SO}_3 \subset \mathbb{R}^{3\times3}$ is closed and bounded. Also we concretely see that SO_3 is smoothly path-connected using the surjectivity of the exponential map proven in Lemma II.77. To see that SO_3 is not simply connected, one can consider first the path $\gamma\colon [0, 2\pi] \to \mathrm{SU}_2$ given by

$$\gamma(t) \;=\; \mathrm{Exp}\big(tS_z\big) \;=\; \begin{bmatrix} e^{-\mathbb{i}t/2} & 0 \\ 0 & e^{\mathbb{i}t/2} \end{bmatrix} \qquad \text{for } t \in [0, 2\pi].$$

This is a path in SU_2 from $\gamma(0) = +\mathbb{I}$ to $\gamma(2\pi) = -\mathbb{I}$. Now

$$f \circ \gamma\colon [0, 2\pi] \to \mathrm{SO}_3$$

is a loop in SO_3: it starts and ends at the neutral element:

$$f(\gamma(0)) \;=\; f(\mathbb{I}) \;=\; \mathbb{I}_{3\times3} \;=\; f(-\mathbb{I}) \;=\; f(\gamma(2\pi)).$$

The reader can check that this loop in SO_3 is non-contractible. □

3.5. Some applications of SO_3 and $\mathfrak{so}_3$

Exercise II.79 (Infinitesimal rotations acting on functions)**.**
Let $C^\infty(\mathbb{R}^3)$ denote the space of smooth complex valued functions on $\mathbb{R}^3$, and on this space, consider the differential operators $\mathcal{J}_x, \mathcal{J}_y, \mathcal{J}_z$ given by $\mathcal{J}_x = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}$, $\mathcal{J}_y = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}$, and $\mathcal{J}_z = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$.

(a) By direct calculation, show that the commutators of the above differential operators are $[\mathcal{J}_x, \mathcal{J}_y] = \mathcal{J}_z$, $[\mathcal{J}_y, \mathcal{J}_z] = \mathcal{J}_x$, $[\mathcal{J}_z, \mathcal{J}_x] = \mathcal{J}_y$.
(b) For $M \in \mathrm{SO}_3$ and $f \in C^\infty(\mathbb{R}^3)$, define $M.f \in C^\infty(\mathbb{R}^3)$ by $(M.f)(\vec{x}) = f(M^{-1}\vec{x})$. Show that $C^\infty(\mathbb{R}^3)$ thus becomes a representation of the group SO_3.
(c) Let
$$M_x^{(\theta)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix},$$
be the rotation by angle θ around the x axis in the positive direction, and let $M_y^{(\theta)}$ and $M_z^{(\theta)}$ be the rotations by θ around y and z-axes, respectively. Show that for any $f \in C^\infty(\mathbb{R}^3)$ we have
$$\frac{\mathrm{d}}{\mathrm{d}\theta}\Big|_{\theta=0} \big(M_x^{(\theta)}.f\big) = \mathcal{J}_x f,$$
and perform similar calculations for the actions of $M_y^{(\theta)}$ and $M_z^{(\theta)}$.

Part III

Representations of Lie algebras and Lie groups

1. Preliminaries on Lie algebra representations

Let us quickly review a few basic notions, before starting to concretely study representations of Lie algebras and Lie groups.

1.1. Representations of Lie algebras and intertwining maps

Recall from Definition II.53 and Remark II.54 that a representation of a Lie algebra $\mathfrak{g}$ over a field $\Bbbk$ on a $\mathbb{K}$-vector space V (with $\Bbbk \subset \mathbb{K}$) is a homomorphism of $\Bbbk$-Lie algebras

$$\vartheta \colon \mathfrak{g} \to \mathrm{End}_{\mathbb{K}}(V).$$

Thus for each $X \in \mathfrak{g}$, the map $\vartheta(X) \colon V \to V$ is $\mathbb{K}$-linear

$$\vartheta(X)(v_1 + v_2) = \vartheta(X)v_1 + \vartheta(X)v_2 \qquad \text{for all } v_1, v_2 \in V, \text{ and}$$
$$\vartheta(X)(cv) = c\,\vartheta(X)v \qquad \text{for all } v \in V,\ c \in \mathbb{K}.$$

The requirement that $X \mapsto \vartheta(X)$ is a Lie algebra homomorhism amounts to $\Bbbk$-linearity

$$\vartheta(aX + bY) = a\,\vartheta(X) + b\,\vartheta(Y) \qquad \text{for all } a, b \in \Bbbk \text{ and } X, Y \in \mathfrak{g}$$

and respecting brackets

$$\vartheta([X,Y]) = \vartheta(X) \circ \vartheta(Y) - \vartheta(Y) \circ \vartheta(X) \qquad \text{for all } X, Y \in \mathfrak{g}.$$

We will mostly be concerned with complex representations, i.e. $\mathbb{K} = \mathbb{C}$, of both real and complex Lie algebras, i.e. $\Bbbk = \mathbb{R}$ or $\Bbbk = \mathbb{C}$.

Definition III.1 (Intertwining map between Lie algebra representations)**.**
Let $\mathfrak{g}$ be a Lie algebra, and $\vartheta_1 \colon \mathfrak{g} \to \mathrm{End}(V_1)$ and $\vartheta_2 \colon \mathfrak{g} \to \mathrm{End}(V_2)$ two representations of $\mathfrak{g}$, on $\mathbb{K}$-vector spaces V_1 and V_2, respectively. A $\mathbb{K}$-linear map $T : V_1 \to V_2$ is said to be an **intertwining map** of representations of $\mathfrak{g}$ if for all $X \in \mathfrak{g}$ we have

$$\vartheta_2(X) \circ T = T \circ \vartheta_1(X).$$

We denote[1] the space of such intertwining maps by $\mathrm{Hom}_{\mathfrak{g}}(V_1, V_2)$.

Subrepresentations are defined in an unsurprising way, starting from the notion of an invariant subspace.

Definition III.2 (Invariant subspace)**.**
Let $\mathfrak{g}$ be a Lie algebra and $\vartheta \colon \mathfrak{g} \to \mathrm{End}(V)$ a representation of it. An **invariant subspace** of the representation is a vector subspace $W \subset V$ such that for all $X \in \mathfrak{g}$ we have $\vartheta(X)W \subset W$.

Definition III.3 (Subrepresentation)**.**
Suppose that $\vartheta \colon \mathfrak{g} \to \mathrm{End}(V)$ is a representation of a Lie algebra $\mathfrak{g}$ on a vector space V, and $W \subset V$ is an invariant subspace. For each $X \in \mathfrak{g}$, the linear map $\vartheta(X) \colon V \to V$ can be restricted to the subspace W, and by the invariance of

[1] Implicit in the notation is the understanding of the field $\mathbb{K}$ over which the representations V_1 and V_2 are defined.

this subspace, the restriction $\vartheta(X)\big|_W =: \tilde{\vartheta}(X)$ defines a map $\tilde{\vartheta}(X)\colon W \to W$. This makes $\tilde{\vartheta}\colon \mathfrak{g} \to \mathrm{End}(W)$ a representation and we correspondingly say that W is a **subrepresentation** in V.

1.2. Complexifications of real Lie algebras

Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{R}$. We will often study the **complexification** of $\mathfrak{g}$,

$$\begin{aligned}\mathfrak{g}_{\mathbb{C}} &:= \mathfrak{g} \oplus \mathbb{i}\,\mathfrak{g} && \text{(a direct sum of } \mathbb{R}\text{-vector spaces)}\\ &= \Big\{ X + \mathbb{i}\,Y \;\Big|\; X, Y \in \mathfrak{g} \Big\}\end{aligned}$$

(equivalently, $\mathfrak{g}$ is the tensor product of the $\mathbb{R}$-vector spaces $\mathfrak{g}$ and $\mathbb{C} = \mathbb{R} \oplus \mathbb{i}\mathbb{R} \cong \mathbb{R}^2$). The complexification $\mathfrak{g}_{\mathbb{C}}$ becomes a $\mathbb{C}$-vector space with the obvious vector addition

$$(X_1 + \mathbb{i}\,Y_1) + (X_2 + \mathbb{i}\,Y_2) = (X_1 + X_2) + \mathbb{i}\,(Y_1 + Y_2) \qquad \text{for all } X_1, Y_1, X_2, Y_2 \in \mathfrak{g}$$

and complex scalar multiplication

$$(a + \mathbb{i}\,b)(X + \mathbb{i}\,Y) = (aX - bY) + \mathbb{i}\,(aY + bX) \qquad \text{for all } a, b \in \mathbb{R} \text{ and } X, Y \in \mathfrak{g}.$$

It moreover becomes a Lie algebra over $\mathbb{C}$ by defining the brackets as

$$[X_1 + \mathbb{i}\,Y_1, X_2 + \mathbb{i}\,Y_2] = ([X_1, X_2] - [Y_1, Y_2]) + \mathbb{i}\,([X_1, Y_2] + [Y_1, X_2]) \qquad \text{for all } X_1, Y_1, X_2, Y_2 \in \mathfrak{g}.$$

If the dimension of the real Lie algebra $\mathfrak{g}$ is $d = \dim_{\mathbb{R}}(\mathfrak{g})$, then the dimension of the complexification $\mathfrak{g}_{\mathbb{C}}$ as a real vector space is $2d$. However, we are more interested in its dimension as a complex vector space, which is just $\dim_{\mathbb{C}}(\mathfrak{g}_{\mathbb{C}}) = d = \dim_{\mathbb{R}}(\mathfrak{g})$.

Example III.4 (Complexification of the Lie algebra $\mathfrak{so}_3$)**.**
Recall from Example II.73 that a basis of the real Lie algebra $\mathfrak{g} = \mathfrak{so}_3$ consists of R_x, R_y, R_z, with nontrivial Lie brackets (II.28)

$$[R_x, R_y] = R_z, \qquad [R_y, R_z] = R_x, \qquad [R_z, R_x] = R_y.$$

The complexification $\mathfrak{g}_{\mathbb{C}} = \mathfrak{so}_3 \oplus \mathbb{i}\,\mathfrak{so}_3$ consists of all complex linear combinations of these basis elements. It contains for example the elements

$$E = -R_y + \mathbb{i}\,R_x \in \mathfrak{g}_{\mathbb{C}}, \qquad F = R_y + \mathbb{i}\,R_x \in \mathfrak{g}_{\mathbb{C}}, \qquad H = 2\mathbb{i}\,R_z \in \mathfrak{g}_{\mathbb{C}}.$$

It is easy to see that E, F, H above are linearly independent (over $\mathbb{C}$) in $\mathfrak{g}_{\mathbb{C}}$, and thus form a basis (over $\mathbb{C}$) of $\mathfrak{g}_{\mathbb{C}}$. The nontrivial Lie brackets in this basis of the complexification $\mathfrak{g}_{\mathbb{C}} = \mathfrak{so}_3 \oplus \mathbb{i}\,\mathfrak{so}_3$ of $\mathfrak{g} = \mathfrak{so}_3$ are

$$\begin{aligned}[H, E] &= [2\mathbb{i}\,R_z\,,\; -R_y + \mathbb{i}\,R_x] = -2\mathbb{i}\,[R_z, R_y] - 2[R_z, R_x] = 2\mathbb{i}\,R_x - 2R_y = 2E\\ [H, F] &= [2\mathbb{i}\,R_z\,,\; R_y + \mathbb{i}\,R_x] = 2\mathbb{i}\,[R_z, R_y] - 2[R_z, R_x] = -2\mathbb{i}\,R_x - 2R_y = -2F\\ [E, F] &= [-R_y + \mathbb{i}\,R_x\,,\; R_y + \mathbb{i}\,R_x] = -\mathbb{i}\,[R_y, R_x] + \mathbb{i}\,[R_x, R_y] = 2\mathbb{i}\,R_z = H.\end{aligned}$$

In view of these relations, we will soon see that the complexification of $\mathfrak{so}_3$ is isomorphic to the complex Lie algebra $\mathfrak{sl}_2(\mathbb{C})$.

The primary justification for studying representations of the complex Lie algebra $\mathfrak{g}_{\mathbb{C}}$ instead of the real Lie algebra $\mathfrak{g}$ is that the complex representations of both are basically the same. Working with complex numbers just makes certain things easier for $\mathfrak{g}_{\mathbb{C}}$.

Lemma III.5 (Complex representations of a real Lie algebra).
Let $\mathfrak{g}$ be a real Lie algebra, and $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \oplus \mathbb{i}\, \mathfrak{g}$ 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 $\vartheta \colon \mathfrak{g} \to \mathrm{End}(V)$ 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 $\vartheta_{\mathbb{C}} \colon \mathfrak{g}_{\mathbb{C}} \to \mathrm{End}(V)$ by setting $\vartheta_{\mathbb{C}}(X + \mathbb{i}\, Y) = \vartheta(X) + \mathbb{i}\, \vartheta(Y)$. 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 intertwining maps, if $T_{\mathbb{C}} \colon V \to W$ is an intertwining map $\mathfrak{g}_{\mathbb{C}}$-representations, then a fortiori it is a an intertwining map of $\mathfrak{g}$-representations. We only need to show the other direction, that if $T \colon V \to W$ is an intertwining map of $\mathfrak{g}$-representations, then it is also an intertwining map of $\mathfrak{g}_{\mathbb{C}}$-representations. This is easy to see by $\mathbb{C}$-linearity of T and the way the representations $\vartheta^V_{\mathbb{C}}$ and $\vartheta^W_{\mathbb{C}}$ extend ϑ^V and ϑ^W. □

1.3. Direct sums of representations

If $\vartheta_1 \colon \mathfrak{g} \to \mathrm{End}(V_1)$ and $\vartheta_2 \colon \mathfrak{g} \to \mathrm{End}(V_2)$ are two representations of a Lie algebra $\mathfrak{g}$, then there is an obvious way to make the direct sum $V_1 \oplus V_2$ a representation of $\mathfrak{g}$ — just set

$$\vartheta(X)(v_1 + v_2) = \vartheta_1(X)v_1 + \vartheta_2(X)v_2 \qquad \text{for all } X \in \mathfrak{g} \text{ and } v_1 \in V_1,\ v_2 \in V_2.$$

The question of when a subrepresentation has a complementary subrepresentation is more involved than in the case of finite groups.

One sufficient condition is the existence of an invariant inner product.

Definition III.6 (Invariant inner product).
Suppose that $\vartheta \colon \mathfrak{g} \to \mathrm{End}(V)$ is a representation of a real Lie algebra $\mathfrak{g}$ on a real or complex inner product space V. The inner product $\langle \cdot, \cdot \rangle$ on V is said to be $\mathfrak{g}$-invariant if we have

$$\langle \vartheta(X)v_1, v_2 \rangle = -\langle v_1, \vartheta(X)v_2 \rangle \quad \text{for all } X \in \mathfrak{g} \text{ and } v_1, v_2 \in V. \tag{III.1}$$

Lemma III.7 (Invariant inner product gives complementary subrepresentations).
Suppose that $\vartheta \colon \mathfrak{g} \to \mathrm{End}(V)$ is a representation of a real Lie algebra $\mathfrak{g}$ on a real or complex vector space V which has a $\mathfrak{g}$-invariant inner product. If $W \subset V$ is a subrepresentation, then also the orthogonal complement

$$W^\perp := \Big\{ u \in V \ \Big|\ \langle u, w \rangle = 0 \quad \forall w \in W \Big\}$$

is a subrepresentation.

Proof. Suppose that $u \in W^\perp$, i.e., that we have $\langle u, w \rangle = 0$ for all $w \in W$. Consider $\vartheta(X)u$ for some $X \in \mathfrak{g}$. Observe that by $\mathfrak{g}$-invariance of the inner product, for any $w \in W$ we get

$$\langle \vartheta(X)u, w \rangle = -\langle u, \vartheta(X)w \rangle = 0,$$

since also $\vartheta(X)w \in W$ by the subrepresentation property. □

1.4. Irreducible representations

The definition of irreducible representations of Lie algebras is similar to the one for groups.

Definition III.8 (Irreducible representation)**.**
A representation $\vartheta \colon \mathfrak{g} \to \mathrm{End}(V)$ of a Lie algebra $\mathfrak{g}$ on a vector space $V \neq \{0\}$ is called **irreducible** if its only invariant subspaces are $\{0\} \subset V$ and $V \subset V$.

When we have a criterion guaranteeing existence of complementary subrepresentations, we obtain complete reducibility for finite-dimensional representations by induction on dimension. As an example, we have the following consequence of Lemma III.7.

Corollary III.9 (Invariant inner product gives complete reducibility)**.**
Suppose that $\vartheta \colon \mathfrak{g} \to \mathrm{End}(V)$ is a representation of a real Lie algebra $\mathfrak{g}$ on a finite-dimensional real or complex vector space V which has a $\mathfrak{g}$-invariant inner product. Then V is the direct sum of subrepresentations,

$$V = W_1 \oplus \cdots \oplus W_\ell,$$

where each $W_j \subset V$, $j = 1, \ldots, \ell$, is irreducible.

Schur's lemmas

Schur's lemmas apply just as well to representations of Lie algebras as they do to representations of groups. The proofs remain identical, since the arguments only use the properties that the kernel and image of an intertwining map are subrepresentations, and that a bijective intertwining map is an isomorphism of representations.

Lemma III.10 (Schur's lemma)**.**
If V and W are irreducible representations of a Lie algebra $\mathfrak{g}$, and $T : V \to W$ is an intertwining map, then either $T = 0$ or T is an isomorphism.

Under the assumption that the representations are defined over an algebraically closed field and that they are finite dimensional, we have the two further formulations.

Lemma III.11 (Schur's lemma over albegraically closed fields)**.**
Let $\mathfrak{g}$ be a Lie algebra, and $\vartheta \colon \mathfrak{g} \to \mathrm{End}_{\mathbb{K}}(V)$ an irreducible representation of $\mathfrak{g}$ in a finite-dimensional vector space V over a field $\mathbb{K}$ such that (AlgClos) *holds.*

Then any intertwining map $T : V \to V$ is necessarily of the form $T = \lambda\, \mathrm{id}_V$ for some $\lambda \in \mathbb{K}$.

Corollary III.12 (Schur's lemma for dimension of intertwining maps)**.**
Let $\mathfrak{g}$ be a Lie algebra, and $\vartheta_V \colon \mathfrak{g} \to \mathrm{End}_{\mathbb{K}}(V)$ and $\vartheta_W \colon \mathfrak{g} \to \mathrm{End}_{\mathbb{K}}(W)$ two irreducible representations of $\mathfrak{g}$ in finite-dimensional vector spaces V and W over a field $\mathbb{K}$ such that (AlgClos) *holds.*

Then the dimension of the space of intertwining maps between these irreducible representations is given by

$$\dim\Big(\mathrm{Hom}_{\mathfrak{g}}(V,W)\Big) = \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{if } V \ncong W. \end{cases}$$

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

We start by analyzing Lie algebra representations in 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 $\mathfrak{su}_2$ — see Example III.4. As such, the complex representations of $\mathfrak{so}_3$ and $\mathfrak{su}_2$ are exactly the same as those of $\mathfrak{sl}_2(\mathbb{C})$, according to Lemma III.5. 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 $\mathfrak{su}_2$, respectively.
- The analysis of all (complex) 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 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 example), this is really the fundamental reason for studying it!

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

By definition, $\mathfrak{sl}_2(\mathbb{C})$ is the set

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

of traceless complex two-by-two matrices, equipped with the Lie bracket $[X, Y] = XY - YX$. As a complex vector space, it is three dimensional, and we will use the basis

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{III.2}$$

for it. The brackets of these basis elements are easily calculated, with the result

$$[H, E] = 2E, \qquad [H, F] = -2F, \qquad [E, F] = H. \tag{III.3}$$

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 III.13 (Defining representation of $\mathfrak{sl}_2(\mathbb{C})$)**.**
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

$$\vartheta(X)\, v = Xv,$$

i.e., we set $\vartheta(X) = X \in \mathbb{C}^{2\times 2} \cong \mathrm{End}(V)$. This two-dimensional representation is called the defining representation of $\mathfrak{sl}_2(\mathbb{C})$.

Example III.14 (Adjoint representation of $\mathfrak{sl}_2(\mathbb{C})$)**.**
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

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

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

$$\vartheta \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}(\mathfrak{sl}_2(\mathbb{C}))$$

becomes, in view of (III.3),

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

2.2. Weight spaces in representations of $\mathfrak{sl}_2(\mathbb{C})$

Let V be a finite dimensional (complex) vector space, which carries a representation of $\mathfrak{sl}_2(\mathbb{C})$. The representation $\vartheta \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}(V)$ gives us linear maps $V \to V$ associated to each Lie algebra element $X \in \mathfrak{sl}_2(\mathbb{C})$. Let us denote the linear maps associated with basis elements H, E, F in (III.2) by

$$\mathcal{H} = \vartheta(H), \qquad \mathcal{E} = \vartheta(E), \qquad \mathcal{F} = \vartheta(F).$$

These are linear maps

$$\mathcal{H} \colon V \to V, \qquad \mathcal{E} \colon V \to V, \qquad \mathcal{F} \colon V \to V,$$

which by virtue of the Lie brackets in (III.3) and homomorphism property of ϑ satisfy the commutation relations

$$\mathcal{H}\mathcal{E} - \mathcal{E}\mathcal{H} = 2\,\mathcal{E}, \qquad \mathcal{H}\mathcal{F} - \mathcal{F}\mathcal{H} = -2\,\mathcal{F}, \qquad \mathcal{E}\mathcal{F} - \mathcal{F}\mathcal{E} = \mathcal{H}. \tag{III.4}$$

Remark III.15 (Remarks on notation)**.**

One obvious reason for preferring $\mathcal{H}, \mathcal{E}, \mathcal{F}$ to $\vartheta(H), \vartheta(E), \vartheta(F)$ is to have a less cumbersome notation, while still distinguishing the matrices $H, E, F \in \mathbb{C}^{2\times 2}$ from the linear maps $\mathcal{H}, \mathcal{E}, \mathcal{F} \in \mathrm{End}(V) \cong \mathbb{C}^{d\times d}$, where $d = \dim(V)$.

Moreover, an advantage of linear maps is that they can be composed, so that for example $\mathcal{E}\mathcal{F}$ makes sense as a linear map $V \to V$, while the product EF is something we should not really ever write — the matrix product $EF \in \mathbb{C}^{2\times 2}$ is not traceless, and so does not lie in $\mathfrak{sl}_2(\mathbb{C})$.[2]

In the long run when studying Lie algebra representations, convenience dictates that we should start using "module notation", i.e., abuse notation and write just Xv for $\vartheta(X)v$ when $v \in V$ and $X \in \mathfrak{g}$ — but due to a non-negligible risk of confusion (think of Examples III.13 and III.14) let us for the moment avoid doing so.

