RUNE HAUGSENG

INTRODUCTION TO LIE THEORY (MA3407)

NTNU

Fall 2022 (last updated: November 18, 2022)

Contents

Sections marked with (⋆) are not relevant for the exam.

Chapter 1

Introduction

In previous courses on algebra you have studied groups, and seen that group actions encode *symmetries*. For example, we might consider the symmetry group of a square under orientation-preserving rigid motions; this is the cyclic group C_4, which is a finite group with 4 elements (corresponding to rotations by 0, $\pi/2$, π and $3\pi/2$ radians). On the other hand, if we consider orientation-preserving rigid motions of a *circle*, we get rotations parametrized by an arbitrary angle $0 \leq \theta \leq 2\pi$, giving the group $\mathbb{R}/2\pi$.[1] There is an important difference between these two situations, however: while in the case of the square the discrete set of 4 symmetries seems to completely encode what's going on, for the circle it is very natural to think of the rotations as depending *continuously* on the angle θ we rotate by, or even *smoothly* (that is, differentiable infinitely many times). We would therefore like to equip the latter group with a topological (or smooth) structure, and say that it acts continuously (or smoothly) on the circle.

Such "smooth" groups are called *Lie groups* after the Norwegian mathematician Sophus Lie. Personally, I think there is a strong argument to be made that Lie was the "greatest", or at least most influential, Norwegian mathematician, despite all the things named after Abel. That said, the nationality of the originator of a subject is hardly a good reason to be interested in it — and in any case much of Lie's work on Lie groups was in collaboration with the German mathematician Friedrich Engel[2], while the modern theory of Lie groups and Lie algebras, as we will (partially) present it in this course, was substantially developed by Wilhelm Killing (German) and Élie Cartan (French). A better reason for learning about Lie groups, Lie algebras and their representations is that it is not only a fascinating topic in its own right, but also a central topic in modern mathematics with connections to many different fields (and many applications in physics, too).

Lie's original motivation for introducing Lie groups was to develop a "Galois

[1] Or equivalently the group $SO_2(\mathbb{R})$ of 2×2 rotation matrices, or the group of unit-length complex numbers under multiplication.

[2] Not, of course, to be confused with Friedrich Engels, the collaborator of Karl Marx.

theory for differential equations". Just as the Galois group of a polynomial permutes its roots, Lie wanted a group that acted on the solutions to a differential equation. However, while this idea has found some applications in analysis, it seems not to have been as interesting as Lie had hoped. Instead, Lie groups have turned out to be important in many areas of mathematics, including differential geometry, algebraic topology, representation theory, harmonic analysis, and even (through the closely related theory of algebraic groups) in algebraic geometry and number theory. Lie groups are also of great importance in physics, since they are the kind of group that arises naturally as the symmetry groups of a physical system or theory.

What will we do in this course?

Setting up the general theory of Lie groups unfortunately requires a substantial background in differential geometry, as well as input from analysis and topology. In this course we will therefore focus on the special case of *matrix groups*, which are, roughly speaking, groups that consist of matrices under matrix multiplication. For this special class of groups we can do a lot using linear algebra together with a bit of basic analysis and topology. This class of groups also includes many, if not most, of the important examples of Lie groups.

In the first part of the course, our goal is to show that many aspects of a matrix group G are controlled by a seemingly much simpler structure: its *Lie algebra* $\mathfrak{g}$. This is a finite-dimensional $\mathbb{R}$-vector space together with a bilinear operation

$$[-,-]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

with certain properties, which encodes the "infinitesimal behaviour" of multiplication in G close to the identity.

In the second part, we will look at *representations* of matrix groups and their Lie algebras. Roughly speaking, a representation of a Lie group G is a continuous linear action of G on a vector space V (over $\mathbb{R}$ or more likely $\mathbb{C}$). It is not so easy to explain why these are so important, but one vague idea is that *linearization* is often an important strategy in mathematics: if G acts on some object X, we might be able to associate a vector space V_X to X, which then typically inherits an action of G and so gives us a representation. From this we can derive (sometimes without losing any information) a representation of the Lie algebra $\mathfrak{g}$ of G. If $\mathfrak{g}$ is nice, there is a classification of its representations, and so it is plausible that we can completely understand V_X as a $\mathfrak{g}$-representation. If we are lucky, we might then be able to learn something interesting about the original object X (with its G-action) from this.

Representations also play an important role in quantum physics: Roughly speaking, if G is the symmetry group of a physical system, the corresponding $\mathbb{C}$-vector space of quantum states will be a representation of G (or of its Lie algebra).

We will focus on representations of *complex semisimple* Lie algebras, which includes many of those that arise from matrix groups. Hopefully we will also have time to take a look at the remarkable classification theorem for this class of Lie algebras, which shows that all but 5 exceptional examples come from three infinite families of matrix groups.

Sources and references

Much of the material here is drawn from Hall's book [2], and these notes should largely be viewed as a supplement to this textbook. Other sources that have been useful include

- a number of articles from the blog [5] (in particular relating to quaternions),
- the books [1] by Fulton and Harris, [3] by Humphreys, and [4] by Stillwell,
- and of course Wikipedia…

Part I

Lie groups and Lie algebras

Chapter 2

Matrix groups

2.1 Matrix groups — basic definitions

Definition 2.1.1. We write $\mathrm{M}_n(\mathbb{R})$ for the set of $n \times n$ real matrices; this is evidently isomorphic to $\mathbb{R}^{n^2}$ as an $\mathbb{R}$-vector space (by viewing an $n \times n$ matrix as a list of its n^2 entries). We view $M_n(\mathbb{R})$ as a topological space with the topology induced by the usual metric on $\mathbb{R}^{n^2}$.

Definition 2.1.2. We write $\mathrm{GL}_n(\mathbb{R})$ for the subset of $\mathrm{M}_n(\mathbb{R})$ consisting of invertible matrices; we regard this as a topological space with the subspace topology inherited from $\mathrm{M}_n(\mathbb{R})$.

Exercise 2.1. Check that $\mathrm{GL}_n(\mathbb{R})$ is a group (the *general linear group* of $\mathbb{R}^n$) under matrix multiplication.

Lemma 2.1.3. $\mathrm{GL}_n(\mathbb{R})$ *is an open subset of* $\mathrm{M}_n(\mathbb{R})$.

Proof. The determinant of an $n \times n$ matrix is by definition a polynomial in the matrix entries, and so $\det\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ is a continuous function. We can regard $\mathrm{GL}_n(\mathbb{R})$ as the preimage $\det^{-1}(\mathbb{R} \setminus \{0\})$, which is open in $\mathrm{M}_n(\mathbb{R})$ since $\mathbb{R} \setminus \{0\}$ is open in $\mathbb{R}$. □

Slightly informally, we want to define a *matrix group* to be a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ for some n, i.e. a closed subset $G \subseteq \mathrm{GL}_n(\mathbb{R})$ that contains the identity matrix and is closed under matrix multiplication and taking inverses.