Although we will not logically need the following fact, it is convenient to be aware of it already at this point.

Fact III.16. The linear map $\mathcal{H} \colon V \to V$ is diagonalizable.

[2]Products of elements in a Lie algebra do not make sense, since Lie algebras are not associative algebras! It would, however, be possible to consider the **universal enveloping algebra** $\mathcal{U}(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$, which *is* an associative algebra constructed in such a way that it has the "same" representations as $\mathfrak{g}$. We avoid introducing the universal enveloping algebra, as it would have relatively few real benefits in the limited scope of the present course.

We will obtain a proof of this fact in Section 2.5 together with the proof of complete reducibility.

For $\mu \in \mathbb{C}$, let us denote the eigenspace of $\mathcal{H}$ corresponding to the eigenvalue μ by

$$V_\mu := \big\{ v \in V \;\big|\; \mathcal{H}v = \mu v \big\}. \tag{III.5}$$

To anticipate some general terminology, we call the eigenvalues μ of $\mathcal{H}$ **weights** (in the representation V), and the corresponding eigenspaces $V_\mu \subset V$ **weight spaces** (in the representation V). Note that if $\mu \in \mathbb{C}$ is not a weight (i.e., not an eigenvalue of $\mathcal{H}$), then $V_\mu = \{0\} \subset V$.

Admitting Fact III.16, one has the eigenspace decomposition

$$V = \bigoplus_\mu V_\mu. \tag{III.6}$$

The decomposition (III.6) completely describes the action of $\mathcal{H}$ on V, and the remaining task is to describe the action of $\mathcal{E}$ and $\mathcal{F}$.

The following lemma describes what $\mathcal{E}$ and $\mathcal{F}$ do to the $\mathcal{H}$-eigenspaces V_μ.

Lemma III.17 (Raising and lowering of weights)**.**
For any $\mu \in \mathbb{C}$, we have

$$\mathcal{E}(V_\mu) \subset V_{\mu+2} \qquad \textit{and} \qquad \mathcal{F}(V_\mu) \subset V_{\mu-2}.$$

Proof. Suppose that $v \in V_\mu$, i.e., $\mathcal{H}v = \mu v$. Consider the vector $\mathcal{E}v \in V$. We can figure out the action of $\mathcal{H}$ on it by an easy but important calculation which uses the commutator (III.4) of $\mathcal{H}$ and $\mathcal{E}$ corresponding to the Lie bracket $[H, E]$ in (III.3).

Fundamental calculation (first time):

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

This calculation shows that if v is an eigenvector of $\mathcal{H}$ with eigenvalue μ, then $\mathcal{E}v$ is either zero, or an eigenvector of $\mathcal{H}$ with eigenvalue $\mu + 2$. This proves that $\mathcal{E}(V_\mu) \subset V_{\mu+2}$.

By an entirely similar calculation we see that $\mathcal{F}(V_\mu) \subset V_{\mu-2}$. □

The above lemma has a consequential corollary for finite-dimensional representations V of $\mathfrak{sl}_2(\mathbb{C})$.

Corollary III.18 (Existence of highest weight vectors for $\mathfrak{sl}_2(\mathbb{C})$)**.**
In any finite-dimensional representation $V \neq \{0\}$ of $\mathfrak{sl}_2(\mathbb{C})$, there exists a non-zero vector $v_0 \in V$ such that $\mathcal{E}v_0 = 0$ and $\mathcal{H}v_0 = \lambda v_0$ for some $\lambda \in \mathbb{C}$.

Proof. Let $\lambda \in \mathbb{C}$ be an eigenvalue of $\mathcal{H}$ with maximal real part.[3] Then the eigenspace $V_\lambda \subset V$ is non-zero, so we can pick a non-zero vector $v_0 \in V_\lambda$ in it. By definition this vector satisfies $\mathcal{H}v_0 = \lambda v_0$. Consider $\mathcal{E}v_0 \in V$. By Lemma III.17 we have $\mathcal{E}v_0 \in V_{\lambda+2}$, but since the real part satisfies $\mathfrak{Re}(\lambda + 2) > \mathfrak{Re}(\lambda)$, we must have $V_{\lambda+2} = \{0\}$. Therefore we have $\mathcal{E}v_0 = 0$, and the proof is complete. □

[3] Note that $\mathcal{H}$ has at least one eigenvalue, since it is a linear map of a non-zero finite-dimensional vector space over the algebraically closed field $\mathbb{C}$.

Definition III.19 (Highest weigth vector for $\mathfrak{sl}_2(\mathbb{C})$).
A vector $v_0 \neq 0$ in a representation V of $\mathfrak{sl}_2(\mathbb{C})$, which satisfies $\mathcal{E}v_0 = 0$ and $\mathcal{H}v_0 = \lambda v_0$, is called a **highest weight vector** of **highest weight** λ.

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

We next start considering finite-dimensional representations which are irreducible. The first observation, however, is to note the following rather explicit description of the subrepresentation generated by a highest weight vector.

Lemma III.20 (Subrepresentation generated by a highest weight vector).
Let $V \neq \{0\}$ be a finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$, and let $v_0 \in V$ be a highest weight vector with highest weight $\lambda \in \mathbb{C}$. For all $j \in \mathbb{Z}_{\geq 0}$, define

$$v_j := \mathcal{F}^j v_0.$$

Then the linear span of $v_0, v_1, v_2, \ldots$ in V is a subrepresentation, and the action of the generators on the vectors v_j, $j \in \mathbb{Z}_{\geq 0}$, is explicitly given by

$$\begin{aligned} \mathcal{H}v_j &= (\lambda - 2j)\, v_j \\ \mathcal{F}v_j &= v_{j+1} \\ \mathcal{E}v_j &= j(\lambda + 1 - j)\, v_{j-1}. \end{aligned} \tag{III.7}$$

Proof. Once we prove formulas (III.7), it follows that the linear span of $v_0, v_1, v_2, \ldots$ in V is a subrepresentation, because it is stable under applying all of the generators of $\mathfrak{sl}_2(\mathbb{C})$.

Let us prove (III.7) by induction on $j \in \mathbb{Z}_{\geq 0}$. The case $j = 0$ is clear by the assumption that v_0 is a highest weight vector, and since $\mathcal{F}v_0 = v_1$ by construction. Consider then a general $j > 0$ and assume that the result has been proven for $j - 1$ already. By the induction assumption we have in particular $\mathcal{H}v_{j-1} = (\lambda - 2j + 2)v_{j-1}$, i.e., $v_{j-1} \in V_{\lambda-2j+2}$. It follows from Lemma III.17 that $v_j = \mathcal{F}v_{j-1} \in V_{\lambda-2j}$, which proves the desired formula $\mathcal{H}v_j = (\lambda - 2j)v_j$. The formula $\mathcal{F}v_j = v_{j+1}$ follows directly by the construction of v_{j+1}. It remains to prove the formula for $\mathcal{E}v_j$, and by induction assumption we have

$$\mathcal{E}v_{j-1} = (j-1)(\lambda + 2 - j)\, v_{j-2}.$$

Now using the last of the commutation relations in (III.4) and the induction assumptions, calculate

$$\begin{aligned} \mathcal{E}v_j = \mathcal{E}\mathcal{F}v_{j-1} &= (\mathcal{F}\mathcal{E} + \mathcal{H})\, v_{j-1} \\ &= \mathcal{F}(\mathcal{E}v_{j-1}) + \mathcal{H}v_{j-1} \\ &= \mathcal{F}\big((j-1)(\lambda + 2 - j)\, v_{j-2}\big) + (\lambda - 2j + 2)\, v_{j-1} \\ &= \Big((j-1)(\lambda + 2 - j) + (\lambda - 2j + 2)\Big)\, v_{j-1} \\ &= \big(j(\lambda + 1 - j)\big)\, v_{j-1}, \end{aligned}$$

which establishes the last remaining formula, and finishes the proof. □

Corollary III.21 (Highest weights in finite-dimensional $\mathfrak{sl}_2(\mathbb{C})$-representations).
Let $V \neq \{0\}$ be a finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$, and let $v_0 \in V$ be a highest weight vector with highest weight λ. Then we have $\lambda \in \mathbb{Z}_{\geq 0}$.

Proof. For $j \in \mathbb{Z}_{\geq 0}$, define $v_j = \mathcal{F}^j v_0$ as in Lemma III.20. Since $v_j \in V_{\lambda-2j}$, each non-zero v_j lies in a different eigenspace of $\mathcal{H}$, so only finitely many of the v_j can be non-zero, due to the finite-dimensionality of V. Let $d \in \mathbb{Z}_{>0}$ be the smallest number such that $v_d = 0$. Then we have $v_{d-1} \neq 0$. By using the formulas from Lemma III.20, we now observe that

$$0 = \mathcal{E} v_d = d(\lambda + 1 - d)\, v_{d-1},$$

which implies that $\lambda + 1 - d = 0$, since $d \neq 0$ and $v_{d-1} \neq 0$. In other words, we have

$$\lambda = d - 1 \in \mathbb{Z}_{\geq 0},$$

which proves the assertion. □

The idea above results in a full classification of all irreducible finite-dimensional representations of $\mathfrak{sl}_2(\mathbb{C})$.

Theorem III.22 (Irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$)**.**
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 $\mathcal{H} = \vartheta_\lambda(H)$, $\mathcal{E} = \vartheta_\lambda(E)$, and $\mathcal{F} = \vartheta_\lambda(F)$ on this basis given by

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

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

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

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. If V is a finite-dimensional irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$, then by Corollary III.18, there exists a highest weight vector v_0 of some highest weight $\lambda \in \mathbb{C}$ in V. By Corollary III.21 we must have $\lambda \in \mathbb{Z}_{\geq 0}$. Denote $v_j = \mathcal{F}^j v_0$ for $j \in \mathbb{Z}_{\geq 0}$ as before. By the same corollary together with Lemma III.20 we get that the vectors $v_0, v_1, \ldots, v_\lambda$ span a subrepresentation in V, which by irreducibility has to be all of V. Since the $\mathcal{H}$-eigenvalues of $v_0, v_1, \ldots, v_\lambda$ are different, these vectors are linearly independent, and therefore form a basis of V. It follows from the same lemma that the action of $\mathcal{H}$, $\mathcal{E}$, and $\mathcal{F}$ on this basis are as in the assertion. Thus indeed any irreducible representation has to be isomorphic to $L(\lambda)$ for some $\lambda \in \mathbb{Z}_{\geq 0}$.

To prove that such a representation $\mathrm{L}(\lambda)$ exists for each $\lambda \in \mathbb{Z}_{\geq 0}$, it suffices to verify that the linear maps $\mathcal{H}$, $\mathcal{E}$, and $\mathcal{F}$ defined by these formulas satisfy the commutation relations (III.4), and thus indeed can be used to define a representation $\vartheta_\lambda \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}(V)$, where V is a vector space with basis $v_0, v_1, \ldots, v_\lambda$. This verification is straightforward.[4] □

We finish our discussion about the irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$ by a few observations which follow immediately from the above, but which are often practical.

Corollary III.23 (Observations about $\mathfrak{sl}_2(\mathbb{C})$ representations)**.**
In an irreducible $\mathfrak{sl}_2(\mathbb{C})$ representation V of dimension $\dim(V) = d < \infty$, the (unique) highest weight is $\lambda = d - 1$. The eigenvalues of $\mathcal{H}$ are

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

[4]Alternatively, an indirect argument for the existence of such a representation $L(\lambda)$ can be given by finding some finite-dimensional representation, which contains a highest weight vector of the given highest weight $\lambda \in \mathbb{Z}_{\geq 0}$. One easy and concrete possibility is to consider the representation $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ (λ times) where each factor is the defining representation $\mathbb{C}^2$ (with standard basis $x = [1\ 0]^\top$, $y = [0\ 1]^\top$), and note that the vector $v_0 := x \otimes \cdots \otimes x$ indeed satisfies $\mathcal{H} v_0 = \lambda v_0$ and $\mathcal{E} v_0 = 0$.

and the multiplicity of each eigenvalue is one. In particular, the $\mathcal{H}$-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 $-\mu$).

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

To illustrate the classification of irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$, Theorem III.22, in concrete cases, let us look at a few examples.

The trivial representation

The trivial representation[5]

$$V = \mathbb{C}$$

is one-dimensional, and therefore necessarily irreducible. By Theorem III.22 (or Corollary III.23) the only 1-dimensional irreducible of $\mathfrak{sl}_2(\mathbb{C})$ corresponds to $\lambda = 0$, so we have

$$V \cong L(0).$$

The defining representation

In Example III.13, 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 as, e.g., in (III.2), 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 defining 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 weight spaces are $V_{+1} = \mathbb{C}x$ and $V_{-1} = \mathbb{C}y$. Also we have $Ex = 0$, so $x \in V$ is a highest weight vector of highest weight $\lambda = 1$, and it therefore generates an irreducible subrepresentation of dimension $d = \lambda + 1 = 2 = \dim(V)$, which by dimensionality therefore is all of V. Therefore the defining representation is irreducible,

$$V \cong L(1).$$

The adjoint representation

In Example III.14, we noted that The vector space

$$V = \mathfrak{sl}_2(\mathbb{C})$$

is a representation of the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ by the adjoint action. The action of the generators is $\mathcal{E} = \mathrm{ad}_E$, $\mathcal{F} = \mathrm{ad}_F$, and $\mathcal{H} = \mathrm{ad}_H$. Note that

$$\begin{aligned}
\mathcal{H}E &= \mathrm{ad}_H(E) = [H, E] = 2\,E, \\
\mathcal{H}H &= \mathrm{ad}_H(H) = [H, H] = 0, \\
\mathcal{H}F &= \mathrm{ad}_H(F) = [H, F] = -2\,F,
\end{aligned}$$

[5] The trivial representation of a Lie algebra $\mathfrak{g}$ on the vector space $\mathbb{C}$ is defined by $\vartheta(X) = 0 \in \mathrm{End}(\mathbb{C})$ for all $X \in \mathfrak{g}$.

so that the $\mathcal{H}$-eigenvalues are $+2$, 0, and -2, each with multiplicity one. The corresponding weight spaces are $V_{+2} = \mathbb{C}E$, $V_0 = \mathbb{C}H$, and $V_{-2} = \mathbb{C}F$. Note that we also have $\mathcal{E}E = \mathrm{ad}_E(E) = [E,E] = 0$, so E is a highest weight vector wiht highest weight $\lambda = 2$, and it therefore generates an irreducible subrepresentation of dimension $d = \lambda + 1 = 3 = \dim(V)$, which by dimensionality therefore is all of V. Therefore the adjoint representation is irreducible,

$$V \cong L(2).$$

Further examples

Exercise III.24 (Homogeneous polynomial representations of $\mathfrak{sl}_2(\mathbb{C})$)**.**
Let $V = \mathbb{C}[x,y]$ be the polynomial algebra in two indeterminates, x and y.

(a) Define a linear map $\mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}(V)$ by setting

$$E \mapsto x\frac{\partial}{\partial y}, \qquad F \mapsto y\frac{\partial}{\partial x}, \qquad H \mapsto x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}.$$

Show that V thus becomes a representation of $\mathfrak{sl}_2(\mathbb{C})$ (albeit ∞-dimensional).

(b) Let $V_m \subset V$ be the subspace of homogeneous polynomials of degree $m \in \mathbb{Z}_{\geq 0}$, i.e., the linear span of monomials $x^i y^j$ with $i+j=m$. Show that $V_m \subset V$ is a finite-dimensional subrepresentation of the representation in part (a).

(c) Show that the finite dimensional subrepresentation $V_m \subset V$ in part (b) is irreducible. Show in particular that V_m is isomorphic to an irreducible finite-dimensional highest weight representation $L(\lambda)$ with highest weight λ, for a certain $\lambda \in \mathbb{Z}_{\geq 0}$.

2.5. Complete reducibility for representations of $\mathfrak{sl}_2(\mathbb{C})$

We now turn to the proof of complete reducibility of finite-dimensional representations of $\mathfrak{sl}_2(\mathbb{C})$. We will need two auxiliary results, which are obtained by fairly direct calculations in Exercise III.25(a) and Exercise III.26 below.

Exercise III.25 (A lemma for diagonalizability of the action of H)**.**
Suppose that V is a complex vector space, and $\mathcal{E}, \mathcal{F}, \mathcal{H}$ are linear operators $V \to V$ which satisfy the commutation relations (III.4).

(a) Show that for any $k \in \mathbb{Z}_{>0}$, we have

$$\mathcal{E}\mathcal{F}^k = \mathcal{F}^k\mathcal{E} + k\mathcal{F}^{k-1}(\mathcal{H} - k + 1),$$
$$\mathcal{F}\mathcal{E}^k = \mathcal{E}^k\mathcal{F} + k(\mathcal{H} - k + 1)\mathcal{E}^{k-1}.$$

(b) Show that for any $k \in \mathbb{Z}_{>0}$, we have

$$\mathcal{E}^k\mathcal{F}^k = k!\,\mathcal{H}\big(\mathcal{H}-1\big)\cdots\big(\mathcal{H}-(k-1)\big) + \mathcal{P}\,\mathcal{E},$$

where $\mathcal{P}$ is some operator $V \to V$ (depending on k) that can be written as a polynomial in the operators $\mathcal{E}, \mathcal{F}, \mathcal{H}$.

Exercise III.26 (Quadratic Casimir in $\mathfrak{sl}_2(\mathbb{C})$ representations)**.**
Suppose that V is a complex vector space, and $\mathcal{E}, \mathcal{F}, \mathcal{H} \in \mathrm{End}(V)$ satisfy $\mathcal{H}\mathcal{E} - \mathcal{E}\mathcal{H} = 2\,\mathcal{E}$, $\mathcal{H}\mathcal{F} - \mathcal{F}\mathcal{H} = -2\,\mathcal{F}$, $\mathcal{E}\mathcal{F} - \mathcal{F}\mathcal{E} = \mathcal{H}$.

(a) Show that the operator

$$\mathcal{Q} = \mathcal{E}\mathcal{F} + \mathcal{F}\mathcal{E} + \frac{1}{2}\mathcal{H}^2$$

commutes with $\mathcal{E}$, $\mathcal{F}$, and $\mathcal{H}$.

(b) Let $\vartheta \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}(V)$ be a finite-dimensional irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$. Show that $\mathcal{Q} = \vartheta(E)\vartheta(F) + \vartheta(F)\vartheta(E) + \frac{1}{2}\vartheta(H)^2$ is a scalar multiple of the identity operator on V, that is $\mathcal{Q} = q \times \mathrm{id}_V$ for some $q \in \mathbb{C}$.

(c) For each finite-dimensional irreducible representation $L(\lambda)$ of $\mathfrak{sl}_2(\mathbb{C})$, find the value of the scalar $q = q_\lambda$ in part (b). Observe in particular that for $\lambda_1 \neq \lambda_2$ we have $q_{\lambda_1} \neq q_{\lambda_2}$.

We are now ready to prove the complete reducibility statement.

Theorem III.27 (Complete reducibility for representations of $\mathfrak{sl}_2(\mathbb{C})$)**.** *Any finite-dimensional representation V of $\mathfrak{sl}_2(\mathbb{C})$ is a direct sum of its irreducible subrepresentations, in particular*

$$V \;\cong\; L(\lambda_1) \oplus \cdots \oplus L(\lambda_s)$$

for some s and $\lambda_1, \ldots, \lambda_s \in \mathbb{Z}_{\geq 0}$.

Proof. As in Exercise III.26, consider the operator

$$\mathcal{Q} = \vartheta(E)\vartheta(F) + \vartheta(F)\vartheta(E) + \frac{1}{2}\vartheta(H)^2 \tag{III.8}$$

on V, and think in particular of the Jordan blocks of $\mathcal{Q}$ on V. By Exercise III.26(a), $\mathcal{Q}$ commutes with the generators of $\mathfrak{sl}_2(\mathbb{C})$ action on V, so each block $\mathrm{Ker}\big((\mathcal{Q} - q\,\mathrm{id}_V)^p\big) \subset V$ (with $p \geq \dim(V)$, say) is a subrepresentation. The blocks with different (generalized) eigenvalues q are complementary to each other (in direct sum), so it suffices to prove complete reducibility for each block separately.

Assume therefore without loss of generality that $\mathcal{Q}$ has only one (generalized) eigenvalue q on V. A non-zero finite-dimensional representation V has to contain some irreducible subrepresentations, so choose one with maximal dimension d, which is correspondingly isomorphic to $L(\lambda)$ with $\lambda = d - 1$. Note also that $\mathcal{Q}$ has eigenvalue q_λ on this irreducible, so we must have $q = q_\lambda$.

Let $V_0 \subset V$ be the direct sum of all irreducible subrepresentations of V which are isomorphic to the above $L(\lambda)$. If $V_0 = V$, then V is completely reducible and we are done, so let us assume that $V_0 \neq V$, i.e., we have a non-trivial quotient $V/V_0 \neq \{0\}$. In the quotient V/V_0, choose some irreducible subrepresentation $\widetilde{W} \subset V/V_0$, and let $W \subset V$ be the corresponding subrepresentation so that $V_0 \subset W$ and $\widetilde{W} = W/V_0 \subset V/V_0$. Note that the only possible generalized eigenvalue of $\mathcal{Q}$ on W and $\widetilde{W}$ is still $q = q_\lambda$, so by Exercise III.26(c) the irreducible $\widetilde{W}$ must also be isomorphic to the same one, $\widetilde{W} \cong L(\lambda)$.

Now note that for $d = \lambda + 1$ we have that $\mathcal{E}^d$ is zero on V_0 (because it is zero on $L(\lambda)$), and similarly for $\widetilde{W}$. Therefore at least $\mathcal{E}^{2d}$ is zero on W. Let k be the largest integer such that $\mathcal{E}^k|_W \neq 0$, so in particular $\mathcal{E}^{k+1}|_W = 0$. Clearly we have $k \geq \lambda$, since W contains copies of $L(\lambda)$. On the other hand we have the formula

$$\mathcal{F}\mathcal{E}^{k+1} - \mathcal{E}^{k+1}\mathcal{F} = (k+1)(\mathcal{H} - k)\mathcal{E}^k$$

of Exercise III.25(a), whose left hand side vanishes on W. Since $\mathcal{E}^k$ does not vanish, the formula implies the existence of an eigenvector of $\mathcal{H}$ of W with eigenvalue k. This is only possible if $k \leq \lambda$, so we conclude that $k = \lambda$.