The reason this definition is informal is that we don't really want to consider the inclusion in $\mathrm{GL}_n(\mathbb{R})$ as part of the data of a matrix group. For one thing, the same group might have several reasonable descriptions in terms of matrices, and we might not want to prefer one over the others. As another justification, consider the following example:

Exercise 2.2.

(i) Show that if we regard a complex number $z = x + iy$ as the 2×2 real matrix $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$ then multiplication in $\mathbb{C}$ corresponds to matrix multiplication.

(ii) ($\star$) Show that the matrices in $M \in \mathrm{M}_2(\mathbb{R})$ that are of the form above are exactly those such that $M^T M = \lambda I$ for some $\lambda \in \mathbb{R}$ and $\det(M) \geq 0$.

(iii) Let $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Show that the matrices in $M \in \mathrm{M}_2(\mathbb{R})$ that are of the form above are exactly those such that $JM = MJ$. (Note that $J \in \mathrm{M}_2(\mathbb{R})$ corresponds to $i \in \mathbb{C}$, so in some sense this says that the matrices that represent complex numbers are those that "commute with i".)

(iv) Show that under this identification the group $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$ of invertible complex numbers (under complex multiplication) corresponds to a closed subgroup of $\mathrm{GL}_2(\mathbb{R})$.

Here we would like to say that $\mathbb{C}^\times$ is a matrix group (rather than just isomorphic to one), and without having to specify this particular description in terms of 2×2 matrices.

Before we get to the final definition, let's first look at two aspects of the informal one: First, why do we only consider *closed* subgroups of $\mathrm{GL}_n(\mathbb{R})$?

To motivate this, let's consider an example of the kind of groups this excludes:

Example 2.1.4. For any n, the set $\mathrm{GL}_n(\mathbb{Q})$ of $n \times n$ invertible matrices with *rational* entries is a non-closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. (This is because the explicit formula for the inverse of a matrix implies that the inverse of a rational matrix is again rational.)

In particular, we see that a non-closed subgroup may well not be a manifold, meaning that it does not have to look locally like an open subset of $\mathbb{R}^k$ for some k. In this sense it is not really a "smooth" group. Another kind of bad behaviour we exclude is "bad" embeddings of nice groups:

Example 2.1.5. Let U_1 be the group of unit-length complex numbers under multiplication. Then we can consider the subgroup

$$G = \left\{(e^{it}, e^{iqt}) : t \in \mathbb{R}\right\} \subseteq U_1 \times U_1$$

where $q \in \mathbb{R}$ is fixed. (We will see in the next section that $U_1 \times U_1$ is indeed a matrix group.) Topologically, U_1 is a circle, $U_1 \times U_1$ is a torus, and G describes a "line of slope q" that winds around the torus. If q is rational, then this line will eventually end up back where it started, and G is just a closed subgroup isomorphic to U_1 again. However, if q is *irrational* then this line winds around infinitely many times, and gets arbitrarily close to any point on the torus. Its closure is therefore all of $U_1 \times U_1$. In particular, G is then not a closed subgroup, though as a group it is actually isomorphic to $\mathbb{R}$ under addition.

Secondly, we might ask why we only look at *real* matrices, instead of complex ones. The answer is that this doesn't matter, because of the following description of complex matrices:

Example 2.1.6. Let $\mathrm{M}_n(\mathbb{C}) \cong \mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$ be the set of $n \times n$ complex matrices, and $\mathrm{GL}_n(\mathbb{C})$ the subset of invertible complex matrices. By expanding each complex number into a 2×2 real matrix as in Exercise 2.2 we get an embedding $\iota\colon \mathrm{M}_n(\mathbb{C}) \hookrightarrow \mathrm{M}_{2n}(\mathbb{R})$, which identifies $\mathrm{M}_n(\mathbb{C})$ with a *closed* subset of $\mathrm{M}_{2n}(\mathbb{R})$; this embedding is moreover $\mathbb{R}$-linear, and is compatible with (complex and real) matrix multiplication (in the sense that $\iota(AB) = \iota(A)\iota(B)$). One can then prove that $\mathrm{GL}_n(\mathbb{C})$ corresponds to $\mathrm{M}_n(\mathbb{C})\cap\mathrm{GL}_{2n}(\mathbb{R})$ via this embedding (for example, by computing that $\det \iota(A) = |\det(A))|^2$, so that a matrix $A \in \mathrm{M}_n(\mathbb{C})$ is invertible if and only if $\iota(A)$ is invertible). We then see that $\mathrm{GL}_n(\mathbb{C})$ is isomorphic to a closed subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$.

Exercise 2.3.

(i) Fill in the details in Example 2.1.6.

(ii) Let $J \in \mathrm{M}_2(\mathbb{R})$ be the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and define $J_n \in \mathrm{M}_{2n}(\mathbb{R})$ as the matrix

$$\begin{pmatrix} J & & 0 \\ & \ddots & \\ 0 & & J \end{pmatrix}.$$

Show that a matrix $M \in \mathrm{M}_{2n}(\mathbb{R})$ is in the image of $\mathrm{M}_n(\mathbb{C})$ under ι if and only if $JM = MJ$.

Any closed subgroup of $\mathrm{GL}_n(\mathbb{C})$ is therefore isomorphic to a closed subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$. Conversely, $\mathrm{GL}_n(\mathbb{R})$ is evidently a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$ via the inclusion $\mathbb{R} \hookrightarrow \mathbb{C}$, so a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ is also a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$. Thus we can equivalently consider matrix groups as either closed subgroups of $\mathrm{GL}_n(\mathbb{R})$ or of $\mathrm{GL}_n(\mathbb{C})$, and we will use whichever description is more convenient in each case.

Now let's turn to the precise definition of a matrix group. To give this we first need to introduce some terminology:

Definition 2.1.7. A *topological group* is a group G together with a topology on the set G such that group multiplication and inverses define continuous maps $G \times G \to G$ and $G \to G$. A *continuous homomorphism* is a homomorphism between topological groups that is also continuous.

Exercise 2.4. Show that $\mathrm{GL}_n(\mathbb{R})$ is a topological group under matrix multiplication. (Hint: Both matrix multiplication and inverses are given entrywise by polynomials in the matrix entries or quotients thereof.)

Definition 2.1.8. We say that a topological group G is a *matrix group* if there exists a continuous isomorphism $G \cong G'$ where G' is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ for some n.

Remark 2.1.9. Note that a particular choice of embedding is *not* part of the data of a matrix group. In particular, the natural notion of a *homomorphism of matrix* *groups* is just that of a continuous homomorphism, which is *not* required to be compatible with the embeddings of the two groups in general linear groups. However, sometimes we *do* want to consider a matrix group G together with a chosen embedding in $\mathrm{GL}_n(\mathbb{R})$:

Definition 2.1.10. An *embedded matrix group* is a pair (G, j) consisting of a topological group G and a continuous homomorphism $j\colon G \to \mathrm{GL}_n(\mathbb{R})$ that is a homeomorphism onto its image and whose image is closed in $\mathrm{GL}_n(\mathbb{R})$.

We will, however, try not to be too pedantic about choices of embeddings, and will often just refer to "a matrix group $G \subseteq \mathrm{GL}_n(\mathbb{R})$ (or $\mathrm{GL}_n(\mathbb{C})$)".

2.2 Examples of matrix groups

Example 2.2.1 (General linear groups)**.** The group $\mathrm{GL}_n(\mathbb{R})$ is tautologically a matrix group, and we saw above that so is $\mathrm{GL}_n(\mathbb{C})$ for all $n \geq 0$.

Example 2.2.2 (Special linear groups)**.** The *special linear group* $\mathrm{SL}_n(\mathbb{R})$ is the subgroup of $\mathrm{GL}_n(\mathbb{R})$ consisting of matrices with determinant 1. Since the determinant is continuous, this is a closed subgroup, being the preimage of the closed subset $\{1\} \subseteq \mathbb{R}$. Similarly, the group $\mathrm{SL}_n(\mathbb{C})$ of complex $n \times n$ matrices with determinant 1 is a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$.

Example 2.2.3 (Finite groups)**.** The symmetric group S_n (that is, the group of permutations of n letters) acts faithfully on $\mathbb{R}^n$ by permuting the coordinates. These symmetries are linear, so this gives an embedding of S_n in $\mathrm{GL}_n(\mathbb{R})$; since S_n is finite the image is obviously closed. Thus S_n is a matrix group. Moreover, since an arbitrary finite group G is always a subgroup of $S_{|G|}$, it follows that all finite groups (with the discrete topology) are matrix groups.

Notation 2.2.4. For $X \in \mathrm{M}_n(\mathbb{C})$, we will write X^T for the *transpose* of X, i.e. the matrix with entries

$$(X^T)_{ij} = X_{ji}.$$

Example 2.2.5 (Orthogonal groups)**.** Recall that a matrix $X \in \mathrm{M}_n(\mathbb{R})$ is *orthogonal* if $X^T X = I$. Equivalently, X is orthogonal if its rows (or its columns) are orthogonal vectors in $\mathbb{R}^n$, or if X preserves the inner product on $\mathbb{R}^n$ in the sense that we have $\langle Xv, Xw\rangle = \langle v, w\rangle$ for all $v, w \in \mathbb{R}^n$. We write $\mathrm{O}_n(\mathbb{R})$ for the set of orthogonal matrices in $\mathrm{M}_n(\mathbb{R})$. From the equation $X^T X = I$ it follows that $(\det X)^2 = 1$, so that X is invertible with the inverse necessarily being X^T, which is thus also orthogonal. Moreover, $\mathrm{O}_n(\mathbb{R})$ is closed under matrix multiplication, since for $X, Y \in \mathrm{O}_n(\mathbb{R})$ we have

$$(XY)^T(XY) = Y^T X^T XY = Y^T Y = I.$$

Thus $\mathrm{O}_n(\mathbb{R})$ is a subgroup of $\mathrm{GL}_n(\mathbb{R})$. Moreover, it is a closed subgroup since the equation $X^T X = I$ is a set of polynomial equations in the entries of X. We further have the subgroup $\mathrm{SO}_n(\mathbb{R}) \subseteq \mathrm{O}_n(\mathbb{R})$ consisting of orthogonal matrices with determinant 1; this is again closed in $\mathrm{GL}_n(\mathbb{R})$, since it is the intersection of the closed subsets $\mathrm{O}_n(\mathbb{R})$ and $\mathrm{SL}_n(\mathbb{R})$. Similarly, we have a closed subgroup $\mathrm{O}_n(\mathbb{C})$ consisting of those matrices $X \in \mathrm{M}_n(\mathbb{C})$ such that $X^T X = I$ (but note that these "orthogonal" matrices are *not* the ones that preserve the complex inner product, which we will rather see in the next example), with a closed subgroup $\mathrm{SO}_n(\mathbb{C}) \subseteq \mathrm{O}_n(\mathbb{C})$ comprising the matrices with determinant 1.

Notation 2.2.6. For $X \in \mathrm{M}_n(\mathbb{C})$, we will write $X^\dagger$ for the *conjugate transpose* of X, i.e. the matrix $\overline{X}^T$ with entries

$$X^\dagger_{ij} = (\overline{X}^T)_{ij} = \overline{x_{ji}}.$$

Example 2.2.7 (Unitary groups). Recall that a matrix $X \in \mathrm{M}_n(\mathbb{C})$ is *unitary* if $X^\dagger X = I$. Equivalently, X is unitary if its rows (or columns) are orthogonal with respect to the *complex* (or *Hermitian*) inner product

$$\langle v, w \rangle = \sum_{i=1}^{n} \overline{v_i} w_i$$

on $\mathbb{C}^n$, or if X preserves this inner product in the sense that $\langle Xv, Xw \rangle = \langle v, w \rangle$ for all $v, w \in \mathbb{C}^n$. We write U_n for the set of unitary matrices in $\mathrm{M}_n(\mathbb{C})$. If X is unitary then (as $\det X^\dagger = \overline{\det X}$) we have $|\det X|^2 = 1$ so that X is in particular invertible, with its inverse necessarily being $X^\dagger$, which is again unitary. Moreover, U_n is closed under matrix multiplication, since for $X, Y \in \mathrm{U}_n$ we have

$$(XY)^\dagger (XY) = Y^\dagger X^\dagger X Y = Y^\dagger Y = I.$$

Thus U_n is a subgroup of $\mathrm{GL}_n(\mathbb{C})$. Moreover, it is a closed subgroup since the equation $X^\dagger X = I$ is a set of polynomial equations in the entries of X and their complex conjugates. We further have the subgroup $\mathrm{SU}_n \subseteq \mathrm{U}_n$ consisting of unitary matrices with determinant 1; this is again closed in $\mathrm{GL}_n(\mathbb{C})$, since it is the intersection of the closed subsets U_n and $\mathrm{SL}_n(\mathbb{C})$.

Exercise 2.5. Show that the group $(\mathbb{R}, +)$ of the real numbers under *addition* is a matrix group. (Hint: Consider the 2×2-matrices of the form $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ for $\lambda \in \mathbb{R}$.)

Example 2.2.8. For $p, q \in \mathbb{N}$ with $p + q = n$, consider the bilinear form $\langle -, - \rangle_{p,q}$ on $\mathbb{R}^n$ given by

$$\langle v, w \rangle_{p,q} = \sum_{i=1}^{p} v_i w_i - \sum_{j=1}^{q} v_{p+j} w_{p+j}.$$

The group $\mathrm{O}_{p,q}(\mathbb{R})$ consists of matrices X that preserve this bilinear form in the sense that we have

$$\langle Xv, Ww \rangle_{p,q} = \langle v, w \rangle_{p,q}.$$

If we write

$$J_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q, \end{pmatrix}$$

then $O_{p,q}(\mathbb{R})$ equivalently consists of those matrices X such that $X^T J_{p,q} X = J_{p,q}$ or (since $J_{p,q}^{-1} = J_{p,q}$)

$$X^{-1} = J_{p,q} X^T J_{p,q}.$$

Its subgroup $SO_{p,q}(\mathbb{R})$ consists of those matrices in $O_{p,q}(\mathbb{R})$ that have determinant 1. Note also that $O_{p,q}(\mathbb{R})$ is isomorphic to $O_{q,p}(\mathbb{R})$ since we have $X^T J_{p,q} X = J_{p,q}$ if and only if $X^T(-J_{p,q})X = -J_{p,q}$ and $-J_{p,q}$ corresponds to $J_{q,p}$ under a reordering of coordinates.