We have thus established that $\mathcal{E}^\lambda|_W \neq 0$ and $\mathcal{E}^{\lambda+1}|_W = 0$. Let now $\tilde{w} \in \widetilde{W} = W/V_0$ be a highest weight vector of $\widetilde{W} \cong L(\lambda)$, and let $w \in W$ be a representative for its equivalence class, i.e., $w + V_0 = \tilde{w}$. Consider $w' = \mathcal{E}^\lambda\mathcal{F}^\lambda w \in W$, and note that it is non-zero, $w' \neq 0$ (indeed $\mathcal{E}^\lambda\mathcal{F}^\lambda\tilde{w} \neq 0$ by the explicit formulas for $L(\lambda)$ given in Theorem III.22). We also have $\mathcal{E}w' = \mathcal{E}^{\lambda+1}\mathcal{F}^\lambda w = 0$ and $(\mathcal{H} - \lambda)w' = (\mathcal{H} - \lambda)\mathcal{E}^\lambda\mathcal{F}^\lambda w = 0$ by the earlier calculation. These properties show that $w' \in W$ is a highest weight vector of highest weight λ, but since $w' \notin V_0$, it generates a copy of the irreducible $L(\lambda)$ in V which was not included in V_0. This is a contradiction which finishes the proof. □

One corollary of complete reducibility is Fact III.16, which we stated earlier. We in fact get a strengthened version.

Proposition III.28 (Diagonalizability of the action of H)**.**
The action $\mathcal{H} = \vartheta(H)$ of H in any finite-dimensional representation V of $\mathfrak{sl}_2(\mathbb{C})$ is diagonalizable. The eigenvalues of $\mathcal{H}$ are integers.

Proof. In any irreducible representation $L(\lambda)$ we have explicitly seen in Theorem III.22 that $\mathcal{H}$ is diagonalizable, with integer eigenvalues. By Theorem III.27, any finite dimensional representation V is a direct sum of such irreducible representatios, so $\mathcal{H}$ is diagonalizable also on V and has integer eigenvalues. □

Clebsch-Gordan coefficients

A frequently important example of complete reducibility is the decomposition of tensor products of representations to direct sums of irreducible representations. The coefficients in such decompositions are known as **Clebsch-Gordan coefficients**. This is the topic of the following exercise.

Exercise III.29 (Tensor products of irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$)**.**
Consider the tensor product $L(\lambda_1) \otimes L(\lambda_2)$ of two irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$.

(a) Find the multiplicities of all $\mathcal{H}$-eigenvalues in $L(\lambda_1) \otimes L(\lambda_2)$.
(b) Given the complete reducibility of finite-dimensional representations of $\mathfrak{sl}_2(\mathbb{C})$, deduce from part (a) the multiplicities of all irreducible representations in the decomposition of $L(\lambda_1) \otimes L(\lambda_2)$ into a direct sum of irreducible subrepresentations.
(c) Find explicitly all vectors $v \in L(\lambda_1) \otimes L(\lambda_2)$ which satisfy $\mathcal{E}v = 0$ and $\mathcal{H}v = \lambda\, v$ for some λ.

Quadratic Casimir and Laplacian in radial coordinates

In applications to rotationally invariant quantum mechanical systems (such as the hydrogen atom) as well as in other questions about many familiar partial differential equations, one often wants to use radial coordinates for the Laplacian operator $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. This is closely related to the quadratic Casimir of $\mathfrak{sl}_2(\mathbb{C})$ (and of $\mathfrak{so}_3$).

Exercise III.30 (Laplacian in radial coordinates)**.**
Let $C^\infty(\mathbb{R}^3 \setminus \{0\})$ denote the space of smooth infinitely differentiable complex valued functions on $\mathbb{R}^3 \setminus \{0\}$, and on this space, consider the differential operators $\mathcal{J}_x, \mathcal{J}_y, \mathcal{J}_z$ as in Exercise II.79.

(a) Show that one can define a representation $\vartheta \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}\big(C^\infty(\mathbb{R}^3 \setminus \{0\})\big)$ of the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ by setting
$$\vartheta(E) \;=\; \mathbb{i}\,\mathcal{J}_x - \mathcal{J}_y, \qquad \vartheta(F) \;=\; \mathbb{i}\,\mathcal{J}_x + \mathcal{J}_y, \qquad H \;\mapsto\; 2\mathbb{i}\,\mathcal{J}_z,$$
and in this representation the operator $\mathcal{Q} = \vartheta(E)\vartheta(F) + \vartheta(F)\vartheta(E) + \frac{1}{2}\vartheta(H)^2$ studied in Exercise III.26 reads $\mathcal{Q} = -2 \times (\mathcal{J}_x^2 + \mathcal{J}_y^2 + \mathcal{J}_z^2)$.
(b) Define also the radial derivative operator $\mathcal{R} = \frac{x}{r}\frac{\partial}{\partial x} + \frac{y}{r}\frac{\partial}{\partial y} + \frac{z}{r}\frac{\partial}{\partial z}$, where $r(x,y,z) = \sqrt{x^2+y^2+z^2}$. Show that the Laplacian $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ admits the expression
$$\Delta = \frac{-1}{2r^2}\,\mathcal{Q} + \mathcal{R}^2 + \frac{2}{r}\,\mathcal{R}.$$

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

In the previous section we 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 $\mathfrak{sl}_3(\mathbb{C})$. Because $\mathfrak{sl}_3(\mathbb{C})$ is the complexification of the (real) Lie algebra $\mathfrak{su}_3$ of the Lie group SU_3, the representations of $\mathfrak{sl}_3(\mathbb{C})$ are needed for example in quantum chromodynamics (QCD) — the theory of strong interactions relevant for nuclear physics. Besides the direct relevance in such applications, 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 roughly 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.

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

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

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

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

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

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

We are also already familiar with a few representations of $\mathfrak{sl}_3(\mathbb{C})$.

Example III.31 (Trivial representation of $\mathfrak{sl}_3(\mathbb{C})$)**.**
As for any Lie algebra, the one-dimensional vector space $V = \mathbb{C}$ carries the trivial representation

$$\vartheta\colon\ \mathfrak{sl}_3(\mathbb{C}) \to \mathrm{End}(\mathbb{C}) \qquad\qquad \vartheta(X) = 0 \quad \forall X \in \mathfrak{sl}_3(\mathbb{C}).$$

Example III.32 (Adjoint representation of $\mathfrak{sl}_3(\mathbb{C})$)**.**
Also, according to a generality about Lie algebras, the vector space $V = \mathfrak{sl}_3(\mathbb{C})$ carries the adjoint representation

$$\begin{aligned} \mathrm{ad}\colon\ \mathfrak{sl}_3(\mathbb{C}) &\to \mathrm{End}(\mathfrak{sl}_3(\mathbb{C})) \\ \mathrm{ad}_X(Y) = [X, Y] &\qquad \forall X \in \mathfrak{sl}_3(\mathbb{C}),\ Y \in V = \mathfrak{sl}_3(\mathbb{C}). \end{aligned}$$

This representation is eight-dimensional, $\dim(V) = \dim\big(\mathfrak{sl}_3(\mathbb{C})\big) = 8$.

Example III.33 (Defining representation of $\mathfrak{sl}_3(\mathbb{C})$)**.**
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

$$\vartheta(X)\, v = Xv,$$

i.e., we set $\vartheta(X) = X \in \mathbb{C}^{3\times 3} \cong \mathrm{End}(V)$. This three-dimensional representation is called the defining representation of $\mathfrak{sl}_3(\mathbb{C})$.

It turns out that the dual V^* of the defining representation $V = \mathbb{C}^3$ provides another interesting example, but we return to this later.

3.2. Auxiliary calculations with elementary matrices

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,

$$E^{kl}_{ij} := \delta_{k,i}\, \delta_{l,j} = \begin{cases} 1 & \text{if } (i,j) = (k,l) \\ 0 & \text{if } (i,j) \neq (k,l). \end{cases} \tag{III.9}$$

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

$$\begin{aligned} [E^{kl}, E^{k'l'}] &= E^{kl} E^{k'l'} - E^{k'l'} E^{kl} \\ &= \delta_{l,k'}\, E^{kl'} - \delta_{l',k}\, E^{k'l}. \end{aligned} \tag{III.10}$$

3.3. Weights and weight spaces in representations of $\mathfrak{sl}_3(\mathbb{C})$

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 $\vartheta \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}(V)$ (including, crucially, the adjoint representation on $V = \mathfrak{sl}_2(\mathbb{C})$ itself) to eigenspaces of $\mathcal{H} := \vartheta(H)$, and figured out how $\mathcal{E} := \vartheta(E)$ and $\mathcal{F} := \vartheta(F)$ acted on the eigenspaces. Our 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 could succeed 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.,

$$\mathfrak{h} = \left\{ \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} \;\middle|\; a_1, a_2, a_3 \in \mathbb{C},\ a_1 + a_2 + a_3 = 0 \right\}. \tag{III.11}$$

All diagonal matrices indeed commute with each other, so $[\mathfrak{h}, \mathfrak{h}] = 0$, and the simultaneous diagonalization of the action of $\vartheta(H)$ for all $H \in \mathfrak{h}$ turns out to be possible (see Fact III.36 below).

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 a vector space V varries a representation $\vartheta\colon \mathfrak{sl}_3(\mathbb{C}) \to \mathrm{End}(V)$, and $v \in V$ is a simultaneous eigenvector for the action of all $H \in \mathfrak{h}$, then we have

$$\vartheta(H)v = \mu(H)v \qquad \text{for all } H \in \mathfrak{h}, \tag{III.12}$$

where $\mu(H) \in \mathbb{C}$ denotes the eigenvalue of the linear map $\vartheta(H)\colon V \to V$ corresponding to the element $H \in \mathfrak{h}$. Obviously $\mu(H)$ then has to depend 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. If a non-zero vector $0 \neq v \in V$ satisfying (III.12) exists in the representation V, we call v a **weight vector**, and we call $\mu \in \mathfrak{h}^*$ its **weight**. For any $\mu \in \mathfrak{h}^*$ we define the **weight space**

$$V_\mu := \left\{ v \in V \;\middle|\; \vartheta(H)v = \mu(H)v \quad \forall H \in \mathfrak{h} \right\}, \tag{III.13}$$

as the space of all weight vectors of weight μ, and weights of V as those $\mu \in \mathfrak{h}^*$ for which $V_\mu \neq \{0\}$.

Example III.34 (Weights in the defining representation of $\mathfrak{sl}_3(\mathbb{C})$).
The three-dimensional space $V = \mathbb{C}^3$ carries the defining representation $\vartheta\colon X \mapsto X$ of $\mathfrak{sl}_3(\mathbb{C})$, see Example III.33. Consider the standard basis vectors

$$u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

of $\mathbb{C}^3$. It is clear that for any diagonal matrix

$$H = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} = a_1 E^{11} + a_2 E^{11} + a_3 E^{33}$$

and any $j \in \{1,2,3\}$, we have

$$H u_j = a_j u_j.$$

Comparing with (III.12), this says that u_j is a weight vector — of weight $\eta^j \in \mathfrak{h}^*$ given by

$$\eta^j \big(a_1 E^{11} + a_2 E^{11} + a_3 E^{33}\big) := a_j. \tag{III.14}$$

The weights in the defining representation $V = \mathbb{C}^3$ of $\mathfrak{sl}_3(\mathbb{C})$ are thus $\eta^1, \eta^2, \eta^3 \in \mathfrak{h}^*$, with respective weight spaces are $V_{\eta^1} = \mathbb{C}u_1$, $V_{\eta^2} = \mathbb{C}u_2$, and $V_{\eta^3} = \mathbb{C}u_3$. The defining representation has the weight space decomposition

$$V = \mathbb{C}^3 = \mathbb{C}u_1 \oplus \mathbb{C}u_2 \oplus \mathbb{C}u_3 = V_{\eta^1} \oplus V_{\eta^2} \oplus V_{\eta^3}.$$

Note that the three weights $\eta^1, \eta^2, \eta^3 \in \mathfrak{h}^*$ defined above by (III.14) lie in the two-dimensional space $\mathfrak{h}^*$, and so can not be linearly independent. The one linear relation among them is $\eta^1 + \eta^2 + \eta^3 = 0$, coming from tracelessness, $0 = \mathrm{tr}(H) = a_1 + a_2 + a_3$, of diagonal matrices $H \in \mathfrak{h} \subset \mathfrak{sl}_3(\mathbb{C})$.

Example III.35 (Weights in the dual of defining representation of $\mathfrak{sl}_3(\mathbb{C})$).
Recall that if $\vartheta\colon \mathfrak{g} \to \mathrm{End}(V)$ is a representation of a Lie algebra $\mathfrak{g}$ on a vector space V, then the dual V^* carries a representation $\vartheta^*\colon \mathfrak{g} \to \mathrm{End}(V^*)$ such that for any $X \in \mathfrak{g}$ and $\varphi \in V^*$, the dual element $\vartheta^*(X)\varphi$ is the functional $v \mapsto -\varphi\big(\vartheta(X)v\big)$ on V.

The dual V^* of the defining 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 $u_1, u_2, u_3 \in V$, i.e., $\varphi_j(u_i) = \delta_{i,j}$ for all $i, j \in \{1,2,3\}$. If $H \in \mathfrak{h}$, then

$$\langle \vartheta^*(H)\varphi_j, u_i \rangle = -\langle \varphi_j, H u_i \rangle = -\langle \varphi_j, \eta^i(H) u_i \rangle = -\eta^i(H)\,\delta_{i,j} = -\eta^j(H)\,\langle \varphi_j, u_i \rangle,$$

which implies that $\vartheta^*(H)\varphi_j = -\eta^j(H)\,\varphi_j$. 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 V^* of the defining representation $V = \mathbb{C}^3$ 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).

Any finite-dimensional representation has a weight space decomposition like in the examples above, i.e., the desired simultaneous diagonalization is always possible. For now, let us state this as a fact (analogous to Fact III.16 for $\mathfrak{sl}_2(\mathbb{C})$).

Fact III.36. Any finite-dimensional representations V of $\mathfrak{sl}_3(\mathbb{C})$ decomposes to a (vector space) direct sum of the weight spaces[6],

$$V = \bigoplus_{\mu \in \mathfrak{h}^*} V_\mu.$$

The proof of this fact is not difficult given what we know of representations of $\mathfrak{sl}_2(\mathbb{C})$. We will see this a little later.

We will next address the weight space decomposition on the adjoint representation, which turns out to be particularly consequential.

3.4. Roots and root spaces of $\mathfrak{sl}_3(\mathbb{C})$

Let us consider the adjoint representation on $V = \mathfrak{sl}_3(\mathbb{C})$ of $\mathfrak{sl}_3(\mathbb{C})$. According to Fact III.36, the adjoint representation also admits a decomposition to weight spaces

$$\mathfrak{sl}_3(\mathbb{C}) = \bigoplus_{\mu} \big(\mathfrak{sl}_3(\mathbb{C})\big)_\mu,$$

and we will verify this explicitly now.

Recall that we took $\mathfrak{h} \subset \mathfrak{sl}_3(\mathbb{C})$ to consist of the diagonal matrices in $\mathfrak{sl}_3(\mathbb{C})$. Diagonal matrices commute with each other, i.e., we have

$$\mathrm{ad}_H(H') = [H, H'] = 0 \qquad \text{for all } H, H' \in \mathfrak{h}.$$

This means that all diagonal matrices $H' \in \mathfrak{h}$ are weight vectors of zero weight $0 \in \mathfrak{h}^*$,

$$\mathfrak{h} \subset \big(\mathfrak{sl}_3(\mathbb{C})\big)_0.$$

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 $\mathfrak{h}$.

As a basis for the rest of $\mathfrak{sl}_3(\mathbb{C})$, we can use off-diagonal elementary matrices E^{ij}, $i \neq j$. These are also weight vectors, since for any diagonal matrix $H = \sum_k a_k E^{kk}$,

[6] There are, of course, only finitely many non-zero direct summands in this decomposition to weight spaces — corresponding exactly the finitely many weights μ of the finite-dimensional representation V.

we have

$$\begin{aligned}[H, E^{ij}] &= \sum_k a_k \, [E^{kk}, E^{ij}] = \sum_k a_k \, (\delta_{ki} E^{kj} - \delta_{jk} E^{ik}) \\ &= (a_i - a_j) \, E^{ij}, \end{aligned} \tag{III.15}$$

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:

$$\mathfrak{sl}_3(\mathbb{C}) = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{C}E^{ij}. \tag{III.16}$$

The non-zero weights appearing in the adjoint representation are called **roots**, and the weight spaces in the adjoint representation other than $\mathfrak{h}$ are correspondingly called **root spaces**. The set of roots is denoted by $\Phi \subset \mathfrak{h}^*$: for $\mathfrak{sl}_3(\mathbb{C})$ we have

$$\Phi = \left\{ \eta^1 - \eta^2, \eta^1 - \eta^3, \eta^2 - \eta^3, \eta^2 - \eta^1, \eta^3 - \eta^1, \eta^3 - \eta^2 \right\}. \tag{III.17}$$

Roots are traditionally denoted by α, so let us introduce the notation

$$\alpha^{ij} := \eta^i - \eta^j \in \Phi \qquad \text{for } i \neq j. \tag{III.18}$$

In this notation, (III.15) says that for any $H \in \mathfrak{h}$ and $i \neq j$

$$[H, E^{ij}] = \alpha^{ij}(H) \, E^{ij}. \tag{III.19}$$

Shifting weights by roots

Consider a representation $\vartheta \colon \mathfrak{sl}_3(\mathbb{C}) \to \mathrm{End}(V)$ of $\mathfrak{sl}_3(\mathbb{C})$ on a finite-dimensional vector space V. According to Fact III.36, the space V decomposes to a direct sum of weight spaces

$$V = \bigoplus_{\mu \in \mathfrak{h}^*} V_\mu,$$

and this exactly describes the actions $\mathcal{H} = \vartheta(H)$ of all elements $H \in \mathfrak{h}$ in the abelian subalgebra of diagonal matrices. To understand how the remaining basis elements act, denote by $\mathcal{E}^{ij} := \vartheta(E^{ij})$, for $i \neq j$.

The following lemma describes what $\mathcal{E}^{ij}$ does to the weight spaces V_μ.

Lemma III.37 (Shifting weights by roots)**.**
For any $\mu \in \mathfrak{h}^*$ and any i, j with $i \neq j$, we have

$$\mathcal{E}^{ij}(V_\mu) \subset V_{\mu + \alpha^{ij}}.$$

Proof. Suppose that $v \in V_\mu$. Let $H \in \mathfrak{h}$, and denote by $\mathcal{H} = \vartheta(H)$ the its action on V. By definition of weight spaces, we have $\mathcal{H}v = \mu(H) \, v$. Then consider the vector $\mathcal{E}^{ij} v \in V$. We can find the action of $\mathcal{H}$ on v using the commutator of $\mathcal{H}$ and $\mathcal{E}^{ij}$

$$\mathcal{H}\mathcal{E}^{ij} - \mathcal{E}^{ij}\mathcal{H} = \alpha^{ij}(H) \, \mathcal{E}^{ij},$$

corresponding to the Lie bracket $[H, E^{ij}]$ in (III.19).

Fundamental calculation (second time):

$$\begin{aligned}\mathcal{H}(\mathcal{E}^{ij}v) &= \mathcal{E}^{ij}(\mathcal{H}v) + [\mathcal{H}, \mathcal{E}^{ij}]v \\ &= \mathcal{E}^{ij}(\mu(H)v) + \alpha^{ij}(H)\,\mathcal{E}^{ij}v \\ &= (\mu + \alpha^{ij})(H)\,\mathcal{E}^{ij}v.\end{aligned}$$

This calculation shows that if v is a weight vector of weight $\mu \in \mathfrak{h}^*$, then $\mathcal{E}^{ij}v$ is either zero or a weight vector of weight $\mu + \alpha^{ij}$. This proves that $\mathcal{E}^{ij}(V_\mu) \subset V_{\mu+\alpha^{ij}}$. □

Root lattice

It follows from Lemma III.37 that in an irreducible representation V of $\mathfrak{sl}_3(\mathbb{C})$, the weights differ from each other by integer multiples of the roots α^{ij}, $i \neq j$. This can be reformulated as saying that the weights in an irreducible lie in some translate of the **root lattice**

$$\Lambda_{\mathrm{R}} = \sum_{\alpha \in \Phi} \mathbb{Z}\alpha. \tag{III.20}$$

Observing that $\alpha^{ji} = -\alpha^{ij}$ and $\alpha^{13} = \alpha^{12} + \alpha^{23}$, one sees that integer linear combinations of α^{12} and α^{23} suffice, and the root lattice can be written as

$$\Lambda_{\mathrm{R}} = \mathbb{Z}\alpha^{12} \oplus \mathbb{Z}\alpha^{23}.$$

We call 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 $\Phi^+ = \{\alpha^{12}, \alpha^{23}, \alpha^{13}\} = \{\alpha^{ij} \,|\, i < j\}$.

Weight lattice

Let us now find necessary conditions on the weights, based on what we know from Section 2 about representations of $\mathfrak{sl}_2(\mathbb{C})$.

Denote $H^{12} = E^{11} - E^{22} \in \mathfrak{h}$, and recall calculations (III.10) and (III.15), 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 above three elements

$$\mathfrak{s}^{12} := \operatorname{span}\left\{H^{12}, E^{12}, E^{21}\right\} \subset \mathfrak{sl}_3(\mathbb{C}) \tag{III.21}$$

is a Lie subalgebra, which is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$ through $H \mapsto H^{12}$, $E \mapsto E^{12}$, $F \mapsto E^{21}$.

Let V be a finite-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$, and consider its weight spaces $V_\mu \subset V$. Note that we can consider V as a representation of the Lie subalgebra $\mathfrak{s}^{12} \subset \mathfrak{sl}_3(\mathbb{C})$. From in the previous section, we know that the eigenvalues of the action of H, now $\mathcal{H}^{12} := \vartheta(H^{12})$, are integers. In particular, if $\mu \in \mathfrak{h}^*$ is a weigth of V, then for a non-zero vector $v \in V_\mu$ we have

$$\mathcal{H}^{12}v = \mu(H^{12})v,$$

and Proposition III.28 therefore implies $\mu(H^{12}) \in \mathbb{Z}$.

Similarly, E^{23}, E^{32}, and $H^{23} = E^{22} - E^{33} \in \mathfrak{h}$ satisfy

$$[H^{23}, E^{23}] = 2\,E^{23}, \qquad [H^{23}, E^{32}] = -2\,E^{32}, \qquad [E^{23}, E^{32}] = H^{23},$$

so also the span of the three elements

$$\mathfrak{s}^{23} := \mathrm{span}\left\{H^{23}, E^{23}, E^{32}\right\} \subset \mathfrak{sl}_3(\mathbb{C}) \tag{III.22}$$

is a Lie subalgebra isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. One correspondingly argues that any weight $\mu \in \mathfrak{h}^*$ has to satisfy $\mu(H^{23}) \in \mathbb{Z}$.

Let us therefore define the weight lattice $\Lambda_{\mathrm{W}} \subset \mathfrak{h}^*$ of $\mathfrak{sl}_3(\mathbb{C})$ by

$$\Lambda_{\mathrm{W}} = \left\{\mu \in \mathfrak{h}^* \;\middle|\; \mu(H^{12}), \mu(H^{23}) \in \mathbb{Z}\right\}. \tag{III.23}$$

One possible basis of $\mathfrak{h}^*$ consists of η^1 and $-\eta^3$. Let us write $\mu = a\eta^1 - b\eta^3$ with coefficients $a, b \in \mathbb{C}$. Note that $\eta^1(H^{12}) = 1$, $\eta^2(H^{12}) = -1$, $\eta^3(H^{12}) = 0$, and similarly $\eta^1(H^{23}) = 0$, $\eta^2(H^{23}) = 1$, $\eta^3(H^{23}) = -1$. With these, we calculate $\mu(H^{12}) = a$ and $\mu(H^{23}) = b$. This calculation gives a concrete expression for the weight lattice

$$\Lambda_{\mathrm{W}} = \left\{a\eta^1 - b\eta^3 \;\middle|\; a, b \in \mathbb{Z}\right\}.$$

Since roots are weights in the adjoint representation, the root lattice generated by roots is a subset of the weight lattice,

$$\Lambda_{\mathrm{R}} \subset \Lambda_{\mathrm{W}} \subset \mathfrak{h}^*.$$

It is easy to see that the inclusion is strict: for example the weights η^1, η^2, η^3 in the defining representation $V = \mathbb{C}^3$ are not integer linear combinations of roots $\alpha \in \Phi$.

Exercise III.38 (Proof of Fact III.36)**.**
Prove Fact III.36 using the Lie subalgebras $\mathfrak{s}^{12}$ and $\mathfrak{s}^{23}$, and representation theory of $\mathfrak{sl}_2(\mathbb{C})$.

3.5. Highest weight vectors for $\mathfrak{sl}_3(\mathbb{C})$

Let again V be a finite-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$. The decomposition $V = \bigoplus_{\mu \in \mathfrak{h}^*} V_\mu$ to weight spaces V_μ tells exactly how any $H \in \mathfrak{h}$ acts on V, and Lemma III.37 tells what the actions $\mathcal{E}^{ij}$ of the remaining basis elements E^{ij}, $i \neq j$, do to the weight spaces.

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 was annihilated by the action of E, 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, based on our choice of simple roots α^{12}, α^{23}: we consider a weight $\mu \in \mathfrak{h}^*$ of the representation V maximal if neither $\mu + \alpha^{12}$ nor $\mu + \alpha^{23}$ are weights.

Corollary III.39 (Existence of highest weight vectors for $\mathfrak{sl}_3(\mathbb{C})$)**.**
In any finite-dimensional representation $V \neq \{0\}$ of $\mathfrak{sl}_3(\mathbb{C})$, there exists a non-zero vector $v_0 \in V$ such that $\mathcal{E}^{ij} v_0 = 0$ whenever $i < j$, and $v_0 \in V_\lambda$ for some $\lambda \in \mathfrak{h}^*$.

Proof. Since the finite-dimensional representation V has finitely many different weights, we can find a weight $\lambda \in \mathfrak{h}^*$ such that neither $\lambda + \alpha^{12}$ nor $\lambda + \alpha^{23}$ are weights. Choose a non-zero vector $v_0 \in V_\lambda$. By Lemma III.37 we then have

$$\mathcal{E}^{12} v_0 \in V_{\lambda+\alpha^{12}} = \{0\} \qquad \text{and} \qquad \mathcal{E}^{23} v_0 \in V_{\lambda+\alpha^{23}} = \{0\},$$

which gives

$$\mathcal{E}^{12} v_0 = 0 \qquad \text{and} \qquad \mathcal{E}^{23} v_0 = 0.$$

The commutation relation $[E^{12}, E^{23}] = E^{13}$ then yields also

$$\mathcal{E}^{13} v_0 = \big(\mathcal{E}^{12}\mathcal{E}^{23} - \mathcal{E}^{23}\mathcal{E}^{12}\big) v_0 = 0.$$

The vector v_0 therefore satisfies all the asserted properties. □

Definition III.40 (Highest weight vector in representation of $\mathfrak{sl}_3(\mathbb{C})$)**.**
Let V is any representation of $\mathfrak{sl}_3(\mathbb{C})$. A **highest weight vector** of **highest weight** $\lambda \in \mathfrak{h}^*$ is a non-zero vector $v_0 \in V_\lambda \subset V$, which satisfies $\mathcal{E}^{ij} v_0 = 0$ for all $i < j$.

Example III.41 (Highest weight vector in the defining representation of $\mathfrak{sl}_3(\mathbb{C})$)**.**
In the defining representation $V = \mathbb{C}^3$ of $\mathfrak{sl}_3(\mathbb{C})$, the vector u_1 is a highest weight vector of highest weight $\lambda = \eta^1$. Indeed, neither $\eta^1 + \alpha^{12} = 2\eta^1 - \eta^2$ nor $\eta^1 + \alpha^{23} = \eta^1 + \eta^2 - \eta^3$ are weights in the defining representation $V = \mathbb{C}^3$ according to Example III.34. Alternatively, one could directly observe that $E^{12} u_1 = 0$ and $E^{23} u_1 = 0$.

Example III.42 (Highest weight vector in the dual of defining representation of $\mathfrak{sl}_3(\mathbb{C})$)**.**
In the dual V^* of the defining representation $V = \mathbb{C}^3$ of $\mathfrak{sl}_3(\mathbb{C})$, the vector φ_3 is a highest weight vector of highest weight $\lambda = -\eta^3$. Indeed, neither $-\eta^3 + \alpha^{12} = \eta^1 - \eta^2 - \eta^3$ nor $-\eta^3 + \alpha^{23} = \eta^2 - 2\eta^3$ are weights in the dual V^* of the defining representation, according to Example III.35.

Example III.43 (Highest weight vector in the adjoint representation of $\mathfrak{sl}_3(\mathbb{C})$)**.**
In the adjoint representation $\mathfrak{sl}_3(\mathbb{C})$, by Equations (III.10) and (III.15), the vector E^{13} is a highest weight vector of highest weight $\lambda = \alpha^{13} = \eta^1 - \eta^3$.

Necessary conditions for highest weights

Again using the two Lie subalgebras $\mathfrak{s}^{12} \subset \mathfrak{sl}_3(\mathbb{C})$ and $\mathfrak{s}^{23} \subset \mathfrak{sl}_3(\mathbb{C})$ we find necessary conditions for a weight to qualify as a highest weight.

Lemma III.44 (Necessary conditions for highest weights)**.**
The highest weight $\lambda \in \mathfrak{h}^*$ of a highest weight vector v_0 in a finite-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$ takes non-negative integer values on the basis H^{12}, H^{23} of $\mathfrak{h} \subset \mathfrak{sl}_3(\mathbb{C})$:

$$\lambda(H^{12}) = a \in \mathbb{Z}_{\geq 0}, \quad \lambda(H^{23}) = b \in \mathbb{Z}_{\geq 0}.$$

Proof. Viewing V as a representation of the Lie subalgebra $\mathfrak{s}^{12} \subset \mathfrak{sl}_3(\mathbb{C})$, the vector $v_0 \in V$ becomes a highest weight vector for the representation of $\mathfrak{sl}_2(\mathbb{C})$, with highest weight $\lambda(H^{12})$. By Corollary III.21 we must have $\lambda(H^{12}) \in \mathbb{Z}_{\geq 0}$. Similarly using the Lie subalgebra $\mathfrak{s}^{23} \subset \mathfrak{sl}_3(\mathbb{C})$ one argues that $\lambda(H^{23}) \in \mathbb{Z}_{\geq 0}$. □

If we write $\lambda = a\eta^1 - b\eta^3$, then the conditions above read $a, b \in \mathbb{Z}_{\geq 0}$. The highest weights in Examples III.41, III.42, and III.43 were explicitly of this form.

Subrepresentation generated by highest weight vector

A highest weight vector v_0 is annihilated by elements in half of the root spaces, the "raising operators" $\mathcal{E}^{ij}$, $i < j$, which shift the weights to the directions of the positive roots $\alpha^{ij} \in \Phi^+$. Like for $\mathfrak{sl}_2(\mathbb{C})$, applying repeatedly on it the elements in the other half of the root spaces, the "lowering operators" $\mathcal{E}^{ji}$, $i < j$, we generate a subrepresentation.

Proposition III.45 (Subrepresentation generated by highest weight vector)**.**
Suppose that v_0 is a highest weight vector in a representation V of $\mathfrak{sl}_3(\mathbb{C})$. Let W denote the subspace spanned by the the vectors obtained by successively applying $\mathcal{E}^{21}$, $\mathcal{E}^{32}$, and $\mathcal{E}^{31}$ on v_0. Then $W \subset V$ is a subrepresentation.

Proof. Note that we have $E^{31} = [E^{32}, E^{21}]$, and consequently $\mathcal{E}^{31} = \mathcal{E}^{32}\mathcal{E}^{21} - \mathcal{E}^{21}\mathcal{E}^{32}$. Thus W could have been alternatively defined as the linear span of vectors obtained by successively applying only $\mathcal{E}^{21}$ and $\mathcal{E}^{32}$ on v_0. For an inductive argument, let W_ℓ denote the linear span of vectors obtained by successively applying on v_0 a word of at most ℓ letters, each equal to $\mathcal{E}^{21}$ or $\mathcal{E}^{32}$. Then W is the sum of W_ℓ, as ℓ ranges over natural numbers. By construction we have $\mathcal{E}^{21}W_\ell \subset W_{\ell+1}$ and $\mathcal{E}^{32}W_\ell \subset W_{\ell+1}$, and then using the fact that $\mathcal{E}^{31} = \mathcal{E}^{32}\mathcal{E}^{21} - \mathcal{E}^{21}\mathcal{E}^{32}$, we get that $\mathcal{E}^{31}W_\ell \subset W_{\ell+2}$. We moreover claim that for any $H \in \mathfrak{h}$ the operator $\mathcal{H} = \vartheta(H)$ preserves these subspaces in the sense that $\mathcal{H}W_\ell \subset W_\ell$. Indeed, the vector obtained by applying a word on the highest weight vector is itself 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_ℓ. It follows that $W = \sum_\ell W_\ell$ 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 elements E^{12}, E^{23}, and E^{13} do to W_ℓ. Moreover, since $E^{13} = [E^{12}, E^{23}]$, it in fact suffices to consider E^{12} and E^{23}.

We claim that $\mathcal{E}^{12}W_\ell \subset W_{\ell-1}$ and $\mathcal{E}^{23}W_\ell \subset W_{\ell-1}$. The proofs are entirely similar, so let us consider only the first of these. The case $\ell = 0$ is clear, since $W_0 = \mathbb{C}v_0$ is the one-dimensional space spanned by the highest weight vector, which is annihilated by $\mathcal{E}^{12}$ (and $\mathcal{E}^{23}$). Proceed by induction on ℓ. Suppose that w is a vector obtained by applying on v_0 a word of ℓ letters, each equal to $\mathcal{E}^{21}$ or $\mathcal{E}^{32}$. Depending on the last letter, we have either $w = \mathcal{E}^{21}w'$ or $w = \mathcal{E}^{32}w'$, with $w' \in W_{\ell-1}$. In the first case we have

$$\begin{aligned}\mathcal{E}^{12}w = \mathcal{E}^{12}\mathcal{E}^{21}w' &= \big(\mathcal{E}^{21}\mathcal{E}^{12} + [\mathcal{E}^{12}, \mathcal{E}^{21}]\big)w' = \big(\mathcal{E}^{21}\mathcal{E}^{12} + \mathcal{H}^{12}\big)w' \\ &= \mathcal{E}^{21}\mathcal{E}^{12}w' + \mathcal{H}^{12}w' \;\in\; \mathcal{E}^{21}W_{\ell-2} + W_{\ell-1} \subset W_{\ell-1},\end{aligned}$$

where we used the induction assumption $\mathcal{E}^{12}W_{\ell-1} \subset W_{\ell-2}$ and the fact that elements of $\mathfrak{h}$ preserve $W_{\ell-1}$. In the second case we have

$$\begin{aligned}\mathcal{E}^{12}w = \mathcal{E}^{12}\mathcal{E}^{32}w' &= \big(\mathcal{E}^{32}\mathcal{E}^{12} + [\mathcal{E}^{12}, \mathcal{E}^{32}]\big)w' = \big(\mathcal{E}^{32}\mathcal{E}^{12} + 0\big)w' \\ &= \mathcal{E}^{32}\mathcal{E}^{12}w' \;\in\; \mathcal{E}^{32}W_{\ell-2} \subset W_{\ell-1},\end{aligned}$$

where we again used the induction assumption $\mathcal{E}^{12}W_{\ell-1} \subset W_{\ell-2}$. By induction, we thus establish that $\mathcal{E}^{12}W_\ell \subset W_{\ell-1}$ and $\mathcal{E}^{23}W_\ell \subset W_{\ell-1}$. By considering a commutator, it follows that also $\mathcal{E}^{13}W_\ell \subset W_{\ell-2}$.

We have thus seen that the subspace $W = \sum_\ell W_\ell$ is invariant also for the action of the elements E^{12}, E^{23}, and E^{13}. It is therefore a subrepresentation. □

The case when a highest weight vector generates the entire representation is of particular interest. Note that this is the case for irreducible representations, in particular. We use the term highest weight representation to refer to such a situation.

Definition III.46 (Highest weight representations of $\mathfrak{sl}_3(\mathbb{C})$)**.**
If a representation V of $\mathfrak{sl}_3(\mathbb{C})$ contains a highest weight vector v_0 of highest weight $\lambda \in \mathfrak{h}^*$, and is generated by this highest weight vector in the sense that

the subrepresentation W descibed in Proposition III.45 is the whole representation, $W = V$, then we call V a **highest weight representation** of **highest weight** λ.

It is not difficult to check that the representations in Examples III.41, III.42, and III.43 are highest weight representations.

Proposition III.45 has particularly nice corollaries for irreducible representations.

Corollary III.47 (Weights in an irreducible representation lie in a cone)**.**
Let V be a finite-dimensional irreducible representation of $\mathfrak{sl}_3(\mathbb{C})$. Then the weights μ of V lie in a cone

$$\lambda - (\mathbb{R}_{\geq 0}\alpha^{12} + \mathbb{R}_{\geq 0}\alpha^{23})$$

seen from the highest weight λ of any highest weight vector $v_0 \in V$. Moreover, the weight space V_λ is one-dimensional.

Proof. Let $v_0 \in V_\lambda$ be a highest weight vector in V. The subspace $W \subset V$ spanned by applications of $\mathcal{E}^{21}$, $\mathcal{E}^{31}$, and $\mathcal{E}^{32}$ on v_0 is a subrepresentation by Proposition III.45. By irreducibility we must have $W = V$. By Lemma III.37, on the other hand, the weights μ of such spanning vectors of W are of the form $\mu = \lambda + n_{21}\alpha^{21} + n_{31}\alpha^{31} + n_{32}\alpha^{32}$ with $n_{21}, n_{31}, n_{32} \in \mathbb{Z}_{\geq 0}$. But we have $\alpha^{31} = \alpha^{21} + \alpha^{32}$ as well as $\alpha^{21} = -\alpha^{12}$ and $\alpha^{32} = -\alpha^{23}$, so this implies $\mu \in \lambda - (\mathbb{Z}_{\geq 0}\alpha^{12} + \mathbb{Z}_{\geq 0}\alpha^{23})$. In view of $W = V$, this implies the assertion about the weights in V.

Note also that all of the spanning vectors of W except from v_0 itself have a weights $\mu \neq \lambda$. This implies one-dimensionality of the weight space $\dim(V_\lambda) = 1$ corresponding to the highest weight. □

Corollary III.48 (Uniqueness of highest weight of an irreducible representation)**.**
An irreducible finite-dimensional representation V of $\mathfrak{sl}_3(\mathbb{C})$ there exists a unique $\lambda \in \mathfrak{h}^*$ which is a highest weight of a highest weight vector in V.

Proof. Suppose that $\lambda, \lambda' \in \mathfrak{h}^*$ are both highest weights of some highest weight vectors in V. By Corollary III.47 we have $\lambda' \in \lambda - (\mathbb{R}_{\geq 0}\alpha^{12} + \mathbb{R}_{\geq 0}\alpha^{23})$ and $\lambda \in \lambda' - (\mathbb{R}_{\geq 0}\alpha^{12} + \mathbb{R}_{\geq 0}\alpha^{23})$. Since $\alpha^{12}, \alpha^{23} \in \mathfrak{h}^*$ are linearly independent, this is only possible if in fact $\lambda = \lambda'$.

□

3.6. Irreducible representations of $\mathfrak{sl}_3(\mathbb{C})$

Consider an irreducible finite-dimensional representation V of $\mathfrak{sl}_3(\mathbb{C})$. By Corollary III.39, there exists a highest weight vector in V, by Corollary III.48 the highest weight $\lambda \in \mathfrak{h}^*$ of such a highest weight vector is unique, and by Corollary III.47 the corresponding weight space V_λ is one-dimensional, so that the highest weight vector is unique up to (non-zero) scalar multiples. Next we show that the highest weight λ determines the irreducible representation up to isomorphism.

Lemma III.49 (Highest weights uniquely determine irreducible representations)**.**
Suppose that V and $\widetilde{V}$ are two irreducible finite-dimensional representations of $\mathfrak{sl}_3(\mathbb{C})$ with the same highest weight $\lambda \in \mathfrak{h}^*$. Then V and $\widetilde{V}$ are isomorphic, $V \cong \widetilde{V}$.

Proof. Let $v_0 \in V_\lambda$ and $\tilde{v}_0 \in \widetilde{V}_\lambda$ be highest weight vectors in the two irreducible representations V and $\widetilde{V}$, respectively. Consider the direct sum $V \oplus \widetilde{V}$. The vector $v_0 + \tilde{v}_0 \in V \oplus \widetilde{V}$ is annihilated by the action of E^{ij}, $i < j$, and is a weight vector of weight λ, i.e., it is a highest weight vector of highest weight λ in $V \oplus \widetilde{V}$. Let $W \subset V \oplus \widetilde{V}$ denote the subrepresentation generated by $v_0 + \tilde{v}_0$. Note that its weight space W_λ corresponding to the highest weight is one-dimensional (by the same argument as in Corollary III.47).

Consider the projection $\varpi \colon W \to V$ defined by $\varpi(v + \tilde{v}) = v$ whenever $v \in V$ and $\tilde{v} \in \widetilde{V}$. This projection is clearly an intertwining map, and non-zero since $\varpi(v_0 + \tilde{v}_0) = v_0 \neq 0$. By irreducibility of V it must therefore be surjective, $\mathrm{Im}(\varpi) = V$. The kernel of ϖ is clearly $\mathrm{Ker}(\varpi) = W \cap \widetilde{V}$. Note that $W \cap \widetilde{V}$ is a subrepresentation of $\widetilde{V}$, so by irreducibility of $\widetilde{V}$ it is either $\{0\}$ or $\widetilde{V}$. The case $W \cap \widetilde{V} = \widetilde{V}$ is impossible, for it would imply that $\tilde{v}_0 \in \mathrm{Ker}(\varpi) \subset W$, and then the one-dimensional weight space W_λ would contain the linearly independent vectors $v_0 + \tilde{v}_0$ and $\tilde{v}_0$. Thus we have $\mathrm{Ker}(\varpi) = W \cap \widetilde{V} = \{0\}$, so $\varpi \colon W \to V$ is injective. We have shown that ϖ is an isomorphism of W and V, so $V \cong W$.

Exactly similarly one shows that $\widetilde{V} \cong W$. Combining the two isomorphisms, we conclude that $V \cong \widetilde{V}$. □

We now know that irreducible finite-dimensional representations of $\mathfrak{sl}_3(\mathbb{C})$ are uniquely determined by their highest weights, and we know that these highest weights $\lambda \in \mathfrak{h}^*$ must satisfy $\lambda(H^{12}) \in \mathbb{Z}_{\geq 0}$ and $\lambda(H^{23}) \in \mathbb{Z}_{\geq 0}$, i.e., we must have

$$\lambda = \lambda_{a,b} := a\eta^1 - b\eta^3 \qquad \text{for } a, b \in \mathbb{Z}_{\geq 0}.$$

The remaining task in our classification of irreducible representations is to show that irreducible highest weight representations of all such highest weights exist. The following example is helpful for that purpose.

Example III.50 (Highest weight vectors of highest weights $\lambda_{a,b}$)**.**
Consider the defining representation $V = \mathbb{C}^3$ of $\mathfrak{sl}_3(\mathbb{C})$ and its dual V^*. Recall from Example III.41 that $u_1 \in V$ is a highest weight vector of highest weight $\eta^1 = \lambda_{1,0}$, and recall from Example III.42 that $\varphi_3 \in V^*$ is a highest weight vector of highest weight $-\eta^3 = \lambda_{0,1}$.

Now let $a, b \in \mathbb{Z}_{\geq 0}$, and 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

$$v_0 = u_1 \otimes \cdots \otimes u_1 \otimes \varphi_3 \otimes \cdots \otimes \varphi_3$$

is annihilated by the action of E^{12}, E^{23}, and E^{13}, and it is an eigenvector of the action of any $H \in \mathfrak{h}$, with eigenvalue $a\,\eta^1(H) - b\,\eta^3(H)$. Therefore, v_0 is a highest weight vector of highest weight $\lambda_{a,b}$, and the subrepresentation generated by v_0 in $V^{\otimes a} \otimes (V^*)^{\otimes b}$ is a highest weight representation of highest weight $\lambda_{a,b}$. Note also that dimension of this highest weight representation is at most 3^{a+b}, the dimension of the tensor product.