Remark 2.2.9. The group $O_{3,1}(\mathbb{R})$ (or $O_{1,3}(\mathbb{R})$) is relevant in physics: together with translations its elements generate the *Lorentz transformations*, the symmetries of special relativity.

Remark 2.2.10. Given any non-degenerate symmetric bilinear form $\beta(-,-)$ on $\mathbb{R}^n$ we can similarly consider the group of matrices $X \in M_n(\mathbb{R})$ such that $\beta(Xv, Xw) = \beta(v, w)$ for all $v, w \in \mathbb{R}^n$. However, the classification of non-degenerate symmetric bilinear forms on $\mathbb{R}^n$ implies that β corresponds to $\langle -,- \rangle_{p,q}$ for some uniquely determined p, q ($p + q = n$) under a suitable change of basis, and this group is hence isomorphic to $O_{p,q}(\mathbb{R})$.

Exercise 2.6. Show that if we analogously define $O_{p,q}(\mathbb{C})$, then we have an isomorphism $O_{p,q}(\mathbb{C}) \cong O_n(\mathbb{C})$ for all p, q.

Example 2.2.11. A *symplectic form* on a vector space V is a bilinear form $\omega \colon V \times V \to V$ that is skew-symmetric ($\omega(v, v') = -\omega(v', v)$) and non-degenerate (if $\omega(v, w) = 0$ for all w then $v = 0$). The standard symplectic form on $\mathbb{R}^{2n}$ is defined by

$$\omega(v, w) = \sum_{i=1}^{n} (v_i w_{n+i} - v_{n+i} w_i).$$

(It can be shown that any symplectic form is equivalent to the standard one under an appropriate choice of basis.) The *symplectic group* $Sp_n(\mathbb{R}) \subseteq M_{2n}(\mathbb{R})$ consists of matrices X that preserve the standard symplectic form, i.e.

$$\omega(Xv, Xw) = \omega(v, w)$$

for all $v, w \in V$. If we define

$$\Omega := \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

then $Sp_n(\mathbb{R})$ equivalently consists of matrices X such that

$$X^T \Omega X = \Omega,$$

or equivalently (since $\Omega^{-1} = -\Omega$)

$$X^{-1} = -\Omega X^T \Omega.$$

It can be shown that we have $\det X = 1$ for any $X \in \mathrm{Sp}_n(\mathbb{R})$. Similarly, we define $\mathrm{Sp}_n(\mathbb{C})$ to be the group of matrices that preserve the standard symplectic form on $\mathbb{C}^{2n}$.

Example 2.2.12. The *compact symplectic group* Sp_n is defined as the intersection

$$\mathrm{USp}_n := \mathrm{Sp}_n(\mathbb{C}) \cap \mathrm{U}_{2n}.$$

This group thus consists of matrices that preserve *both* the standard symplectic form and the Hermitian inner product on $\mathbb{C}^{2n}$. (The compact symplectic groups can also be described as the "unitary groups of the quaternions", as we will see in the next section.)

Let's also look at a couple of examples of homomorphisms between matrix groups:

Example 2.2.13. The determinant satisfies $\det(AB) = \det(A)\det(B)$ for $A, B \in \mathrm{M}_n(\mathbb{R})$ and so defines a homomorphism

$$\det\colon \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times,$$

where $\mathbb{R}^\times$ is the group of non-zero real numbers under multiplication (i.e. $\mathrm{GL}_1(\mathbb{R})$); it is continuous since the determinant of a matrix is given by a polynomial in its entries. Similarly, we have a continuous homomorphism

$$\det\colon \mathrm{GL}_n(\mathbb{C}) \to \mathbb{C}^\times$$

where $\mathbb{C}^\times = \mathrm{GL}_1(\mathbb{C})$ is the group of non-zero complex numbers.

Exercise 2.7. Show that the function that sends $\theta \in \mathbb{R}$ to the matrix $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ (which encodes rotation by an angle θ) defines a continuous homomorphism $(\mathbb{R}, +) \to \mathrm{SO}_2(\mathbb{R})$.

Example 2.2.14 (Products)**.** We can embed the product $\mathrm{M}_n(\mathbb{R}) \times \mathrm{M}_m(\mathbb{R})$ in $\mathrm{M}_{n+m}(\mathbb{R})$ by taking a pair (A, B) to the block matrix

$$\mu(A, B) := \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

The map μ is then $\mathbb{R}$-linear and compatible with matrix multiplication. Moreover, it identifies $\mathrm{M}_n(\mathbb{R}) \times \mathrm{M}_m(\mathbb{R})$ with a *closed* subset of $\mathrm{M}_{n+m}(\mathbb{R})$, since the image consists precisely of those matrices where certain entries are 0. A matrix in the image of μ is invertible if and only if it is the image of a pair of invertible matrices (for instance since we can compute $\det \mu(A, B) = (\det A)(\det B)$). We have therefore identified the product $\mathrm{GL}_n(\mathbb{R}) \times \mathrm{GL}_m(\mathbb{R})$ (with the product topology) as a matrix group. Consequently, if G and H are matrix groups, then so is their product $G \times H$.

2.3 Quaternions

Let $\mathbb{H}$ be a 4-dimensional $\mathbb{R}$-vector space with basis $1, i, j, k$. The *quaternions* are obtained by equipping $\mathbb{H}$ with an $\mathbb{R}$-bilinear multiplication generated by the relations

$$i^2 = j^2 = k^2 = ijk = -1. \tag{2.1}$$

It is easy (but tedious) to check by hand that this multiplication is associative (that is, we have $a(bc) = (ab)c$ for $a, b, c \in \mathbb{H}$), but it is not commutative, since we get

$$\begin{aligned} ij &= -ijk^2 = k, \\ jk &= -i^2 jk = i, \\ ji &= -jij^2 = -jkj = -ij = -k. \end{aligned}$$

The quaternions were first introduced by William Rowan Hamilton in 1843, who used them to study geometry in 3 dimensions. Here we will see that several of the matrix groups we've looked at earlier have interesting alternative descriptions in terms of the quaternions.

The starting point will be to view the quaternions as certain 2×2 complex matrices (just as we earlier viewed the complex numbers as certain 2×2 *real* matrices).