Theorem III.51 (Irreducible representations of $\mathfrak{sl}_3(\mathbb{C})$)**.**
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 an irreducible finite-dimensional representation $L(\lambda_{a,b})$ of $\mathfrak{sl}_3(\mathbb{C})$ with highest weight $\lambda_{a,b}$, and such a representation is unique up to isomorphism. 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. Corollaries III.39 and III.48 imply that an irreducible finite-dimensional representation is a highest weight representation of a uniquely determined highest weight $\lambda \in \mathfrak{h}^*$. Lemma III.44 shows that $\lambda = \lambda_{a,b}$ for some $a, b \in \mathbb{Z}_{\geq 0}$. Uniqueness of an irreducible highest weight representation of highest weight $\lambda_{a,b}$ was shown in Lemma III.49. It remains to show existence.

In Example III.50 we showed that there exist finite-dimensional highest weight representations of highest weights $\lambda_{a,b}$, $a, b \in \mathbb{Z}_{\geq 0}$. A standard argument then implies the existence of corresponding irreducible highest weight representations: the quotient of a highest weight representation by its maximal proper subrepresentation must be both irreducible and a highest weight representation of the same highest weight. The quotient of a finite-dimensional represenation is moreover finite-dimensional. This finishes the proof. □

We finally state the complete reducibility for representations of $\mathfrak{sl}_3(\mathbb{C})$. The proof is best done through general theory, and we omit it for the moment.

Fact III.52 (Complete reducibility for representations of $\mathfrak{sl}_3(\mathbb{C})$)**.**
Any finite-dimensional representation V of $\mathfrak{sl}_3(\mathbb{C})$ is a direct sum of its irreducible subrepresentations, in particular

$$V \;\cong\; L(\lambda_{a_1,b_1}) \oplus \cdots \oplus L(\lambda_{a_s,b_s})$$

for some s and $a_1, \ldots, a_s, b_1, \ldots, b_s \in \mathbb{Z}_{\geq 0}$.

Given complete reducibility we can in fact easily show that any finite-dimensional highest weight representation is automatically irreducible. In particular, the subrepresentation $W \subset V$ described in Proposition III.45 is irreducible if V is finite-dimensional.

Lemma III.53 (Irreducibility of highest weight representations)**.**
Any finite-dimensional highest weight representation V of $\mathfrak{sl}_3(\mathbb{C})$ is irreducible.

Proof. Suppose, by contrapositive, that a finite-dimensional highest weight representation V of highest weight λ is not irreducible. By complete reducibility, it is therefore the direct sum of two non-zero subrepresentations $\{0\} \neq W \subset V$ and $\{0\} \neq \widetilde{W} \subset V$,

$$V = W \oplus \widetilde{W}.$$

The weight space V_λ is one-dimensional by Corollary III.47: in concrete terms $V_\lambda = \mathbb{C}v_0$ for a highest weight vector v_0. The weight spaces of W and $\widetilde{W}$ are given by $W_\mu = V_\mu \cap W$ and $\widetilde{W}_\mu = V_\mu \cap \widetilde{W}$, and we have $V_\mu = W_\mu \oplus \widetilde{W}_\mu$. One-dimensionality of V_λ thus implies that either W_λ or $\widetilde{W}_\lambda$ is zero. Then the highest weight vector v_0 must lie in the other one, say $v_0 \in W_\lambda \subset W$. But the highest weight vector v_0 generates the whole representation V, so we get $W = V$, a contradiction. □

4. Lie algebras of compact type

We now proceed to more general structure theory and representation theory of (semisimple) Lie algebras. The analysis of $\mathfrak{sl}_3(\mathbb{C})$ in Section 3 serves as the model case — almost everything generalizes quite directly.

The most standard way would be to define semisimple Lie algebras, and study them purely algebraically. We choose a different route: we add one assumption, which is valid for the Lie algebra of any compact Lie group. With this assumption, some proofs become shorter. Moreover, it could be seen a posteriori that all complex semisimple Lie algebras are obtained as complexifications of the ones that satisfy this assumption, so in a way the simplifications come without loss of generality (recall that we anyway study representations of Lie algebras via their complexifications).

4.1. Algebraic notions

Definition III.54 (Lie subalgebra)**.**
A **Lie subalgebra** of a Lie algebra $\mathfrak{g}$ is a vector subspace $\mathfrak{s} \subset \mathfrak{g}$ such that $[\mathfrak{s}, \mathfrak{s}] \subset \mathfrak{s}$.

Definition III.55 (Ideal in a Lie algebra)**.**
An *ideal* of a Lie algebra $\mathfrak{g}$ is a vector subspace $\mathfrak{j} \subset \mathfrak{g}$ such that $[\mathfrak{j}, \mathfrak{g}] \subset \mathfrak{j}$.

Definition III.56 (Simple Lie algebra)**.**
A Lie algebra $\mathfrak{g}$ is **simple** if it is not abelian and it has no other ideals except $\{0\} \subset \mathfrak{g}$ and $\mathfrak{g} \subset \mathfrak{g}$.

Lemma III.57 (Bracket is surjective in simple Lie algebra)**.**
If $\mathfrak{g}$ is a simple Lie algebra, then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.

Proof. The vector subspace $\mathfrak{j} := [\mathfrak{g}, \mathfrak{g}]$ is clearly an ideal, since all Lie brackets take values in it. Since a simple Lie algebra is not abelian, this subspace is non-zero, $\mathfrak{j} \neq \{0\}$. As a non-zero ideal in a simple Lie algebra, it then has to be equal to the whole Lie algebra, $\mathfrak{j} = \mathfrak{g}$. □

Definition III.58 (Killing form)**.**
Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{k}$. The bilinear form $\mathsf{K} \colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{k}$ defined by

$$\mathsf{K}(X, Y) := \mathrm{tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y) \tag{III.24}$$

is called the **Killing form** on $\mathfrak{g}$.

Lemma III.59 (Symmetry and ad-invariance of the Killing form)**.**
The Killing form is symmetric,

$$\mathsf{K}(X, Y) = \mathsf{K}(Y, X) \qquad \textit{for all } X, Y \in \mathfrak{g}$$

and ad*-invariant*

$$\mathsf{K}([Z, X], Y) + \mathsf{K}(X, [Z, Y]) = 0 \qquad \textit{for all } X, Y, Z \in \mathfrak{g}.$$

Proof. Symmetry follows directly from cyclicity of trace,

$$\mathsf{K}(X,Y) = \mathrm{tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y) = \mathrm{tr}(\mathrm{ad}_Y \circ \mathrm{ad}_X) = \mathsf{K}(Y,X).$$

To prove ad-invariance, observe first that the Jacobi-identity implies $\mathrm{ad}_{[Z,X]} = \mathrm{ad}_Z \circ \mathrm{ad}_X - \mathrm{ad}_X \circ \mathrm{ad}_Z$ and similarly for $\mathrm{ad}_{[Z,Y]}$. With this, calculate

$$\begin{aligned}\mathsf{K}([Z,X],Y) + \mathsf{K}(X,[Z,Y]) &= \mathrm{tr}\big(\mathrm{ad}_{[Z,X]} \circ \mathrm{ad}_Y\big) + \mathrm{tr}\big(\mathrm{ad}_X \circ \mathrm{ad}_{[Z,Y]}\big)\\ &= \mathrm{tr}\big(\mathrm{ad}_Z \circ \mathrm{ad}_X \circ \mathrm{ad}_Y\big) - \mathrm{tr}\big(\mathrm{ad}_X \circ \mathrm{ad}_Z \circ \mathrm{ad}_Y\big)\\ &\quad + \mathrm{tr}\big(\mathrm{ad}_X \circ \mathrm{ad}_Z \circ \mathrm{ad}_Y\big) - \mathrm{tr}\big(\mathrm{ad}_X \circ \mathrm{ad}_Y \circ \mathrm{ad}_Z\big).\end{aligned}$$

The second and third term cancel immediately, and the first and the last term cancel by cyclicity of trace. This proves the ad-invariance of K. □

4.2. Definition and examples of Lie algebras of compact type

The assumption that we make about the Lie algebra is encapsulated in the following definition.

Definition III.60 (Lie algebra of compact type).
A finite-dimensional real Lie algebra $\mathfrak{g}$ is said to be a **Lie algebra of compact type** if there exists an ad-invariant inner product on the vector space $\mathfrak{g}$, i.e., a symmetric bilinear positive definite form

$$\mathsf{B}\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}, \qquad (X,Y) \mapsto \mathsf{B}(X,Y)$$

such that

$$\mathsf{B}([Z,X],Y) + \mathsf{B}(X,[Z,Y]) = 0 \qquad \text{for all } X,Y,Z \in \mathfrak{g}.$$

We have two familiar examples.

Example III.61 (The Lie algebra $\mathfrak{so}_n$ is of compact type).
Consider the Lie algebra

$$\mathfrak{so}_n = \Big\{ X \in \mathbb{R}^{n\times n} \;\Big|\; X^\top = -X \Big\}$$

of the special orthogonal group SO_n, see Examples II.8 and II.43.

Define the bilinear form on $\mathfrak{so}_n$ by

$$\mathsf{B}(X,Y) := -\mathrm{tr}(XY) \qquad \text{for } X,Y \in \mathfrak{so}_n.$$

By cyclicity of trace, we have

$$\mathsf{B}(X,Y) := -\mathrm{tr}(XY) = -\mathrm{tr}(YX) = \mathsf{B}(Y,X),$$

so the form B is symmetric. To see that it is also positive definite, calculate for $X \in \mathfrak{so}_n$, using the defining property $X^\top = -X$,

$$\begin{aligned}\mathsf{B}(X,X) = -\mathrm{tr}(XX) = +\mathrm{tr}(X^\top X) &= \sum_{i=1}^n (X^\top X)_{ii}\\ &= \sum_{i,j=1}^n (X^\top)_{ij}\, X_{ji}\\ &= \sum_{i,j=1}^n X_{ji}\, X_{ji} = \sum_{i,j=1}^n X_{ji}^2 \geq 0,\end{aligned}$$

and observe that equality only arises if all entries of X are zero. We have thus shown that B is a symmetric positive definite bilinear form, i.e., an inner product on $\mathfrak{so}_n$.

To see that this inner product B is ad-invariant, let $X, Y, Z \in \mathfrak{so}_n$ and calculate

$$\begin{aligned}\mathsf{B}(\mathrm{ad}_Z(X), Y) + \mathsf{B}(X, \mathrm{ad}_Z(Y)) &= -\operatorname{tr}([Z,X]Y) - \operatorname{tr}(X[Z,Y]) \\ &= -\operatorname{tr}\big(ZXY - XZY\big) - \operatorname{tr}\big(XZY - XYZ\big) \\ &= -\operatorname{tr}(ZXY) + \operatorname{tr}(XYZ) \\ &= 0,\end{aligned}$$

where the last equality holds again by cyclicity of trace.

Therefore B is an ad-invariant inner product on $\mathfrak{so}_n$, so $\mathfrak{so}_n$ is a Lie algebra of compact type.

Example III.62 (The Lie algebra $\mathfrak{u}_n$ is of compact type)**.**
Consider the Lie algebra

$$\mathfrak{u}_n = \Big\{ X \in \mathbb{C}^{n \times n} \;\Big|\; X^\dagger = -X \Big\}$$

of the unitary group U_n, see Example II.9 and Exercise II.44. Define the bilinear form on $\mathfrak{u}_n$ by

$$\mathsf{B}(X, Y) := -\operatorname{tr}(XY) \qquad \text{for } X, Y \in \mathfrak{u}_n.$$

Although the matrices $X, Y \in \mathfrak{u}_n$ are complex, this form is actually real valued: the defining property requires $X^\dagger = -X$, and $Y^\dagger = -Y$, and traces satisfy $\operatorname{tr}(XY) = \operatorname{tr}((XY)^\top) = \operatorname{tr}(Y^\top X^\top)$, so we get

$$\overline{\mathsf{B}(X,Y)} = -\overline{\operatorname{tr}(XY)} = -\overline{\operatorname{tr}(Y^\top X^\top)} = -\operatorname{tr}(Y^\dagger X^\dagger) = -\operatorname{tr}(YX).$$

It then follows from cyclicity of trace that this expression equals $\mathsf{B}(X, Y)$ again, so we see that $\mathsf{B}(X, Y) \in \mathbb{R}$.

The form B is symmetric (by cyclicity of trace), and ad-invariant (by a calculation similar to the previous example). For positive definiteness, using the defining property $X^\dagger = -X$ of $X \in \mathfrak{u}_n$, calculate

$$\begin{aligned}\mathsf{B}(X, X) = -\operatorname{tr}(XX) = +\operatorname{tr}(X^\dagger X) &= \sum_{i,j=1}^{n} (X^\dagger)_{ij}\, X_{ji} \\ &= \sum_{i,j=1}^{n} \overline{X_{ji}}\, X_{ji} = \sum_{i,j=1}^{n} |X_{ji}|^2 \geq 0,\end{aligned}$$

and observe that equality only arises if all entries of X are zero. Therefore B is an ad-invariant inner product on $\mathfrak{u}_n$, so $\mathfrak{u}_n$ is a Lie algebra of compact type.

Note also that for any $c \in \mathbb{R}$, the imaginary multiple of the unit matrix $\mathbb{i}c\mathbb{I}$ is satisfies $(\mathbb{i}c\mathbb{I})^\dagger = -\mathbb{i}c\mathbb{I}$, and therefore belongs to $\mathfrak{u}_n$. Multiples of unit matrix commute with any other matrix, so the Lie bracket of $\mathbb{i}c\mathbb{I}$ with any other element vanishes. This shows that the centre $\mathfrak{z} \subset \mathfrak{u}_n$ of $\mathfrak{u}_n$ is non-trivial, $\mathfrak{z} \neq \{0\}$. One can show that the centre is in fact one-dimensional, $\mathfrak{z} = \mathbb{i}\mathbb{R}\,\mathbb{I}$, and that correspondingly $\mathfrak{su}_n \subset \mathfrak{u}_n$ has trivial centre.

More generally the Lie algebra of any closed subgroup of U_n or SO_n is a Lie algebra of compact type.

Lemma III.63 (Ideals are subrepresentations of the adjoint representation)**.**
A subspace $\mathfrak{j} \subset \mathfrak{g}$ is an ideal if and only if it is a subrepresentation $\mathfrak{j} \subset \mathfrak{g}$ of the adjoint representation of $\mathfrak{g}$.

Proposition III.64 (Simple pieces of a Lie algebra of compact type)**.**
A Lie algebra of compact type is the direct sum of an abelian Lie algebra (its centre) and simple Lie algebras of compact type.

Proof. Since the space $\mathfrak{g}$ carries an invariant inner product B, it follows from Lemma III.7 that the orthogonal complement $\mathfrak{j}^\perp$ of a subrepresentation $\mathfrak{j} \subset \mathfrak{g}$ of the adjoint representation $\mathfrak{g}$ is a complementary subrepresentation, i.e., we have $\mathfrak{g} = \mathfrak{j} \oplus \mathfrak{j}^\perp$. By induction this implies that we have a direct sum decomposition

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n,$$

where $\mathfrak{g}_0$ is the direct sum of all trivial subrepresentations (i.e., the centre $\mathfrak{g}_0 = \mathfrak{z}$ of $\mathfrak{g}$), and $\mathfrak{g}_1, \ldots, \mathfrak{g}_n$ are irreducible subrepresentations of the adjoint representation, which are not trivial (in particular contain no non-zero central elements). Since $\mathfrak{g}_j$ is a subrepresentation of the adjoint representation for each j, we have $[\mathfrak{g}, \mathfrak{g}_j] \subset \mathfrak{g}_j$. In particular we get $[\mathfrak{g}_i, \mathfrak{g}_j] = \mathfrak{g}_j \cap \mathfrak{g}_i = \{0\}$ for $i \neq j$. It follows that if $\mathfrak{j}_j \subset \mathfrak{g}_j$ is an ideal (i.e., $\mathrm{ad}_{\mathfrak{g}_j}$-invariant subspace), then it is an $\mathrm{ad}_{\mathfrak{g}}$-invariant subspace. For $j > 0$, by irreducibility we must therefore have $\mathfrak{j}_j = 0$ or $\mathfrak{j}_j = \mathfrak{g}_j$. Since there are no non-zero central elements in $\mathfrak{g}_j$, and $\mathfrak{g}_j$ is irreducible, we must have $\mathfrak{j}_j = \mathfrak{g}_j$. This shows that $\mathfrak{g}_1, \ldots, \mathfrak{g}_n$ are simple Lie algebras. Lie subalgebras of a Lie algebra of compact type are themselves automatically of compact type — one can just restrict the invariant inner product to the subspace. □

From the proof above we get also the following corollaries.

Corollary III.65. *For a Lie algebra $\mathfrak{g}$ of compact type, we have $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{z}^\perp$.*

In view of the above corollary, the following case is particularly important.

Definition III.66 (Lie algebra of compact semisimple type)**.**
We say that a real Lie algebra $\mathfrak{g}$ is a **Lie algebra of compact semisimple type** if it is a Lie algebra of compact type and its centre $\mathfrak{z} \subset \mathfrak{g}$ is trivial, $\mathfrak{z} = \{0\}$.

Corollary III.67. *For a Lie algebra of compact semisimple type we have $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.*

A property that we will soon use, is that the existence of invariant inner product implies that the adjoint actions are given by skew-symmetric operators in an orthonormal basis.

Lemma III.68 (Skew-symmetry of adjoint actions for compact type)**.**
Suppose that $\mathfrak{g}$ is a Lie algebra of compact type. Let $Z_1, \ldots, Z_d$ be an orthonormal basis of $\mathfrak{g}$ with respect to the ad*-invariant inner product B. For any $Z \in \mathfrak{g}$, the matrix $A \in \mathbb{R}^{d \times d}$ of the linear map $\mathrm{ad}_Z \colon \mathfrak{g} \to \mathfrak{g}$ in this basis is antisymmetric, $A_{ij} = -A_{ji}$ for all $i, j = 1, \ldots, d$.*

Proof. The matrix elements A_{kj} of ad_X are by definition the coefficients in $\mathrm{ad}_X(Z_j) = \sum_k A_{kj} Z_k$. Taking inner products with Z_i and using orthonormality of the basis, we find the expression

$$A_{ij} = \mathsf{B}(Z_i, \mathrm{ad}_X(Z_j))$$

for them. By ad-invariance and symmetry of B, we then get

$$A_{ij} = \mathsf{B}(Z_i, \mathrm{ad}_X(Z_j)) = -\mathsf{B}(\mathrm{ad}_X(Z_i), Z_j) = -\mathsf{B}(Z_j, \mathrm{ad}_X(Z_i)) = -A_{ji},$$

which proves antisymmetry. □

This has implications about the Killing form.

Proposition III.69 (Killing form and semisimplicity).
A Lie algebra $\mathfrak{g}$ of compact type has trivial center $\mathfrak{z} = \{0\}$ if and only if its Killing form K is strictly negative definite.

Proof. Since $\mathfrak{g}$ is of compact type, there exists an invariant inner product, and the matrices of the adjoint actions of all elements are antisymmetric in an orthonormal basis $Z_1, \ldots, Z_d$ of $\mathfrak{g}$. Let $X \in \mathfrak{g}$, and denote the matrix of ad_X in this basis by $A = (A_{ij})_{i,j=1}^d$. Antisymmetry means $A^\top = -A$. To evaluate the Killing form, we can calculate traces in this basis. We find, e.g.,

$$\begin{aligned}\mathsf{K}(X,X) := \mathrm{tr}(\mathrm{ad}_X \circ \mathrm{ad}_X) = \mathrm{tr}(AA) = \mathrm{tr}\big((-A^\top)A\big)\\ = -\sum_{i,j=1}^d (A^\top)_{ij} A_{ji} = -\sum_{i,j=1}^d A_{ji}^2 \leq 0,\end{aligned}$$

and equality can only occur if $A \equiv 0$. But note that $A \equiv 0$ means $\mathrm{ad}_X = 0$, which occurs if and only if X is in the centre $X \in \mathfrak{z}$. Thus trivial centre $\mathfrak{z} = \{0\}$ implies that the Killing form is strictly negative definite, whereas non-trivial centre $\mathfrak{z} \neq \{0\}$ implies that the Killing form has no strict definiteness. □

Therefore in a Lie algebra of compact semisimple type, we can use $-\mathsf{K}$ as an invariant inner product, and it will be convenient to do so.

4.3. Complexification of Lie algebras of compact semisimple type

Let us now assume that $\mathfrak{g}$ is of compact semisimple type, i.e., has an invariant inner product and trivial centre. We saw above that in this case $\mathfrak{g}$ is a direct sum of simple Lie algebras, and we can use the negative of the Killing form as an invariant inner product.

The complexification of the real Lie algebra $\mathfrak{g}$ is the complex Lie algebra

$$\mathfrak{g}_\mathbb{C} := \mathfrak{g} \oplus \mathbb{i}\mathfrak{g}.$$

Let us define a real-linear (anti-)involution $Z \mapsto Z^*$ on the complexification $\mathfrak{g}_\mathbb{C}$ by

$$(X + \mathbb{i}Y)^* := -X + \mathbb{i}Y \qquad \text{for } X, Y \in \mathfrak{g}.$$

Note that we have $(\mathbb{i}Z)^* = -\mathbb{i}Z^*$ for any $Z \in \mathfrak{g}_\mathbb{C}$, so the map $Z \mapsto Z^*$ is in fact conjugate linear.[7] This map reverses brackets (the term anti-involution rather than involution is used for this reason), as we prove next.

Lemma III.70 (Anti-involution).
For any $Z, W \in \mathfrak{g}_\mathbb{C}$ we have

$$[Z, W]^* = -[Z^*, W^*].$$

Proof. Write $Z = X_1 + \mathbb{i}Y_1$ and $W = X_2 + \mathbb{i}Y_2$ for $X_1, Y_1, X_2, Y_2 \in \mathfrak{g}$. Calculate first

$$\begin{aligned}[Z^*, W^*] &= [-X_1 + \mathbb{i}Y_1,\ -X_2 + \mathbb{i}Y_2]\\ &= \big([X_1, X_2] - [Y_1, Y_2]\big) - \mathbb{i}\big([X_1, Y_2] + [Y_1, X_2]\big).\end{aligned}$$

[7] Indeed, for $Z = X + \mathbb{i}Y$, we have $(\mathbb{i}Z)^* = (-Y + \mathbb{i}X)^* = Y + \mathbb{i}X = \mathbb{i}(X - \mathbb{i}Y) = -\mathbb{i}Z^*$.

Then calculate for comparison

$$\begin{aligned}[Z,W]^* &= \big([X_1 + \mathbb{i}Y_1,\, X_2 + \mathbb{i}Y_2]\big)^* \\ &= \Big(\big([X_1,X_2] - [Y_1,Y_2]\big) + \mathbb{i}\big([X_1,Y_2] + [Y_1,X_2]\big)\Big)^* \\ &= -\big([X_1,X_2] - [Y_1,Y_2]\big) + \mathbb{i}\big([X_1,Y_2] + [Y_1,X_2]\big).\end{aligned}$$

The two expressions are the opposites of each other, which proves the claim. □

We also define the real Lie subalgebra $\mathfrak{g}_{\mathbb{R}} \subset \mathfrak{g}_{\mathbb{C}}$ by

$$\mathfrak{g}_{\mathbb{R}} := \big\{ Z \in \mathfrak{g}_{\mathbb{C}} \;\big|\; Z^* = Z \big\}. \tag{III.25}$$

Note that $\mathfrak{g}_{\mathbb{R}} = \mathbb{i}\mathfrak{g}$ — this real Lie subalgebra is <u>not</u> $\mathfrak{g}$!

Proposition III.71 (Killing form for compact semisimple type)**.**
Let $\mathfrak{g}$ be a Lie algebra of compact semisimple type and $\mathfrak{g}_{\mathbb{C}}$ its complexification. Then the Killing form K defined by $\mathsf{K}(Z,W) = \mathrm{tr}(\mathrm{ad}_Z \circ \mathrm{ad}_Y)$ has the following properties.

(i) K *is strictly positive definite on* $\mathfrak{g}_{\mathbb{R}}$.
(ii) K *is non-degenerate on* $\mathfrak{g}_{\mathbb{C}}$.

Proof. Recall from Lemma III.69 that K is strictly negative definite on $\mathfrak{g}$. If $Z \in \mathfrak{g}_{\mathbb{R}} = \mathbb{i}\,\mathfrak{g}$, then $Z = \mathbb{i}Y$ with $Y \in \mathfrak{g}$, and therefore

$$\mathsf{K}(Z,Z) = \mathsf{K}(\mathbb{i}Y, \mathbb{i}Y) = -\mathsf{K}(Y,Y) \geq 0,$$

and equality only arises if $Y = 0 \in \mathfrak{g}$, i.e., if $Z = 0 \in \mathfrak{g}_{\mathbb{R}}$. This proves part (i).

For non-degeneracy on the complexification $\mathfrak{g}_{\mathbb{C}}$, assume that $Z \in \mathfrak{g}_{\mathbb{C}}$. Write $Z = X + \mathbb{i}Y$ with $X, Y \in \mathfrak{g}$ and recall that $Z^* = -X + \mathbb{i}Y$. Calculate

$$\mathsf{K}(Z^*,Z) = \mathsf{K}(-X + \mathbb{i}Y, X + \mathbb{i}Y) = -\mathsf{K}(X,X) - \mathsf{K}(Y,Y) \geq 0, \tag{III.26}$$

where the inequality is again by negative definiteness of K on $\mathfrak{g}$. Equality only holds here if $X = Y = 0 \in \mathfrak{g}$ (strict negative-definiteness), i.e., if $Z = 0 \in \mathfrak{g}_{\mathbb{C}}$. This proves non-degeneracy, part (ii). □

On the complex vector space $\mathfrak{g}_{\mathbb{C}}$, define

$$\langle Z, W \rangle := \mathsf{K}(Z^*, W) \qquad \text{for } Z, W \in \mathfrak{g}_{\mathbb{C}}. \tag{III.27}$$

Lemma III.72 (Inner product on the complexification)**.**
If $\mathfrak{g}$ is a Lie algebra of compact semisimple type, then the formula (III.27) *defines an inner product on its complexification $\mathfrak{g}_{\mathbb{C}}$.*

Proof. Clearly (III.27) is $\mathbb{C}$-linear in W, and conjugate linear in Z (since $Z \mapsto Z^*$ is). Positive definiteness follows from the calculation (III.26), which showed that $\langle Z, Z\rangle \geq 0$ with equality only if $Z = 0$. To prove the conjugate symmetry property, write first $Z = X + \mathbb{i}Y$ and $W = X' + \mathbb{i}Y'$ with $X, Y, X', Y' \in \mathfrak{g}$ and calculate

$$\begin{aligned}\langle Z, W\rangle = \mathsf{K}(Z^*, W) &= \mathsf{K}(-X + \mathbb{i}Y, X' + \mathbb{i}Y') \\ &= -\mathsf{K}(X,X') - \mathsf{K}(Y,Y') - \mathbb{i}\,\mathsf{K}(X,Y') + \mathbb{i}\,\mathsf{K}(Y,X')\end{aligned}$$

The Killing form is real on $\mathfrak{g}$, so the complex conjugate of the above is

$$\begin{aligned}\overline{\langle Z, W\rangle} &= -\mathsf{K}(X,X') - \mathsf{K}(Y,Y') + \mathbb{i}\,\mathsf{K}(X,Y') - \mathbb{i}\,\mathsf{K}(Y,X') \\ &= \mathsf{K}(X + \mathbb{i}Y, -X' + \mathbb{i}Y') \\ &= \mathsf{K}(Z, W^*) \\ &= \langle W, Z\rangle,\end{aligned}$$

where we used symmetricity of the Killing form in the last step. This finishes the proof. □

We next show that with respect to this complex Hilbert space structure on $\mathfrak{g}_{\mathbb{C}}$, the Hilbert space adjoints of the operators $\mathrm{ad}_Z : \mathfrak{g}_{\mathbb{C}} \to \mathfrak{g}_{\mathbb{C}}$ are obtained by applying the anti-involution $Z \mapsto Z^*$. This explains our a priori strange looking definition of the anti-involution.

Proposition III.73 (Hilbert space adjoints of adjoint actions)**.**
For any $Z \in \mathfrak{g}_{\mathbb{C}}$ and $W_1, W_2 \in \mathfrak{g}_{\mathbb{C}}$, we have

$$\langle \mathrm{ad}_Z(W_1), W_2 \rangle = \langle W_1, \mathrm{ad}_{Z^*}(W_2) \rangle,$$

i.e., the Hilbert space adjoint of ad_Z *is* ad_{Z^*}*. In particular for all $Z \in \mathfrak{g}_{\mathbb{R}}$, the operators* ad_Z *are self-adjoint.*

Proof. Recall the anti-involution property of Lemma III.70, and the ad-invariance of the Killing form from Lemma III.59. Then the first assertion is proven by the calculation

$$\begin{aligned}\langle \mathrm{ad}_Z(W_1), W_2 \rangle &= \mathsf{K}([Z, W_1]^*, W_2) \\ &= -\mathsf{K}([Z^*, W_1^*], W_2) \\ &= +\mathsf{K}(W_1^*, \mathrm{ad}_{Z^*}(W_2)) = \langle W_1, \mathrm{ad}_{Z^*}(W_2) \rangle.\end{aligned}$$

Self-adjointness of ad_Z for $Z \in \mathfrak{g}_{\mathbb{R}}$ then follows directly from the defining property $Z^* = Z$ in this real subalgebra. □

4.4. Structure of Lie algebras of compact semisimple type

Assume still that $\mathfrak{g}$ is of compact semisimple type. Let us choose a maximal abelian subalgebra $\mathfrak{h} \subset \mathfrak{g}$, called a **Cartan subalgebra**. Then its complexification $\mathfrak{h}_{\mathbb{C}}$ is an abelian Lie subalgebra of the complex Lie algebra $\mathfrak{g}_{\mathbb{C}}$, and $\mathfrak{h}_{\mathbb{R}} = \mathbb{i}\mathfrak{h}$ is an abelian Lie subalgebra of $\mathfrak{g}_{\mathbb{R}} = \mathbb{i}\mathfrak{g}$.

By virtue of Proposition III.73, the collection

$$\Big\{ \mathrm{ad}_H : \mathfrak{g}_{\mathbb{C}} \to \mathfrak{g}_{\mathbb{C}} \;\Big|\; H \in \mathfrak{h}_{\mathbb{R}} \Big\}$$

consists of commuting self-adjoint operators. Therefore these operators can be simultaneously diagonalized, i.e., the space $\mathfrak{g}_{\mathbb{C}}$ can be written as a direct sum of joint eigenspaces of the form

$$\mathfrak{g}_\alpha := \Big\{ Z \in \mathfrak{g}_{\mathbb{C}} \;\Big|\; \mathrm{ad}_H(Z) = \alpha(H) Z \quad \forall H \in \mathfrak{h}_{\mathbb{R}}, \Big\} \tag{III.28}$$

where $\alpha(H) \in \mathbb{R}$ are the real eigenvalues of the self-adjoint operators ad_H. Obviously $\alpha(H)$ depends linearly on $H \in \mathfrak{h}_{\mathbb{R}}$, so defines a functional $\alpha \in \mathfrak{h}_{\mathbb{R}}^*$.

The eigenvalues on $\mathfrak{h}_{\mathbb{C}}$ are zeroes by abelianity, i.e, the corresponding functional is $0 \in \mathfrak{h}_{\mathbb{R}}^*$ and we have $\mathfrak{h}_{\mathbb{C}} \subset \mathfrak{g}_0$. Since the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ was chosen to be a maximal abelian Lie subalgebra, there are no other elements which have zero as the joint eigenvalue, i.e., we have $\mathfrak{h}_{\mathbb{C}} = \mathfrak{g}_0$. The rest of the joint eigenvalues will be of great interest to us.

Definition III.74 (Roots)**.**
Those non-zero $\alpha \in \mathfrak{h}_{\mathbb{R}}^*$ for which the joint eigenspace (III.28) is non-trivial, $\mathfrak{g}_\alpha \neq \{0\}$, are called **roots** of $\mathfrak{g}$. The finite set of roots of $\mathfrak{g}$ is denoted by $\Phi \subset \mathfrak{h}_{\mathbb{R}}^*$. The corresponding vector subspaces $\mathfrak{g}_\alpha \subset \mathfrak{g}_{\mathbb{C}}$ are called **root spaces**.

The resulting decomposition (direct sum of vectors spaces)

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha \tag{III.29}$$

is called the **root space decomposition**.

Lemma III.75 (Opposite root)**.**
If α is a root, then also $-\alpha$ is a root, i.e., we have $-\Phi = \Phi \subset \mathfrak{h}_{\mathbb{R}}^*$.

Proof. If $\alpha \in \Phi$, then we can choose an elements $X_\alpha \in \mathfrak{g}_\alpha$ such that $\langle X_\alpha, X_\alpha \rangle = 1$. Consider the non-zero element $X_\alpha^* \in \mathfrak{g}_{\mathbb{C}}$. For any $H \in \mathfrak{h}_{\mathbb{R}}$ we have, using $H^* = H$ and Lemma III.70,

$$\mathrm{ad}_H(X_\alpha^*) = [H, X_\alpha^*] = [H^*, X_\alpha^*] = -[H, X_\alpha]^* = -\alpha(H) X_\alpha^*.$$

This shows that $X_\alpha^* \in \mathfrak{g}_{-\alpha}$. We conclude that $\mathfrak{g}_{-\alpha} \neq \{0\}$ and $-\alpha \in \Phi$. □

Non-degeneracy of the inner product on $\mathfrak{h}_{\mathbb{R}}$ permits us to make the following definition.

Definition III.76 (Root vectors)**.**
For any root $\alpha \in \Phi$, there exists a unique $H_\alpha \in \mathfrak{h}_{\mathbb{R}}$ such that

$$\alpha(H) = \langle H_\alpha, H \rangle \qquad \text{for all } H \in \mathfrak{h}_{\mathbb{R}}.$$

We call H_α the **root vector** corresponding to the root α.

Lemma III.77 (Root vectors span Cartan subalgebra)**.**
The roots and root vectors span $\mathfrak{h}_{\mathbb{R}}^*$ and $\mathfrak{h}_{\mathbb{R}}$, respectively, i.e.,

$$\mathfrak{h}_{\mathbb{R}} = \mathrm{span}_{\mathbb{R}} \left\{ H_\alpha \;\middle|\; \alpha \in \Phi \right\}$$
$$\mathfrak{h}_{\mathbb{R}}^* = \mathrm{span}_{\mathbb{R}} \left\{ \alpha \;\middle|\; \alpha \in \Phi \right\}.$$

Proof. Let us start by proving the second formula. Let $H \in \mathfrak{h}_{\mathbb{R}}$, and observe that since ad_H acts as multiplication by scalar $\alpha(H) \in \mathbb{R}$ on the root space $\mathfrak{g}_\alpha$, calculating traces using the root space decomposition (III.29) is easy, and we find

$$\mathsf{K}(H, H) = \mathrm{tr}(\mathrm{ad}_H \circ \mathrm{ad}_H) = 0 + \sum_{\alpha \in \Phi} \dim(\mathfrak{g}_\alpha)\, \alpha(H)^2.$$

The Killing form K is positive definite on $\mathfrak{h}_{\mathbb{R}} \subset \mathfrak{g}_{\mathbb{R}}$ by Lemma III.71, i.e., for any non-zero $H \neq 0$ we must have $\mathsf{K}(H, H) > 0$. Therefore the above calculation implies that there must exist some $\alpha \in \Phi$ for which $\alpha(H) \neq 0$. It follows that the span of all $\alpha \in \Phi$ is the dual $\mathfrak{h}_{\mathbb{R}}^*$ entirely. Since the root vectors $H_\alpha \in \mathfrak{h}_{\mathbb{R}}$ were obtained from the roots $\alpha \in \mathfrak{h}_{\mathbb{R}}^*$ via the identification based on the non-degenerate inner product, the root vectors also must span $\mathfrak{h}_{\mathbb{R}}$. This finishes the proof. □

Lemma III.78 (Brackets of root spaces)**.**
For any $\alpha, \beta \in \mathfrak{h}_{\mathbb{R}}^*$, we have

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}.$$

In particular we have $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{h}_{\mathbb{C}}$, and if $\alpha + \beta \notin \Phi \cup \{0\}$, then $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \{0\}$.

Proof. Assume that $X_\alpha \in \mathfrak{g}_\alpha$ and $X_\beta \in \mathfrak{g}_\beta$. For $H \in \mathfrak{h}_\mathbb{R}$, calculate using Jacobi identity

$$\begin{aligned}\mathrm{ad}_H([X_\alpha, X_\beta]) &= \big[H, [X_\alpha, X_\beta]\big] \\ &= -\big[X_\alpha, [X_\beta, H]\big] - \big[X_\beta, [H, X_\alpha]\big] \\ &= +\beta(H)\,[X_\alpha, X_\beta] - \alpha(H)\,[X_\beta, X_\alpha] \\ &= \big(\alpha(H) + \beta(H)\big)\,[X_\alpha, X_\beta].\end{aligned}$$

In view of $\mathfrak{g}_0 = \mathfrak{h}_\mathbb{C}$ and $\mathfrak{g}_\gamma = \{0\}$ for $\gamma \notin \Phi \cup \{0\}$, the rest of the assertions are obvious consequences. given □

Lemma III.79 (Brackets from opposite root spaces)**.**
Let $\alpha \in \Phi$ be a root, and let H_α be the corresponding root vector (see Definition III.76). Choose $X_\alpha \in \mathfrak{g}_\alpha$ such that $\langle X_\alpha, X_\alpha\rangle = 1$. Then we have

$$[X_\alpha, X_\alpha^*] = H_\alpha.$$

Proof. In the proof of Lemma III.75 we saw that $X_\alpha \in \mathfrak{g}_\alpha$ implies $X_\alpha^* \in \mathfrak{g}_{-\alpha}$. By Lemma III.78, we thus have $[X_\alpha, X_\alpha^*] \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{g}_0 = \mathfrak{h}_\mathbb{C}$. Moreover by the anti-involution property we have

$$[X_\alpha, X_\alpha^*]^* = -[X_\alpha^*, X_\alpha^{**}] = -[X_\alpha^*, X_\alpha] = [X_\alpha, X_\alpha^*],$$

so this bracket in fact lies in the real Lie subalgebra, $[X_\alpha, X_\alpha^*] \in \mathfrak{h}_\mathbb{R}$.

Now for $H \in \mathfrak{h}_\mathbb{R}$ calculate the following inner product using ad-invariance of the Killing form

$$\begin{aligned}\big\langle [X_\alpha, X_\alpha^*],\, H\big\rangle = \mathsf{K}([X_\alpha, X_\alpha^*]^*, H) &= \mathsf{K}([X_\alpha, X_\alpha^*], H) \\ &= -\mathsf{K}(X_\alpha^*, [X_\alpha, H]) \\ &= \alpha(H)\,\mathsf{K}(X_\alpha^*, X_\alpha) \\ &= \alpha(H)\,\langle X_\alpha, X_\alpha\rangle \\ &= \alpha(H).\end{aligned}$$

This is the defining property of the root vector H_α, se we indeed have $[X_\alpha, X_\alpha^*] = H_\alpha$. □

Lemma III.80 (One-dimensionality of root spaces)**.**
For any root $\alpha \in \Phi$ we have

$$\dim(\mathfrak{g}_\alpha) = 1.$$

Moreover, the only integers $j \in \mathbb{Z}$ for which $j\alpha \in \Phi$ is also a root are $j = \pm 1$.

Proof. Choose again $X_\alpha \in \mathfrak{g}_\alpha$ such that $\langle X_\alpha, X_\alpha\rangle = 1$. Denote $X_{-\alpha} := X_\alpha^*$ and recall that then $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$. Consider the vector subspace of $\mathfrak{g}_\mathbb{C}$ defined by

$$\mathbb{C}X_{-\alpha} \oplus \mathfrak{h}_\mathbb{C} \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{2\alpha} \oplus \cdots \oplus \mathfrak{g}_{k\alpha}, \tag{III.30}$$

where $k \in \mathbb{N}$ is the largest integer for which $k\alpha$ is a root. From Lemma III.78 it follows that ad_{X_α} preserves the subspace (III.30). To see that also $\mathrm{ad}_{X_{-\alpha}}$ preserves the subspace (III.30), we make three observations:

$$\begin{aligned}\mathrm{ad}_{X_{-\alpha}}(X_{-\alpha}) &= [X_{-\alpha}, X_{-\alpha}] = 0, \\ \mathrm{ad}_{X_{-\alpha}}(H) &= [X_{-\alpha}, H] = +\alpha(H)X_{-\alpha} \qquad \text{for all } H \in \mathfrak{h}_\mathbb{C},\end{aligned}$$

and finally $\mathrm{ad}_{X_{-\alpha}}$ maps the space $\mathfrak{g}_\alpha \oplus \cdots \oplus \mathfrak{g}_{k\alpha}$ into $\mathfrak{h}_\mathbb{C} \oplus \mathfrak{g}_\alpha \oplus \cdots \oplus \mathfrak{g}_{(k-1)\alpha}$ by Lemma III.78 again.

Therefore both ad_{X_α} and $\mathrm{ad}_{X_{-\alpha}}$ can be restricted to linear maps of the subspace (III.30) to itself, and then so can their commutator $[\mathrm{ad}_{X_\alpha}, \mathrm{ad}_{X_{-\alpha}}] = \mathrm{ad}_{[X_\alpha, X_{-\alpha}]} = \mathrm{ad}_{H_\alpha}$, where the last equality follows from Lemma III.79. Calculate the trace of ad_{H_α} on the subspace (III.30), which must vanish (the trace of a commutator is zero by cyclicity of trace):

$$\begin{aligned} 0 = & -\alpha(H_\alpha) + 0 + \alpha(H_\alpha)\dim(\mathfrak{g}_\alpha) + 2\alpha(H_\alpha)\dim(\mathfrak{g}_{2\alpha}) + \cdots + k\alpha(H_\alpha)\dim(\mathfrak{g}_{k\alpha}) \\ = & \, \alpha(H_\alpha)\Big(-1 + \dim(\mathfrak{g}_\alpha) + 2\dim(\mathfrak{g}_{2\alpha}) + \cdots + k\dim(\mathfrak{g}_{k\alpha})\Big). \end{aligned}$$

Since $\alpha(H_\alpha) = \langle H_\alpha, H_\alpha\rangle > 0$, the other factor must vanish, which is only possible if $\dim(\mathfrak{g}_\alpha) = 1$ and $\dim(\mathfrak{g}_{j\alpha}) = 0$ for all $j > 1$. Replacing α by $-\alpha$ similarly shows that $\dim(\mathfrak{g}_{-\alpha}) = 1$ and $\dim(\mathfrak{g}_{-j\alpha}) = 0$ for all $j > 1$, completing the proof. □

By calculating traces as in the proof of Lemma III.77 using the root space decomposition, one obtains in particular the following expression for the Killing form on the Cartan subalgebra.

Corollary III.81 (Killing form on Cartan subalgebra)**.**
For any $H_1, H_2 \in \mathfrak{h}_\mathbb{R}$ we have

$$\mathsf{K}(H_1, H_2) = \sum_{\alpha\in\Phi} \alpha(H_1)\alpha(H_2).$$

Convenient normalizations and $\mathfrak{sl}_2(\mathbb{C})$ subalgebras

Let us finally change to a very convenient normalization. For a root $\alpha \in \Phi$, choose again $X_\alpha \in \mathfrak{g}_\alpha$ such that $\langle X_\alpha, X_\alpha\rangle = 1$, and denote $X_{-\alpha} := X_\alpha^* \in \mathfrak{g}_{-\alpha}$. Let also $H_\alpha \in \mathfrak{h}_\mathbb{R}$ be the corresponding root vector and denote $\|H_\alpha\| = \sqrt{\langle H_\alpha, H_\alpha\rangle}$.
Define

$$H'_\alpha := \frac{2}{\|H_\alpha\|^2}\, H_\alpha, \qquad\qquad X'_{\pm\alpha} := \frac{\sqrt{2}}{\|H_\alpha\|}\, X_{\pm\alpha}. \tag{III.31}$$

Note that the normalization of H'_α implies in particular

$$\alpha(H'_\alpha) = \langle H_\alpha, H'_\alpha\rangle = \frac{2}{\|H_\alpha\|^2}\,\langle H_\alpha, H_\alpha\rangle = 2.$$

Lemma III.82 (The $\mathfrak{sl}_2(\mathbb{C})$ subalgebras associated to roots)**.**
The elements (III.31) *satisfy*

$$\begin{aligned} [H'_\alpha,\, X'_{\pm\alpha}] &= \pm 2\, X'_{\pm\alpha} \\ [X'_{+\alpha}, X'_{-\alpha}] &= H'_\alpha. \end{aligned}$$

In particular the span of $H'_\alpha, X'_{+\alpha}, X'_{-\alpha}$ in $\mathfrak{g}_\mathbb{C}$ is a Lie subalgebra isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.