We can give a concrete description of the quaternions as an algebra of matrices (similar to the description of $\mathbb{C}$ we gave in Exercise 2.2):

Exercise 2.8. Show that if we identify the quaternion $a + bi + cj + dk \in \mathbb{H}$ with the matrix

$$\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \tag{2.2}$$

where $z = a + bi$, $w = c + di$, then quaternion multiplication corresponds to matrix multiplication in $\mathrm{M}_2(\mathbb{C})$.

Note that under this identification we have the following correspondence of bases (over $\mathbb{R}$):

$$1 \longleftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad i \longleftrightarrow \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \qquad j \longleftrightarrow \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad k \longleftrightarrow \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{2.3}$$

Here is one characterization of the matrices that correspond to quaternions:

Proposition 2.3.1. *$M \in \mathrm{M}_2(\mathbb{C})$ is of the form (2.2) if and only if the following conditions hold:*

(1) $M^\dagger M = \lambda I$ *for some* $\lambda \in \mathbb{R}$,

(2) $\det(M) \in \mathbb{R}_{\geq 0}$,

Proof. We first check that a matrix of the form (2.2) satisfies the given conditions:

$$\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}^{\dagger} \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} = \begin{pmatrix} \overline{z} & -w \\ \overline{w} & z \end{pmatrix} \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} = \begin{pmatrix} |z|^2 + |w|^2 & 0 \\ 0 & |z|^2 + |w|^2, \end{pmatrix} = (|z|^2+|w|^2)I,$$

$$\det \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} = |z|^2 + |w|^2 \geq 0.$$

Next, suppose $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in M_2(\mathbb{C})$ satisfies the conditions. Then we have

$$M^{\dagger}M = \begin{pmatrix} \overline{\alpha} & \overline{\gamma} \\ \overline{\beta} & \overline{\delta} \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \overline{\alpha}\alpha + \overline{\gamma}\gamma & \overline{\alpha}\beta + \overline{\gamma}\delta \\ \overline{\beta}\alpha + \overline{\delta}\gamma & \overline{\beta}\beta + \overline{\delta}\delta \end{pmatrix} = \lambda I,$$

so that

$$|\alpha|^2 + |\gamma|^2 = |\beta|^2 + |\delta|^2, \tag{2.4}$$

$$\overline{\alpha}\beta + \overline{\gamma}\delta = 0, \tag{2.5}$$

and in addition $\det M$ is real and

$$\det M = \alpha\delta - \beta\gamma \geq 0. \tag{2.6}$$

If $\beta \neq 0$ then (2.5) gives

$$\overline{\alpha} = -\frac{\overline{\gamma}\delta}{\beta}, \qquad \alpha = -\frac{\gamma\overline{\delta}}{\overline{\beta}}, \tag{2.7}$$

and substituting both conjugates in (2.4) gives

$$\frac{|\gamma|^2|\delta|^2}{|\beta|^2} + |\gamma|^2 = \frac{|\gamma|^2}{|\beta|^2}(|\delta|^2 + |\beta|^2) = |\beta|^2 + |\delta|^2,$$

so that $|\gamma| = |\beta|$. Putting this back in (2.4) we also get $|\alpha| = |\delta|$.

Now substituting (2.7) in (2.6) we get

$$\alpha\delta - \beta\gamma = -\frac{\gamma\overline{\delta}\delta}{\overline{\beta}} - \beta\gamma = -\frac{\gamma}{\overline{\beta}}(|\delta|^2 + |\beta|^2) \geq 0,$$

which means that we must have $\gamma = -\lambda\overline{\beta}$ for some $\lambda \in \mathbb{R}_{\geq 0}$. But we know $|\gamma| = |\beta|$, which forces

$$\gamma = -\overline{\beta}.$$

Putting this back in (2.5) gives

$$\overline{\alpha}\beta - \beta\delta = \beta(\overline{\alpha} - \delta) = 0$$

and so $\overline{\alpha} = \delta$. Thus M is of the form (2.2).

On the other hand, if $\beta = 0$ then (2.5) gives $\gamma\overline{\delta} = 0$ so that either $\gamma = 0$ or $\delta = 0$. If $\delta = 0$ then (2.4) forces $\alpha = \gamma = 0$ too, so that M is the zero matrix, while if $\gamma = 0$ then (2.4) gives $|\alpha| = |\delta|$. We can then write $\alpha = \lambda u$ and $\delta = \lambda v$ for some $\lambda \geq 0$ with $|u| = |v| = 1$, so that (2.6) implies

$$\alpha\delta = \lambda^2 uv \geq 0.$$

Then we must have $uv = 1$ and so $u = \overline{v}$ and $\overline{\alpha} = \delta$, which means M is again of the required form (2.2). □

If $q = a + bi + cj + dk$ $(a, b, c, d \in \mathbb{R})$ is a quaternion, its *conjugate* is defined by

$$\overline{q} = a - bi - cj - dk.$$

Exercise 2.9. Check that for quaternions q, q' we have $\overline{qq'} = \overline{q'} \cdot \overline{q}$, and that

$$q\overline{q} = \overline{q}q = |q|^2 = a^2 + b^2 + c^2 + d^2, \qquad |qq'| = |q||q'|.$$

Conclude that the subset $\{q \in \mathbb{H} : |q| = 1\}$ is a group under quaternion multiplication.

Corollary 2.3.2. *The group* SU_2 *is isomorphic to the group of unit-length quaternions under quaternion multiplication.*

Proof. If a quaternion q corresponds to a matrix M under the isomorphism of Proposition 2.3.1, then our calculations above show that

$$M^\dagger M = |q|^2 I, \qquad \det M = |q|^2.$$

The group of unit-length quaternions hence corresponds to the group of matrices $M \in \mathrm{M}_2(\mathbb{C})$ such that

$$M^\dagger M = I, \qquad \det M = 1,$$

which is precisely the group SU_2 of unitary matrices with determinant 1. □

Corollary 2.3.3. *As a topological space,* SU_2 *is homeomorphic to the 3-dimensional sphere*

$$S^3 := \{(a, b, c, d) \in \mathbb{R}^4 : a^2 + b^2 + c^2 + d^2 = 1\}.$$

We can use this quaternionic description of SU_2 to define an important homomorphism $\mathrm{SU}_2 \to \mathrm{SO}_3(\mathbb{R})$, which explains what the 4-dimensional quaternions have to do with rotations in 3 dimensions:

Construction 2.3.4. Let V be the 3-dimensional subspace (over $\mathbb{R}$) of $\mathbb{H}$ spanned by the basis vectors i, j, k; then a quaternion x lies in V if and only if we have $\overline{x} = -x$. Now suppose q is a unit-length quaternion. Then we have $q^{-1} = \overline{q}$, and it follows that the map $r_q \colon x \mapsto qxq^{-1} = qx\overline{q}$ takes V to itself, since we have

$$\overline{r_q(x)} = \overline{qxq^{-1}} = \overline{q^{-1}}\,\overline{x}\,\overline{q} = q(-x)q^{-1} = -r_q(x).$$

r_q is also clearly an $\mathbb{R}$-linear map, and we have $r_q \circ r_{q'} = r_{qq'}$. In particular r_q is invertible with inverse $r_{q^{-1}}$, so that we have defined a homomorphism $\mathrm{SU}_2 \to \mathrm{GL}(V)$, where $\mathrm{GL}(V) \cong \mathrm{GL}_3(\mathbb{R})$ is the group of invertible $\mathbb{R}$-linear endomorphisms of V, which is moreover continuous (since the components of the matrix for r_q are the components of qxq^{-1} when $x = i, j, k$, and these certainly depend continuously on the components of q). Moreover, r_q preserves lengths, since we have

$$|r_q(x)| = |qxq^{-1}| = |q||x||q|^{-1} = |x|.$$

Thus we can view r as a homomorphism $\mathrm{SU}_2 \to O_3(\mathbb{R})$. In fact, we always have $\det r_q = 1$ — one way to see this is to observe that SU_3 is connected (since topologically it is S^3), and so the continuous function $\det \circ r\colon \mathrm{SU}_2 \to \{\pm 1\}$ must be constant with value 1 (as that is the value at I). Thus r gives a homomorphism $\mathrm{SU}_2 \to \mathrm{SO}_3(\mathbb{R})$. Note, however, that this is not injective: we clearly have $r_q = r_{-q}$.

Exercise 2.10. Show that if $r_q = \mathrm{id}$ then $q = \pm 1$ by computing $r_q(x)$ for $x = i, j, k$. Conclude that r is a 2-to-1 homomorphism.

Remark 2.3.5. One can show that rotation in $\mathbb{R}^3$ by an angle θ about an axis represented by the unit vector $v = (x, y, z)$ corresponds to $r_{\pm q}$ where

$$q = \cos(\theta/2) + \sin(\theta/2)(xi + yj + zk).$$

Note that this representation is much more straightforward than writing down the corresponding 3×3 rotation matrix explicitly, and gives a way to find the axis and angle for a composite of rotations by multiplying quaternions.

Exercise 2.11. Verify this in the case of rotations about the coordinate axes, i.e. show that if $q = \cos(\theta/2) + \sin(\theta/2)i$ then r_q is a rotation by θ about the x-axis (and the same for j and k).

Since it can be shown that any rotation in $\mathbb{R}^3$ can be expressed as a composition of rotations about the coordinate axes, it follows from this exercise that r is surjective. We have therefore proved the following:

Proposition 2.3.6. *r is a 2-to-1 surjective continuous homomorphism* $\mathrm{SU}_2 \to \mathrm{SO}_3(\mathbb{R})$.

Corollary 2.3.7. $\mathrm{SO}_3(\mathbb{R})$ *is homeomorphic to the 3-dimensional real projective space* $\mathbb{RP}^3$ *(i.e. the space of lines through the origin in* $\mathbb{R}^4$*).*

Proof. The kernel of r is $\{\pm 1\}$, so that r defines a continuous bijection

$$\mathrm{SU}_2/\{\pm I\} \to \mathrm{SO}_3(\mathbb{R}).$$

Both sides are compact Hausdorff spaces, so this is a homeomorphism. If we identify SU_2 with S^3 then $\mathrm{SU}_2/\{\pm I\}$ is the quotient of S^3 where we identify antipodal points on the sphere, which is precisely $\mathbb{RP}^3$. (This is because every line through the origin in $\mathbb{R}^4$ intersects the unit sphere S^3 in exactly two antipodal points.) □

Construction 2.3.8. Let $\mathrm{M}_n(\mathbb{H})$ denote the $n \times n$ matrices with quaternion entries. We can define addition and matrix multiplication by the usual formulas, as well as scalar multiplication by $\mathbb{H}$ (but note that since $\mathbb{H}$ is not commutative, we have to be careful about this!). Now, just as we embedded $\mathrm{M}_n(\mathbb{C})$ in $\mathrm{M}_{2n}(\mathbb{R})$ earlier, we can use the matrix description Proposition 2.3.1 to embed $\mathrm{M}_n(\mathbb{H})$ in $\mathrm{M}_{2n}(\mathbb{C})$ by expanding each quaternion entry into a 2×2 block. This gives an embedding

$$\eta_n \colon \mathrm{M}_n(\mathbb{H}) \hookrightarrow \mathrm{M}_{2n}(\mathbb{C}),$$

which is $\mathbb{R}$-linear and compatible with matrix multiplication.

Exercise 2.12. Show that $\eta_1(\overline{q}) = \eta_1(q)^\dagger$ for $q \in \mathbb{H}$, and that if we define the conjugate transpose in $\mathrm{M}_n(\mathbb{H})$ in the obvious way, then $\eta_n(M^\dagger) = \eta_n(M)^\dagger$ for all n.

Exercise 2.13.

(i) Let $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Show that a matrix $M \in M_2(\mathbb{C})$ is in the image of η_1 if and only if we have

$$JM = \overline{M}J.$$

(ii) Let $J_n \in \mathrm{M}_{2n}(\mathbb{C})$ be the matrix

$$\begin{pmatrix} J & & 0 \\ & \ddots & \\ 0 & & J \end{pmatrix}.$$

Show that a matrix $M \in \mathrm{M}_{2n}(\mathbb{C})$ is in the image of η_n if and only if we have

$$J_n M = \overline{M} J_n.$$

Proposition 2.3.9. *The compact symplectic group* $\mathrm{USp}_n \subseteq \mathrm{GL}_{2n}(\mathbb{C})$ *is the image under* η_n *of the group* $\mathrm{U}_n(\mathbb{H})$ *of unitary* $n \times n$ *quaternion matrices, i.e.* $M \in \mathrm{M}_n(\mathbb{H})$ *such that*

$$M^\dagger M = I.$$

Proof. Under the embedding η_n, the condition $M^\dagger M = I$ translates into $\eta_n(M)^\dagger \eta_n(M) = I$ by Exercise 2.12. From Exercise 2.13 we therefore see that a matrix $X \in \mathrm{M}_{2n}(\mathbb{C})$ is the image under η_n of a unitary quaternion matrix if and only if we have

$$X^\dagger X = I, \qquad J_n X = \overline{X} J_n.$$

The first condition here says precisely that X is unitary, so that we have $X^{-1} = X^\dagger$, and hence $\overline{X} = (X^T)^{-1}$. Given that X is unitary we can therefore rephrase the second equation as

$$X^T J_n X = J_n,$$

which says that X preserves the symplectic form

$$\omega(x, y) = \sum_{i=1}^{n} (x_{2i-1} y_{2i} - x_{2i} y_{2i-1}).$$

This corresponds to the standard sympletic form we considered above under a reordering of the standard basis of $\mathbb{C}^{2n}$. □

2.4 Topological properties

In this section we will look at three important topological properties that matrix groups may (or may not) have, namely compactness, path-connectedness, and simple connectedness.

Definition 2.4.1. Recall that a topological space is *compact* if every open cover has a finite subcover. Since $M_n(\mathbb{R})$ is homeomorphic to $\mathbb{R}^{n^2}$, the Heine–Borel theorem implies that a subset of $M_n(\mathbb{R})$ is compact if and only if it is closed and bounded (with respect to the usual metric on $\mathbb{R}^{n^2}$.) Thus a matrix group G is compact if and only if it is closed in $M_n(\mathbb{R})$ and there exists a constant C such that if $A \in G$ then we have $|A_{ij}| \leq C$ for all the entries of A.

Examples 2.4.2.

(i) The group $GL_n(\mathbb{R})$ is not compact — indeed, it is a (non-empty and proper) *open* subset of $M_n(\mathbb{R})$, and so cannot be closed.

(ii) The group $SL_n(\mathbb{R})$ is a closed subset of $M_n(\mathbb{R})$ (being the preimage of the closed set $\{1\} \subseteq \mathbb{R}$ under the determinant map). However, it is not compact for $n > 1$, since the diagonal matrix with entries $(m, 1/m, 1, \ldots, 1)$ is in $SL_n(\mathbb{R})$ for $m = 1, 2, \ldots$ and so it is not bounded.

(iii) The groups $(S)O_n(\mathbb{R})$, $(S)O_n(\mathbb{C})$, $(S)U_n$ and USp_n are all compact. For example, if $A \in O_n(\mathbb{R})$ then the columns of A are required to be unit vectors, and so we necessarily have $|A_{ij}| \leq 1$ for all i, j, which shows that $O_n(\mathbb{R})$ is bounded; our argument that it is a matrix group actually showed that it is closed in $M_n(\mathbb{R})$ and so it is compact.

Exercise 2.14. Check that U_n and USp_n are compact.

Definition 2.4.3. A topological space is *path-connected* if any pair of points can be connected by a continuous path.

Remark 2.4.4. We will later show that any matrix group G is a topological manifold. This implies that it is in particular locally path-connected, and from this it follows that its connected components and path components coincide. In particular, a matrix group is path-connected if and only if it is *connected* in the sense that it cannot be written as the disjoint union of two open subsets. We will therefore often use this shorter term when talking about matrix groups.

Proposition 2.4.5. *Let G be a topological group, and let G_0 be the path component of the identity. Then G_0 is a normal subgroup of G. If G is a matrix group, then so is G_0.*

Proof. We first check that G_0 is a subgroup. Suppose X and Y lie in G_0 so that there exist continuous paths $\xi, \eta\colon [0,1] \to G$ with $\xi(0) = \eta(0) = I$ and $\xi(1) = X, \xi(1) = Y$. Since multiplication is continuous, $t \mapsto \xi(t)\eta(t)$ is then a continuous path from I to XY, so that $XY \in G_0$. Similarly, as inversion is continuous we get that $\xi(t)^{-1}$ is a continuous path from I to X^{-1}, so $X^{-1} \in G_0$. To see this subgroup is normal, observe that furthermore $t \mapsto Y\xi(t)Y^{-1}$ gives a continuous path from I to YXY^{-1} for any $Y \in G$. By Remark 2.4.4, if G is a matrix group then the path-component G_0 is also a connected component. Then it is closed in G, which implies that it is again a matrix group. □

Example 2.4.6. $\mathrm{GL}_n(\mathbb{C})$ and $\mathrm{SL}_n(\mathbb{C})$ are connected: For $A, B \in \mathrm{GL}_n(\mathbb{C})$ consider the function $z \mapsto \det(zB + (1-z)A)$ from $\mathbb{C}$ to $\mathbb{C}$. This is a non-zero polynomial of degree n in z, and so has at most n zeros in $\mathbb{C}$. We can then choose a continuous path $\zeta(t)\colon [0,1] \to \mathbb{C}$ with $\zeta(0) = 0, \zeta(1) = 1$, that avoids these n points. Then $t \mapsto P(t) := \zeta(t)B + (1 - \zeta(t))A$ is a path from A to B in $\mathrm{GL}_n(\mathbb{C})$. If A and B lie in $\mathrm{SL}_n(\mathbb{C})$ we can instead consider

$$t \mapsto \left(\det P(t)\right)^{-1/n} P(n)$$

to get a path in $\mathrm{SL}_n(\mathbb{C})$ (provided we choose ζ to not wind around 0 so that we can define a continous branch of the nth root).

Example 2.4.7. $\mathrm{GL}_n(\mathbb{R})$ is *not* connected: since the determinant is continuous we can't have a path that connects a matrix with determinant > 0 to one whose determinant is < 0. For the same reason, $\mathrm{O}_n(\mathbb{R})$ is not connected.

Remark 2.4.8. It can be shown that $\mathrm{SL}_n(\mathbb{R})$ is connected, and that the connected component of the identity in $\mathrm{GL}_n(\mathbb{R})$ is the group $\mathrm{GL}_n^+(\mathbb{R})$ of matrices with determinant > 0.

Exercise 2.15. Assuming $\mathrm{SL}_n(\mathbb{R})$ is connected, show that so is $\mathrm{GL}_n^+(\mathbb{R})$.

Example 2.4.9. U_n and SU_n are connected: It can be shown that any unitary matrix has a (complex-)orthonormal basis of eigenvectors (with all eigenvalues necessarily having norm 1). If $X \in \mathrm{U}_n$ we can therefore write

$$X = YDY^{-1}$$

where Y is unitary and D is diagonal with diagonal entries $(\lambda_1, \ldots, \lambda_n)$. If $\lambda_j = e^{i\theta_j}$ then we can set $D(t)$ to be diagonal with entries $(e^{it\theta_1}, \ldots, e^{it\theta_n})$, so that

$$t \mapsto YD(t)Y^{-1}$$

is a continuous path in $U(n)$ from I to X. For SU_n we use $\frac{1}{\det D(t)} YD(t)Y^{-1}$ instead.

Exercise 2.16. Show that $\mathrm{SO}_n(\mathbb{R})$ is connected by induction on n; see [2, Chapter 2, Exercise 13] or [4, Section 3.2] for more details on the argument.

Definition 2.4.10. A topological space X is *simply connected* if it is path-connected and in addition every loop is homotopic to a constant one, that is given $\gamma\colon S^1 \to X$ there exists $h\colon S^1 \times [0,1] \to X$ such that $h(-,0) = \gamma$ and $h(-,1)$ is constant.

Example 2.4.11. By Corollary 2.3.2 the group SU_2 is homeomorphic ot the 3-sphere S^3 as a topological space. It is therefore simply connected. It can be shown that SU_n is simply connected for all $n \geq 2$.

Example 2.4.12. By Corollary 2.3.7 the group $\mathrm{SO}_3(\mathbb{R})$ is homeomorphic to $\mathbb{RP}^3$, which is *not* simply connected. (In fact, this means $\pi_1\mathrm{SO}_3(\mathbb{R}) = \mathbb{Z}/2$, and this is true for $\mathrm{SO}_n(\mathbb{R})$ for all $n \geq 3$.)