Proof. Recall that $\alpha(H_\alpha) = \langle H_\alpha, H_\alpha\rangle = \|H_\alpha\|^2$. Noting that we have $X'_{\pm\alpha} \in \mathfrak{g}_{\pm\alpha}$, calculate

$$[H'_\alpha,\, X'_{\pm\alpha}] = \frac{2}{\|H_\alpha\|^2}\,[H_\alpha,\, X'_{\pm\alpha}] = \frac{\pm 2\,\alpha(H_\alpha)}{\|H_\alpha\|^2}\, X'_{\pm\alpha} = \pm 2\, X'_{\pm\alpha},$$

which proves the first formula. The other formula is straightforward in view of Lemma III.79,

$$[X'_{-\alpha},\, X'_{-\alpha}] = \Big(\frac{\sqrt{2}}{\|H_\alpha\|}\Big)^2\,[X_{-\alpha},\, X_{-\alpha}] = \frac{2}{\|H_\alpha\|^2}\, H_\alpha = H'_\alpha.$$

The Lie brackets of $H'_\alpha, X'_{+\alpha}, X'_{-\alpha}$ thus have exactly the same form as those of $H, E, F \in \mathfrak{sl}_2(\mathbb{C})$ given in (III.3), and the proof is complete. □

Proposition III.83 (Cartan integers).
For any two roots $\alpha, \beta \in \Phi$, we have

$$\beta(H'_\alpha) \in \{0, \pm 1, \pm 2, \pm 3, \pm 4\},$$

and in case $\beta(H'_\alpha) = \pm 4$ we have $\beta = \pm\alpha$.

Proof. Let us first prove that $\beta(H'_\alpha) \in \mathbb{Z}$. For this, consider the Lie subalgebra spanned by $H'_\alpha, X'_{+\alpha}, X'_{-\alpha}$, which is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. The adjoint representation $\mathfrak{g}_\mathbb{C}$ can in particular be seen as a representation of this Lie subalgebra. It is finite-dimensional, so we know that the eigenvalues of the H-generator in $\mathfrak{sl}_2(\mathbb{C})$ are integers by Proposition III.28. For $Z \in \mathfrak{g}_\beta$ we have $\mathrm{ad}_{H'_\alpha}(Z) = \beta(H'_\alpha)Z$, so $\beta(H'_\alpha)$ is such an eigenvalue in the adjoint representation $\mathfrak{g}_\mathbb{C}$. This proves integrality $\beta(H'_\alpha) \in \mathbb{Z}$.

Let us now point out that

$$\beta(H'_\alpha) = \frac{2\langle H_\beta, H_\alpha\rangle}{\|H_\alpha\|^2} \qquad \text{and similarly} \qquad \alpha(H'_\beta) = \frac{2\langle H_\alpha, H_\beta\rangle}{\|H_\beta\|^2}.$$

From these expressions it is clear that either both are simultaneously non-zero, $\beta(H'_\alpha) \neq 0 \neq \alpha(H'_\beta)$, or both vanish simultaneously $\beta(H'_\alpha) = 0 = \alpha(H'_\beta)$. The case of $\beta(H'_\alpha) = 0$ is allowed as a possibility in the statement, so let us focus on the case when both are non-zero.

From the above expressions we can form the product, which satisfies

$$\beta(H'_\alpha)\,\alpha(H'_\beta) = 4\,\frac{\langle H_\alpha, H_\beta\rangle^2}{\|H_\alpha\|^2\,\|H_\beta\|^2} \leq 4$$

by Cauchy-Schwarz inequality, and equality only occurs if $H_\beta = cH_\alpha$ for some $c \in \mathbb{R}$. Combined with integrality of the two factors, in the case that both are non-vanishing we already find that $\beta(H'_\alpha), \alpha(H'_\beta) \in \{\pm 1, \pm 2, \pm 3, \pm 4\}$.

Let us consider the case of equality in Cauchy-Schwarz above, $\beta(H'_\alpha)\,\alpha(H'_\beta) = 4$. Then we must have proportionality $H_\beta = cH_\alpha$, and correspondingly $\beta = c\alpha$ and $H'_\beta = \frac{1}{c}H'_\alpha$. The two integer factors $\beta(H'_\alpha), \alpha(H'_\beta)$ must either both be ± 2, or one of them has to be ± 4 and the other ± 1. Consider the latter possibility and assume without loss of generality $\beta(H'_\alpha) = \pm 4$ and $\alpha(H'_\beta) = \pm 1$. But then we find $\pm 4 = \beta(H'_\alpha) = c\alpha(H'_\alpha) = 2c$, which implies $c = \pm 2$, and we get $\beta = \pm 2\alpha$, which contradicts Lemma III.80. Therefore the only remaining possibility in this case is $\beta(H'_\alpha) = \pm 2 = \alpha(H'_\beta)$. In this case we find $\pm 2 = \beta(H'_\alpha) = c\alpha(H'_\alpha) = 2c$, which implies $c = \pm 1$, and we get $\beta = \pm\alpha$, which is a possibility allowed in the statement. □

Lemma III.84 (Proportional roots are equal or opposite).
If $\alpha \in \Phi$ is a root, then the only real multiples of it which are also roots are $\pm\alpha$.

Proof. The proof is basically the same as above: if $\beta = c\alpha$ is also a root, then both $2c = \beta(H'_\alpha) \in \mathbb{Z}$ and $\frac{2}{c} = \alpha(H'_\beta) \in \mathbb{Z}$ are integers. This already implies $c \in \left\{\pm\frac{1}{2}, \pm 1, \pm 2\right\}$. The case $c = \pm 2$ was ruled out by invoking Lemma III.80 in the previous proof, and the case $c = \pm\frac{1}{2}$ is ruled out by interchanging the roles of α and β. This only leaves the possibility $c = \pm 1$. □

4.5. Weyl group

Assume still that $\mathfrak{g}$ is of compact semisimple type.

The inner product $\langle\cdot,\cdot\rangle$ on $\mathfrak{h}_\mathbb{R}$ allowed us to identify the dual $\mathfrak{h}^*_\mathbb{R}$ with $\mathfrak{h}_\mathbb{R}$, through writing $\mu \in \mathfrak{h}^*_\mathbb{R}$ as $\mu(\cdot) = \langle H, \cdot\rangle$ with a certain unique $H \in \mathfrak{h}_\mathbb{R}$. We used this in particular to associate to a root $\alpha \in \Phi \subset \mathfrak{h}^*_\mathbb{R}$ the root vector $H_\alpha \in \mathfrak{h}_\mathbb{R}$. One can translate the inner product from $\mathfrak{h}_\mathbb{R}$ to $\mathfrak{h}^*_\mathbb{R}$ by setting $\langle\mu_1, \mu_2\rangle := \langle H_1, H_2\rangle$ for any $\mu_1, \mu_2 \in \mathfrak{h}^*_\mathbb{R}$ and the corresponding $H_1, H_2 \in \mathfrak{h}_\mathbb{R}$.

For $\alpha \in \Phi$ a root of a Lie algebra $\mathfrak{g}$, define the linear map

$$\varsigma_\alpha \colon \mathfrak{h}_{\mathbb{R}}^* \to \mathfrak{h}_{\mathbb{R}}^* \qquad \text{by} \qquad \varsigma_\alpha(\mu) = \mu - 2\,\frac{\langle \mu, \alpha\rangle}{\|\alpha\|^2}\,\alpha. \tag{III.32}$$

Lemma III.85 (Reflection in a hyperplane).
The map ς_α reflects the line $\mathbb{R}\alpha \subset \mathfrak{h}_{\mathbb{R}}^*$ across the origin, and restricts to the identity transformation on the hyperplane $\{\mu \in \mathfrak{h}_{\mathbb{R}}^* \mid \langle \mu, \alpha\rangle = 0\}$ orthogonal to this line.

Proof. From definition (III.32) we get in particular

$$\varsigma_\alpha(\alpha) = \alpha - 2\,\frac{\langle \alpha, \alpha\rangle}{\|\alpha\|^2}\,\alpha = \alpha - 2\alpha = -\alpha,$$

proving the first part. To get the second, note that if $\langle \mu, \alpha\rangle = 0$, then we obviously have $\varsigma_\alpha(\mu) = \mu - 0\alpha = \mu$. □

Proposition III.86 (Invariance of the set of roots under reflections).
If $\alpha, \beta \in \Phi$ are roots, then also $\varsigma_\alpha(\beta)$ is a root.

Proof. Consider the vector subspace

$$\bigoplus_{k\in\mathbb{Z}} \mathfrak{g}_{\beta+k\alpha} \subset \mathfrak{g}_{\mathbb{C}}.$$

It follows from Lemma III.78 that this space is invariant for the Lie subalgebra spanned by $H'_\alpha, X'_{+\alpha}, X'_{-\alpha}$, and thus forms a representation of this Lie subalgebra. By Lemma III.82, this Lie subalgebra is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$, so from Proposition III.28 it follows that the eigenvalues of $\mathrm{ad}_{H'_\alpha}$ are integers, and symmetric under reflection across zero. The eigenvalue of $\mathrm{ad}_{H'_\alpha}$ on the root space $\mathfrak{g}_\beta \neq \{0\}$ is $\beta(H'_\alpha)$, so there has to exist some $k \in \mathbb{Z}$ so that $\mathfrak{g}_{\beta+k\alpha} \neq \{0\}$, and on this space $\mathrm{ad}_{H'_\alpha}$ has eigenvalue $-\beta(H'_\alpha)$. By definition, however, the eigenvalue of $\mathrm{ad}_{H'_\alpha}$ on $\mathfrak{g}_{\beta+k\alpha}$ equals

$$(\beta + k\alpha)(H'_\alpha) = \beta(H'_\alpha) + k\alpha(H'_\alpha) = \beta(H'_\alpha) + 2k.$$

Equating the two expressions $-\beta(H'_\alpha) = \beta(H'_\alpha) + 2k$ for the eigenvalue and solving, one gets $k = -\beta(H'_\alpha)$. Since we have shown that $\mathfrak{g}_{\beta+k\alpha} \neq \{0\}$, we have that $\beta + k\alpha \in \Phi$ is a root. But with what we observed above, we can rewrite this root as

$$\beta + k\alpha = \beta - \beta(H'_\alpha)\alpha = \beta - \frac{2\langle H_\beta, H_\alpha\rangle}{\|H_\alpha\|^2}\alpha = \beta - 2\,\frac{\langle \beta, \alpha\rangle}{\|\alpha\|^2}\alpha = \varsigma_\alpha(\beta).$$

This proves the assertion. □

4.6. Classification of Lie algebras of compact semisimple type

We saw in Proposition III.83 that for any two roots $\alpha, \beta \in \Phi$, the numbers

$$n_{\beta\alpha} := \frac{2\langle H_\beta, H_\alpha\rangle}{\|H_\alpha\|^2} = \frac{2\langle \beta, \alpha\rangle}{\|\alpha\|^2} \tag{III.33}$$

must be integers from a small list of possibilities. Combined with the invariance of the set of roots under all of the reflections $\{\varsigma_\alpha \mid \alpha \in \Phi\}$ (and in fact the finite group generated by them), which we saw in Proposition III.86, this is the starting point of a full combinatorial classification of all Lie algebras of compact semisimple type (and in fact of all semisimple Lie algebras) based on only Euclidian geometry of the finite-dimensional inner product space $\mathfrak{h}_{\mathbb{R}}^*$. In fact we know already know from Lemma III.64 that Lie algebras of compact semisimple type are direct sums of simple Lie algebras of compact type, so it suffices to classify these.

The result of the classification — done in detail in, e.g., [**FH91**] — contains a list of four infinite series of simple Lie algebras, of which we have seen many (but not all) already:

type	simple Lie algebra of compact type	complex simple Lie algebra
A_r $(r \geq 1)$	$\mathfrak{su}_{r+1}$	$\mathfrak{sl}_{r+1}(\mathbb{C})$
B_r $(r \geq 2)$	$\mathfrak{so}_{2r+1}$	$\mathfrak{so}_{2r+1}(\mathbb{C})$
C_r $(r \geq 3)$	$\mathfrak{usp}_r$	$\mathfrak{sp}_{2r}(\mathbb{C})$
D_r $(r \geq 4)$	$\mathfrak{so}_{2r}$	$\mathfrak{so}_{2r}(\mathbb{C})$

Here $r \in \mathbb{N}$ signifies the rank of the Lie algebra $\mathfrak{g}$ in question, defined as the dimension $r = \dim(\mathfrak{h})$ of its Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$. The restrictions on $r \in \mathbb{N}$ indicated are there only to avoid listing isomorphic Lie algebras twice. In addition to the four infinite series of simple Lie algebras above, there are five exceptional types, of which we provide the labels, ranks, and dimensions in the following

type	simple Lie algebra	rank: $\dim(\mathfrak{h})$	dimension: $\dim(\mathfrak{g})$
G_2	$\mathfrak{g}_2$	2	14
F_4	$\mathfrak{f}_4$	4	52
E_6	$\mathfrak{e}_6$	6	78
E_7	$\mathfrak{e}_7$	7	133
E_8	$\mathfrak{e}_8$	8	248

For each of the above simple Lie algebras (more generally semisimple Lie algebras), finite-dimensional representations are completely reducible, i.e., direct sums of irreducibles. Moreover, the finite-dimensional irreducible representations are highest weight representations denoted by $L(\lambda)$, and the possible highest weights $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$ are determined by non-negativity and integrality conditions — in very close parallel to our analysis of representations of $\mathfrak{sl}_3(\mathbb{C})$ (i.e., the type A_2).

Appendix A

Background

During most parts of this course, vector spaces are over the field $\mathbb{C}$ of complex numbers. Often any other algebraically closed field of characteristic zero could be used instead. Sometimes, however, none of these assumptions are needed, and $\Bbbk$ could be any field. This ground field should always be clear from the context, and we usually omit explicitly mentioning it.

Definition A.1 (Space of linear maps between vector spaces)**.**
Let V, W be $\Bbbk$-vector spaces. The **space of linear maps** (i.e. the space of homomorphisms of $\Bbbk$-vector spaces) from V to W is denoted by

$$\mathrm{Hom}(V, W) \;=\; \big\{ T : V \to W \;\big|\; T \text{ is a } \Bbbk\text{-linear map} \big\} .$$

By pointwise addition and scalar multiplication, $\mathrm{Hom}(V, W)$ itself becomes a $\Bbbk$-vector space.

Definition A.2 (Dual of a vector space)**.**
The (algebraic) **dual** of a vector space V is the space of linear maps from V to the ground field,

$$V^* \;=\; \mathrm{Hom}(V, \Bbbk).$$

We denote the duality pairing by brackets $\langle \cdot, \cdot \rangle$. A dual vector $\varphi \in V^*$ is thus a linear function $V \to \Bbbk$, whose value at a vector $v \in V$ is (usually) denoted by $\langle \varphi, v \rangle$.

Definition A.3 (Transpose)**.**
For $T : V \to W$ a linear map, the transpose is the linear map $T^* : W^* \to V^*$ defined by

$$\langle T^*(\varphi), v \rangle \;=\; \langle \varphi, T(v) \rangle \qquad \text{for all } \varphi \in W^*,\ v \in V.$$

1. On tensor products of vector spaces

A crucial concept in this course is that of a tensor product of vector spaces. Here, vector spaces can be over any field $\Bbbk$, but it should be noted that the concept of tensor product depends of the field. In this course we primarily use tensor products of complex vector spaces.

Definition A.4 (Bilinear maps)**.**
Let V_1, V_2, W be vector spaces. A map $\beta : V_1 \times V_2 \to W$ is called **bilinear** if for all $v_1 \in V_1$ the map $v_2 \mapsto \beta(v_1, v_2)$ is linear $V_2 \to W$ and for all $v_2 \in V_2$ the map $v_1 \mapsto \beta(v_1, v_2)$ is linear $V_1 \to W$.

Multilinear maps $V_1 \times V_2 \times \cdots \times V_n \to W$ are defined similarly.

The tensor product is a space which allows us to replace some bilinear (more generally multilinear) maps by linear maps.

Definition A.5 (Tensor product of vector spaces).
Let V_1 and V_2 be two vector spaces. A **tensor product** of V_1 and V_2 is a vector space U together with a bilinear map $\phi : V_1 \times V_2 \to U$ such that the following universal property holds: for any bilinear map $\beta : V_1 \times V_2 \to W$, there exists a unique linear map $\bar{\beta} : U \to W$ such that the diagram

$$\begin{tikzcd} V_1 \times V_2 \arrow[rr, "\beta"] \arrow[dr, "\phi"'] & & W \\ & U \arrow[ur, dotted, "\bar{\beta}"'] & \end{tikzcd}$$

commutes, that is $\beta = \bar{\beta} \circ \phi$.

Proving the uniqueness (up to canonical isomorphism) of an object defined by a universal property is a standard exercise in abstract nonsense.[1]

And since the tensor product is unique (up to canonical isomorphism), we can and will use the following notations (without danger of confusion)

$$\begin{aligned} U \;&=\; V_1 \otimes V_2 \qquad \text{and} \\ V_1 \times V_2 \;\ni\; (v_1, v_2) \;&\overset{\phi}{\mapsto}\; v_1 \otimes v_2 \;\in\; V_1 \otimes V_2. \end{aligned}$$

An explicit construction, which shows that tensor products indeed exist, is done in Exercise A.6. The same exercise establishes two fundamental properties of the tensor product:

- If $(v_i^{(1)})_{i \in I}$ is a linearly independent collection in V_1 and $(v_j^{(2)})_{j \in J}$ is a linearly independent collection in V_2, then the collection $\big(v_i^{(1)} \otimes v_j^{(2)}\big)_{(i,j) \in I \times J}$ is linearly independent in $V_1 \otimes V_2$.
- If the collection $(v_i^{(1)})_{i \in I}$ spans V_1 and the collection $(v_j^{(2)})_{j \in J}$ spans V_2, then the collection $\big(v_i^{(1)} \otimes v_j^{(2)}\big)_{(i,j) \in I \times J}$ spans the tensor product $V_1 \otimes V_2$.

It follows that if $(v_i^{(1)})_{i \in I}$ and $(v_j^{(2)})_{j \in J}$ are bases of V_1 and V_2, respectively, then

$$\Big(v_i^{(1)} \otimes v_j^{(2)}\Big)_{(i,j) \in I \times J}$$

[1] Indeed, if we suppose U' with a bilinear map $\phi' : V_1 \times V_2 \to U'$ is another tensor product, then the universal property of U gives a linear map $\bar{\phi}' : U \to U'$ such that $\phi' = \bar{\phi}' \circ \phi$. Likewise, the universal property of U' gives a linear map $\bar{\phi} : U' \to U$ such that $\phi = \bar{\phi} \circ \phi'$. Combining these we get

$$\mathrm{id}_U \circ \phi \;=\; \phi \;=\; \bar{\phi} \circ \phi' \;=\; \bar{\phi} \circ \bar{\phi}' \circ \phi.$$

But here are two ways of factorizing the map ϕ itself, so by the uniqueness requirement in the universal property we must have equality $\mathrm{id}_U = \bar{\phi} \circ \bar{\phi}'$. By a similar argument we get $\mathrm{id}_{U'} = \bar{\phi}' \circ \bar{\phi}$. We conclude that $\bar{\phi}$ and $\bar{\phi}'$ are isomorphisms (and inverses of each other).

is a basis of the tensor product $V_1 \otimes V_2$. In particular if V_1 and V_2 are finite dimensional, then

$$\dim(V_1 \otimes V_2) = \dim(V_1)\ \dim(V_2). \tag{A.1}$$

Exercise A.6 (A construction of the tensor product)**.**
We saw that the tensor product of vector spaces, defined by the universal property, is unique (up to isomorphism) if it exists. The purpose of this exercise is to show existence by an explicit construction, under the simplifying assumption that V and W are function spaces (it is easy to see that this can be assumed without loss of generality).

For any set X, denote by $\Bbbk^X$ the vector space of $\Bbbk$-valued functions on X, with addition and scalar multiplication defined pointwise. Assume that $V \subset \Bbbk^X$ and $W \subset \Bbbk^Y$ for some sets X and Y. For $f \in \Bbbk^X$ and $g \in \Bbbk^Y$, define $f \otimes g \in \Bbbk^{X\times Y}$ by

$$\big(f \otimes g\big)(x,y) = f(x)\, g(y).$$

Also set

$$V \otimes W = \mathrm{span}\big(f \otimes g \;\big|\; f \in V,\ g \in W\big),$$

so that the map $(f,g) \mapsto f \otimes g$ is a bilinear map $V \times W \to V \otimes W \subset \Bbbk^{X\times Y}$.

(a) Show that if $(f_i)_{i\in I}$ is a linearly independent collection in V and $(g_j)_{j\in J}$ is a linearly independent collection in W, then the collection $(f_i \otimes g_j)_{(i,j)\in I\times J}$ is linearly independent in $V \otimes W$.
(b) Show that if $(f_i)_{i\in I}$ is a collection that spans V and $(g_j)_{j\in J}$ is collection that spans W, then the collection $(f_i \otimes g_j)_{(i,j)\in I\times J}$ spans $V \otimes W$.
(c) Conclude that if $(f_i)_{i\in I}$ is a basis of V and $(g_j)_{j\in J}$ is a basis of W, then $(f_i \otimes g_j)_{(i,j)\in I\times J}$ is a basis of $V \otimes W$. Conclude furthermore that $V \otimes W$, equipped with the bilinear map $\phi(f,g) = f \otimes g$ from $V \times W$ to $V \otimes W$, satisfies the universal property defining the tensor product.

A tensor of the form $v^{(1)} \otimes v^{(2)}$ is called a **simple tensor**. By part (b) of the above exercise, any $t \in V_1 \otimes V_2$ can be written as a linear combination of simple tensors

$$t = \sum_{\alpha=1}^{n} v^{(1)}_\alpha \otimes v^{(2)}_\alpha,$$

for some $v^{(1)}_\alpha \in V_1$ and $v^{(2)}_\alpha \in V_2$, $\alpha = 1, 2, \ldots, n$. Note, however, that such an expression is by no means unique! The smallest n for which it is possible to write t as a sum of simple tensors is called the **rank** of the tensor, denoted by $n = \mathrm{rank}(t)$. An obvious upper bound is $\mathrm{rank}(t) \le \dim(V_1)\dim(V_2)$. One can do much better in general, as follows from the following useful observation.

Lemma A.7 (Linear independence in minimal expressions of tensors)**.**
Suppose that

$$t = \sum_{\alpha=1}^{n} v^{(1)}_\alpha \otimes v^{(2)}_\alpha,$$

where $n = \mathrm{rank}(t)$*. Then both* $(v^{(1)}_\alpha)_{\alpha=1}^n$ *and* $(v^{(2)}_\alpha)_{\alpha=1}^n$ *are linearly independent collections.*

Proof. Suppose, by contraposition, that there is a linear relation

$$\sum_{\alpha=1}^{n} c_\alpha v^{(1)}_\alpha = 0,$$

where not all the coefficients are zero. We may assume that $c_n = 1$. Thus we have $v_n^{(1)} = -\sum_{\alpha=1}^{n-1} c_\alpha v_\alpha^{(1)}$, and using bilinearity we simplify t as

$$t = \sum_{\alpha=1}^{n-1} v_\alpha^{(1)} \otimes v_\alpha^{(2)} + v_n^{(1)} \otimes v_n^{(2)} = \sum_{\alpha=1}^{n-1} v_\alpha^{(1)} \otimes v_\alpha^{(2)} - \sum_{\alpha=1}^{n-1} c_\alpha \, v_\alpha^{(1)} \otimes v_n^{(2)}$$
$$= \sum_{\alpha=1}^{n-1} v_\alpha^{(1)} \otimes \left(v_\alpha^{(1)} - c_\alpha v_n^{(2)} \right)$$

which contradicts minimality of $n = \mathrm{rank}(t)$. The linear independence of $(v_\alpha^{(2)})$ is proven similarly. □

As a consequence we get a better upper bound

$$\mathrm{rank}(t) \leq \min \{ \dim (V_1) , \dim (V_2) \} .$$

Taking tensor products with the one-dimensional vector space $\Bbbk$ does basically nothing: for any vector space V we can canonically identify

$$\begin{aligned} V \otimes \Bbbk &\cong V \\ v \otimes \lambda &\mapsto \lambda v \end{aligned} \qquad \text{and} \qquad \begin{aligned} \Bbbk \otimes V &\cong V \\ \lambda \otimes v &\mapsto \lambda v. \end{aligned}$$

By the obvious correspondence of bilinear maps $V_1 \times V_2 \to W$ and $V_2 \times V_1 \to W$, one also always gets a canonical identification

$$V_1 \otimes V_2 \cong V_2 \otimes V_1 .$$

Almost equally obvious correspondences give the canonical identifications

$$(V_1 \otimes V_2) \otimes V_3 \cong V_1 \otimes (V_2 \otimes V_3)$$

etc., which allow us to omit parentheses in multiple tensor products.

A slightly more interesting property than the above obvious identifications, is the existence of an embedding

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

which is obtained by associating to $v_2 \otimes \varphi$ the linear map

$$v_1 \mapsto \langle \varphi, v_1 \rangle \, v_2$$

(and extending linearly from the simple tensors to all tensors). The following exercise verifies among other things that this is indeed an embedding and that in the finite dimensional case the embedding becomes an isomorphism.

Exercise A.8 (The relation between $\mathrm{Hom}(V, W)$ and $W \otimes V^*$)**.**

(a) For $w \in W$ and $\varphi \in V^*$, we associate to $w \otimes \varphi$ the following map $V \to W$

$$v \mapsto \langle \varphi, v \rangle \, w.$$

Show that the linear extension of this defines an injective linear map

$$W \otimes V^* \longrightarrow \mathrm{Hom}(V, W).$$

(b) Show that if both V and W are finite dimensional, then the injective map in (a) is an isomorphism

$$W \otimes V^* \cong \mathrm{Hom}(V, W).$$

Show that under this identification, the rank of a tensor $t \in W \otimes V^*$ is the same as the rank of a matrix of the corresponding linear map $T \in \mathrm{Hom}(V, W)$.

Definition A.9 (Tensor product of linear maps)**.**
When
$$f : V_1 \to W_1 \qquad \text{and} \qquad g : V_2 \to W_2$$
are linear maps, then there is a linear map
$$f \otimes g : V_1 \otimes V_2 \to W_1 \otimes W_2$$
defined by the condition
$$(f \otimes g)(v_1 \otimes v_2) \;=\; f(v_1) \otimes g(v_2) \qquad \text{for all } v_1 \in V_1,\; v_2 \in V_2.$$

The above map clearly depends bilinearly on (f, g), so we get a canonical map
$$\mathrm{Hom}(V_1, W_1) \otimes \mathrm{Hom}(V_2, W_2) \;\hookrightarrow\; \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2),$$
which is easily seen to be injective. When all the vector spaces V_1, W_1, V_2, W_2 are finite dimensional, then the dimensions of both sides are given by
$$\dim(V_1)\;\dim(V_2)\;\dim(W_1)\;\dim(W_2)\,,$$
so in this case the canonical map is an isomorphism
$$\mathrm{Hom}(V_1, W_1) \otimes \mathrm{Hom}(V_2, W_2) \;\cong\; \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2).$$

As a particular case of the above, interpreting the dual of a vector space V as $V^* = \mathrm{Hom}(V, \Bbbk)$ and using $\Bbbk \otimes \Bbbk \cong \Bbbk$, we see that the tensor product of duals sits inside the dual of the tensor product. Explicitly, if V_1 and V_2 are vector spaces and $\varphi_1 \in V_1^*$, $\varphi_2 \in V_2^*$, then
$$v_1 \otimes v_2 \;\mapsto\; \langle \varphi_1, v_1\rangle \, \langle \varphi_2, v_2\rangle$$
defines an element of the dual of $V_1 \otimes V_2$. To summarize, we have an embedding
$$V_1^* \otimes V_2^* \;\hookrightarrow\; (V_1 \otimes V_2)^*.$$
If V_1 and V_2 are finite dimensional this becomes an isomorphism
$$V_1^* \otimes V_2^* \;\cong\; (V_1 \otimes V_2)^*.$$

The transpose behaves well under the tensor product of linear maps.

Lemma A.10. *When $f : V_1 \to W_1$ and $g : V_2 \to W_2$ are linear maps, then the map*
$$f \otimes g : V_1 \otimes V_2 \to W_1 \otimes W_2$$
has a transpose $(f \otimes g)^*$ which makes the following diagram commute

$$\begin{tikzcd}
(W_1 \otimes W_2)^* \arrow[r, "(f\otimes g)^*"] & (V_1 \otimes V_2)^* \\
W_1^* \otimes W_2^* \arrow[u, hook] \arrow[r, "f^*\otimes g^*"] & V_1^* \otimes V_2^* \arrow[u, hook]
\end{tikzcd}$$

Proof. Indeed, for $\varphi \in W_1^*$, $\psi \in W_2^*$ and any simple tensor $v_1 \otimes v_2 \in V_1 \otimes V_2$ we compute
$$\begin{aligned}
\langle (f^* \otimes g^*)(\varphi \otimes \psi), v_1 \otimes v_2\rangle &= \langle f^*(\varphi) \otimes g^*(\psi), v_1 \otimes v_2\rangle \\
&= \langle f^*(\varphi), v_1\rangle \, \langle g^*(\psi), v_2\rangle \\
&= \langle \varphi, f(v_1)\rangle \, \langle \psi, g(v_2)\rangle \\
&= \langle \varphi \otimes \psi, f(v_1) \otimes g(v_2)\rangle \\
&= \langle \varphi \otimes \psi, (f \otimes g)(v_1 \otimes v_2)\rangle \\
&= \langle (f \otimes g)^*(\varphi \otimes \psi), v_1 \otimes v_2\rangle.
\end{aligned}$$

□

2. On diagonalization of matrices

In this section, vector spaces are over the field $\mathbb{C}$ of complex numbers.

Recall first the following definitions.

Definition A.11 (Characteristic polynomial and minimal polynomial).
The **characteristic polynomial** of a matrix $A \in \mathbb{C}^{n\times n}$ is

$$p_A(x) \;=\; \det\big(x\mathbb{I} - A\big).$$

The **minimal polynomial** of a matrix A is the polynomial q_A of smallest positive degree such that[2] $q_A(A) = 0$, with the coefficient of highest degree term equal to 1.

The Cayley-Hamilton theorem states that the characteristic polynomial evaluated at the matrix itself is the zero matrix, that is $p_A(A) = 0$ for any square matrix A. An equivalent statement is that the minimal polynomial $q_A(x)$ divides the characteristic polynomial $p_A(x)$. These facts follow explicitly from the Jordan normal form discussed later in this section.

Motivation and definition of generalized eigenvectors

Given a square matrix A, it is often convenient to diagonalize A. This means finding an invertible matrix P ("a change of basis"), such that the conjugated matrix $P\,A\,P^{-1}$ is diagonal. If, instead of matrices, we think of a linear operator A from vector space V to itself, the equivalent question is finding a basis for V consisting of eigenvectors of A.

Recall from basic linear algebra that (for example) any real symmetric matrix can be diagonalized. Unfortunately, this is not the case with all matrices.

Example A.12 (A non-diagonalizable matrix).
Let $\lambda \in \mathbb{C}$ and

$$A \;=\; \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \;\in\; \mathbb{C}^{3\times 3}.$$

The characteristic polynomial of A is

$$p_A(x) \;=\; \det(x\mathbb{I} - A) \;=\; (x-\lambda)^3,$$

so we know that A has no other eigenvalues but λ. It follows from $\det(A - \lambda\mathbb{I}) = 0$ that the eigenspace pertaining to the eigenvalue λ is nontrivial, $\dim\big(\mathrm{Ker}\,(A-\lambda\mathbb{I})\big) > 0$. Note that

$$A - \lambda\mathbb{I} \;=\; \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

[2]Square matrices can be multiplied together, and multiplied by scalars. In particular it is meaningful to evaluate a polynomial at a square matrix: the polynomial $c_0 + c_1 x + \cdots + c_m x^m$ evaluated at the matrix $A \in \mathbb{C}^{n\times n}$ is the matrix $c_0 + c_1 A + \cdots + c_m A^m$.

so that the image of A is two dimensional, $\dim(\mathrm{Im}(A - \lambda\mathbb{I})) = 2$. By rank-nullity theorem,

$$\dim(\mathrm{Im}(A - \lambda\mathbb{I})) + \dim(\mathrm{Ker}(A - \lambda\mathbb{I})) \;=\; \dim\left(\mathbb{C}^3\right) \;=\; 3,$$

so the eigenspace pertaining to λ must be one-dimensional. Thus the maximal number of linearly independent eigenvectors of A we can have is one — in particular, there doesn't exist a basis of $\mathbb{C}^3$ consisting of eigenvectors of A.

We still take a look at the action of A in some basis. Let

$$w_1 \;=\; \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad w_2 \;=\; \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad w_3 \;=\; \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then the following "string" indicates how $A - \lambda\mathbb{I}$ maps these vectors

$$w_3 \overset{A-\lambda}{\mapsto} w_2 \overset{A-\lambda}{\mapsto} w_1 \overset{A-\lambda}{\mapsto} 0.$$

In particular we see that $(A - \lambda\mathbb{I})^3 = 0$.

The "string" in the above example illustrates and motivates the following definition.

Definition A.13 (Generalized eigenvectors and generalized eigenspaces)**.**
Let V be a vector space and $A : V \to V$ be a linear map. A vector $v \in V$ is said to be a **generalized eigenvector** of eigenvalue $\lambda \in \mathbb{C}$ if for some positive integer p we have $(A - \lambda\mathbb{I})^p\, v = 0$. The set of these generalized eigenvectors is called the **generalized eigenspace** of A pertaining to eigenvalue λ.

With $p = 1$ the above would correspond to the usual eigenvectors.

The Jordan canonical form

Although not every matrix has a basis of eigenvectors, we will see that every complex square matrix has a basis of generalized eigenvectors. More precisely, if V is a finite dimensional complex vector space and $A : V \to V$ is a linear map, then there exists eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ of A (not necessarily distinct) and a basis $\{w_m^{(j)} : 1 \le j \le k, 1 \le m \le n_j\}$ of V which consists of "strings" as follows

$$\begin{array}{ccccccccccc}
w_{n_1}^{(1)} & \overset{A-\lambda_1}{\mapsto} & w_{n_1-1}^{(1)} & \overset{A-\lambda_1}{\mapsto} & \cdots & \overset{A-\lambda_1}{\mapsto} & w_2^{(1)} & \overset{A-\lambda_1}{\mapsto} & w_1^{(1)} & \overset{A-\lambda_1}{\mapsto} & 0 \\
w_{n_2}^{(2)} & \overset{A-\lambda_2}{\mapsto} & w_{n_2-1}^{(2)} & \overset{A-\lambda_2}{\mapsto} & \cdots & \overset{A-\lambda_2}{\mapsto} & w_2^{(2)} & \overset{A-\lambda_2}{\mapsto} & w_1^{(2)} & \overset{A-\lambda_2}{\mapsto} & 0 \\
 & & & & & & \vdots & & \vdots & & \vdots \\
w_{n_k}^{(k)} & \overset{A-\lambda_k}{\mapsto} & w_{n_k-1}^{(k)} & \overset{A-\lambda_k}{\mapsto} & \cdots & \overset{A-\lambda_k}{\mapsto} & w_2^{(k)} & \overset{A-\lambda_k}{\mapsto} & w_1^{(k)} & \overset{A-\lambda_k}{\mapsto} & 0.
\end{array} \tag{A.2}$$

Note that in this basis the matrix of A takes the "block diagonal form"

$$A \;=\; \begin{bmatrix} J_{\lambda_1;n_1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & J_{\lambda_2;n_2} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & J_{\lambda_3;n_3} & & \mathbf{0} \\ \vdots & \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & J_{\lambda_k;n_k} \end{bmatrix}, \tag{A.3}$$

where the blocks correspond to the subspaces spanned by $w_1^{(j)}, w_2^{(j)}, \ldots, w_{n_j}^{(j)}$ and the matrices of the blocks are the following "Jordan blocks"

$$J_{\lambda_j;n_j} \;=\; \begin{bmatrix} \lambda_j & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_j & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_j & & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_j & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_j \end{bmatrix} \in \mathbb{C}^{n_j \times n_j}.$$

Definition A.14 (Jordan canonical form)**.**
A matrix of the form (A.3) is said to be in **Jordan canonical form** (or **Jordan normal form**).

The characteristic polynomial of the a matrix A in Jordan canonical form is

$$p_A(x) \;=\; \det\left(x\mathbb{I} - A\right) \;=\; \prod_{j=1}^{k} (x - \lambda_j)^{n_j}.$$

Note also that if we write a block $J_{\lambda;n} = \lambda\mathbb{I} + N$ as a sum of diagonal part $\lambda\mathbb{I}$ and upper triangular part N, then the latter is nilpotent: $N^n = 0$. In particular the assertion $p_A(A) = 0$ of the Cayley-Hamilton theorem can be seen immediately for matrices which are in Jordan canonical form.

Definition A.15 (Similarity of matrices)**.**
Two $n \times n$ square matrices A and B are said to be **similar** if $A = P\,B\,P^{-1}$ for some invertible matrix P.

It is in this sense that any complex square matrix can be put to Jordan canonical form, the matrix P implements a change of basis to a basis consisting of the strings of the above type. Below is a short and concrete proof.

Theorem A.16 (Jordan canonical form)**.**
Given any complex $n \times n$ matrix A, there exists an invertible matrix P such that the conjugated matrix $P\,A\,P^{-1}$ is in Jordan canonical form.

Proof. In view of the above discussion it is clear that the statement is equivalent to the following: if V is a finite dimensional complex vector space and $A : V \to V$ a linear map, then there exists a basis of V consisting of strings as in (A.2).

We prove the statement by induction on $n = \dim(V)$. The case $n = 1$ is clear. As an induction hypothesis, assume that the statement is true for all linear maps of vector spaces of dimension less than n.

Take any eigenvalue λ of A (any root of the characteristic polynomial). Note that

$$\dim\left(\mathrm{Ker}\left(A - \lambda\mathbb{I}\right)\right) > 0,$$

and since $n = \dim\left(\mathrm{Ker}\left(A - \lambda\mathbb{I}\right)\right) + \dim\left(\mathrm{Im}\left(A - \lambda\mathbb{I}\right)\right)$, the dimension of the image of $A - \lambda\mathbb{I}$ is strictly less than n. Denote

$$R \;=\; \mathrm{Im}\left(A - \lambda\mathbb{I}\right) \qquad \text{and} \qquad r \;=\; \dim(R) \;<\; n.$$

Note that R is an invariant subspace for A, that is $A\,R \subset R$ (indeed, $A\,(A-\lambda\mathbb{I})\,v = (A-\lambda\mathbb{I})\,A\,v$). We can use the induction hypothesis to the restriction of A to R, to find a basis

$$\{w_m^{(j)} \,:\, 1 \le j \le k, 1 \le m \le n_j\}$$

of R in which the action of A is described by the strings as in (A.2).

Let $q = \dim(R \cap \operatorname{Ker}(A-\lambda\mathbb{I}))$. This means that in R there are q linearly independent eigenvectors of A with eigenvalue λ. The vectors at the right end of the strings span the eigenspaces of A in R, so we assume without loss of generality that the last q strings correspond to eigenvalue λ and others to different eigenvalues: $\lambda_1, \lambda_2, \ldots, \lambda_{k-q} \neq \lambda$ and $\lambda_{k-q+1} = \lambda_{k-q+2} = \cdots = \lambda_k = \lambda$. For all j such that $k-q < j \le k$ the vector $w_{n_j}^{(j)}$ is in R, so we can choose

$$y^{(j)} \in V \quad \text{such that} \quad (A-\lambda\mathbb{I})\,y^{(j)} \;=\; w_{n_j}^{(j)}.$$

The vectors $y^{(j)}$ extend the last q strings from the left.

Find vectors

$$z^{(1)}, z^{(2)}, \ldots, z^{(n-r-q)}$$

which complete the linearly independent collection

$$w_1^{(k-q+1)}, \ldots, w_1^{(k-1)}, w_1^{(k)}$$

to a basis of $\operatorname{Ker}(A-\lambda\mathbb{I})$. We have now found n vectors in V, which form strings as follows

$$\begin{array}{ccccccccc}
 & & & & & & z^{(1)} & \overset{A-\lambda}{\mapsto} & 0 \\
 & & & & & & \vdots & & \vdots \\
 & & & & & & z^{(n-r-q)} & \overset{A-\lambda}{\mapsto} & 0 \\
 & & w_{n_1}^{(1)} & \overset{A-\lambda_1}{\mapsto} & \cdots & \overset{A-\lambda_1}{\mapsto} & w_1^{(1)} & \overset{A-\lambda_1}{\mapsto} & 0 \\
 & & \vdots & & & & \vdots & & \vdots \\
 & & w_{n_{k-q}}^{(k-q)} & \overset{A-\lambda_{k-q}}{\mapsto} & \cdots & \overset{A-\lambda_{k-q}}{\mapsto} & w_1^{(k-q)} & \overset{A-\lambda_{k-q}}{\mapsto} & 0 \\
y^{(k-q+1)} & \overset{A-\lambda}{\mapsto} & w_{n_{k-q+1}}^{(k-q+1)} & \overset{A-\lambda}{\mapsto} & \cdots & \overset{A-\lambda}{\mapsto} & w_1^{(k-q+1)} & \overset{A-\lambda}{\mapsto} & 0 \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
y^{(k)} & \overset{A-\lambda}{\mapsto} & w_{n_k-1}^{(k)} & \overset{A-\lambda}{\mapsto} & \cdots & \overset{A-\lambda}{\mapsto} & w_1^{(k)} & \overset{A-\lambda}{\mapsto} & 0.
\end{array}$$

It suffices to show that these vectors are linearly independent. Suppose that a linear combination of them vanishes

$$\sum_{j=k-q+1}^{k} \alpha_j\, y^{(j)} + \sum_{j,m} \beta_{j,m}\, w_m^{(j)} + \sum_{l=1}^{n-r-q} \gamma_l\, z^{(l)} \;=\; 0.$$

From the string diagram we see that the image of this linear combination under $A-\lambda\mathbb{I}$ is a linear combination of the vectors $w_m^{(j)}$, which are linearly independent, and since the coefficient of $w_{n_j}^{(j)}$ is α_j, we get $\alpha_j = 0$ for all j. Now recalling that $\{w_m^{(j)}\}$ is a basis of R, and $\{w_1^{(j)} : k-q < j \le k\} \cup \{z^{(l)}\}$ is a basis of $\operatorname{Ker}(A-\lambda\mathbb{I})$, and $\{w_1^{(j)} : k-q < j \le k\}$ is a basis of $R \cap \operatorname{Ker}(A-\lambda\mathbb{I})$, we see that all the coefficients in the linear combination must vanish. This finishes the proof. □

Exercise A.17 (Around the Jordan normal form)**.**

(a) Find two matrices $A, B \in \mathbb{C}^{n\times n}$, which have the same minimal polynomial and the same characteristic polynomial, but which are not similar.

(b) Show that the Jordan normal form of a matrix $A \in \mathbb{C}^{n\times n}$ is unique up to permutation of the Jordan blocks. In other words, if $C_1 = P_1\,A\,P_1^{-1}$ and $C_2 = P_2\,A\,P_2^{-1}$ are both in Jordan normal form, C_1 with blocks $J_{\lambda_1,n_1}, \ldots J_{\lambda_k;n_k}$ and C_2 with blocks $J_{\lambda'_1,n'_1}, \ldots J_{\lambda'_l;n'_l}$, then $k = l$ and there is a permutation $\sigma \in S_k$ such that $\lambda_j = \lambda'_{\sigma(j)}$ and $n_j = n'_{\sigma(j)}$ for all $j = 1, 2, \ldots, k$.

(c) Show that any two matrices with the same Jordan normal form up to permutation of blocks are similar.

In view of Theorem A.16 and the remark after Definition A.14, the Cayley-Hamilton theorem follows for complex matrices (and for real matrices, as a particular case — although the Jordan canonical form can not be used directly).

Theorem A.18 (Cayley-Hamilton theorem)**.**
Let $A \in \mathbb{C}^{n\times n}$ be a complex square matrix, and denote by $p_A(x) := \det\big(x\mathbb{I} - A\big)$ its characteristic polynomial. Then we have

$$p_A(A) = 0.$$

In other words, the minimal polynomial $q_A(x)$ of A divides the characteristic polynomial, $q_A(x) \mid p_A(x)$.

Another very useful consequence is the following sufficient condition for diagonalizability.

Proposition A.19 (Simple roots guarantee diagonalizability)**.**
Suppose that the characteristic polynomial $p_A(x) := \det\big(x\mathbb{I} - A\big)$ of a complex square matrix $A \in \mathbb{C}^{n\times n}$ has no multiple roots. Then the matrix A is diagonalizable.

Bibliography

[FH91] William Fulton and Joe Harris. *Representation theory: a first course*. Graduate Texts in Mathematics. Springer-Verlag, 1991.

[Sim96] Barry Simon. *Representations of Finite and Compact Groups*. Graduate studies in mathematics. American Mathematical Society, 1996.