2.5 Manifolds and Lie groups

We claimed earlier that Lie groups are supposed to be "smooth" groups, but when we introduced matrix groups we only considered them as topological groups. In this section we try to explain this discrepancy by looking briefly at what precisely we mean by a smooth structure, and state some results (some very hard and some easy enough that we will prove a version of them) that justify our definitions above.

Definition 2.5.1. A topological space X is a *topological manifold* if it has an open cover $\{U_i\}$ such that each U_i is homeomorphic to an open subset of $\mathbb{R}^n$ for some n. A *smooth manifold* is a topological manifold X together with such a cover and homeomorphisms $\phi_i\colon U_i \xrightarrow{\sim} V_i \subseteq \mathbb{R}^n$, for which the transition functions

$$\phi_i(U_i \cap U_j) \xrightarrow{\phi_i^{-1}} U_i \cap U_j \xrightarrow{\phi_j} \phi_j(U_i \cap U_j)$$

are smooth, that is infinitely differentiable, functions between open subsets of $\mathbb{R}^n$. (More precisely, we should say that X is equipped with an equivalence class of this data, or with a maximal choice thereof, but we will not go into the details here.)

If M and N are smooth manifolds, then a map $M \to N$ is *smooth* if it is locally given by smooth maps between subsets of $\mathbb{R}^n$'s. We won't spell out the details, but assuming this makes sense we can make the following definition:

Definition 2.5.2. A *Lie group* is a topological group G together with a smooth manifold structure on G such that the multiplication $G \times G \to G$ and inverse $G \to G$ are smooth maps. The natural maps between Lie groups are then *smooth homomorphisms*, i.e. group homomorphisms that are also smooth maps.

Remark 2.5.3. While being a topological manifold is a *property* of a topological space, a smooth manifold is a topological space with extra *structure*. For example, perhaps one of the most surprising theorems in mathematics is that while $\mathbb{R}^n$ has a unique smooth structure for all $n \neq 4$, the topological manifold $\mathbb{R}^4$ can be made into a smooth manifold in uncountably many distinct ways.

For topological groups, however, it turns out that being a Lie group is actually a property:

Fact 2.5.4 (Hilbert's Fifth Problem)**.** *Suppose G is a topological group whose underlying space is a topological manifold. Then G can be promoted to a Lie group in a unique way (in fact, G can be uniquely promoted to a group in real analytic manifolds).*

This is a theorem of von Neumann (1933) if G is compact, and in general of Gleason, Montgomery, and Zippin (1955). We will later prove the (far easier!) claim that any matrix group is a Lie group; this was originally proved by von Neumann (1929), and can be regarded as a special case of the following result of Cartan (1930):

Fact 2.5.5 (Closed Subgroup Theorem)**.** *Any closed subgroup of a Lie group is again a Lie group, with the smooth structure inherited from the larger group.*

A consequence of this theorem is:

Fact 2.5.6. *Any continuous homomorphism between Lie groups is smooth.*

While we will not prove this in this course, we will see later that a continuous homomorphism between *matrix* groups is smooth.

We end this section with an example of a Lie group that is *not* a matrix group:

Example 2.5.7. Let $\mathbb{T}$ be the subspace of $\mathbb{C}$ consisting of complex numbers z with $|z| = 1$. Consider the manifold $\mathbb{R} \times \mathbb{R} \times \mathbb{T}$ with the group structure given by

$$(x, y, u) \cdot (x', y', u') = (x + x', y + y', e^{ixy'} uu').$$

It is easy to check that this is a Lie group, but it can be shown (see [2, Section 4.8]) that it is not a matrix group.

Remark 2.5.8. If G is a Lie group, then (as for any reasonable topological space) there is a universal simply connected space $\widetilde{G}$ with a map to G, namely the *universal cover* $\widetilde{G} \to G$. It can be shown that $\widetilde{G}$ is again a Lie group, with the canonical map $\widetilde{G} \to G$ a continuous (and so smooth) homomorphism. However, it is possible that G is a matrix group but $\widetilde{G}$ is not. For example, $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ is *not* a matrix group.

Chapter 3

The matrix exponential and Lie algebras

3.1 The matrix exponential

If $X \in \mathrm{M}_n(\mathbb{C})$ is a square matrix, we want to define its *exponential* by the usual power series

$$\exp(X) = e^X := \sum_{k=0}^{\infty} \frac{1}{k!} X^k. \tag{3.1}$$

We need to check that this is well-defined, in the sense that each matrix entry in the partial sums converges in $\mathbb{C}$. For this it is convenient to introduce a norm on $\mathrm{M}_n(\mathbb{C})$, which will correspond to the usual norm on $\mathbb{C}^{n^2}$:

Definition 3.1.1. The *Hilbert–Schmidt* norm of $X \in \mathrm{M}_n(\mathbb{C})$ is

$$\|X\| := \sqrt{\sum_{i,j=1}^{n} |X_{ij}|^2} = \sqrt{\mathrm{tr}(X^\dagger X)}.$$

From the triangle inequality in $\mathbb{C}^{n^2}$ we get

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

Exercise 3.1. Use the Cauchy–Schwarz inequality to show that we have

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

Exercise 3.2. Show that for a sequence of matrices X_m in $\mathrm{M}_n(\mathbb{C})$ we have that $\|X - X_m\|$ converges to 0 as m goes to ∞ for some matrix X if and only if for all $0 \leq i, j \leq n$ the sequence $(X_m)_{ij}$ converges to X_{ij} in $\mathbb{C}$.

Exercise 3.3. If $u_1, \ldots, u_n$ is an orthonormal basis of $\mathbb{C}^n$, show that for $X \in \mathrm{M}_n(\mathbb{C})$ we have

$$\|X\|^2 = \sum_{i,j=1}^{n} \left|\langle u_i, X u_j \rangle\right|^2$$

(where "orthonormal" and the inner product are in the complex (sesquilinear) sense). Conclude that if λ is an eigenvalue of X then $|\lambda| \leq \|X\|$.

Lemma 3.1.2. *The exponential sequence (3.1) converges for any matrix $X \in \mathrm{M}_n(\mathbb{C})$.*

Proof. We have

$$\left\| \sum_{k=0}^{K} \frac{1}{k!} X^k \right\| \leq \sum_{k=0}^{K} \frac{1}{k!} \|X\|^k \tag{3.2}$$

which converges, so the series converges absolutely. □

Remark 3.1.3. In this section we will work with complex matrices, since working over $\mathbb{C}$ will occasionally be useful later on. However, it is clear from the definition that for $X \in \mathrm{M}_n(\mathbb{R}) \subseteq \mathrm{M}_n(\mathbb{C})$ the matrix e^X also lies in $\mathrm{M}_n(\mathbb{R})$. Moreover, since the embedding $\iota\colon \mathrm{M}_n(\mathbb{C}) \hookrightarrow \mathrm{M}_{2n}(\mathbb{R})$ is continuous and compatible with multiplication, we have

$$\iota e^X = e^{\iota X}.$$

Proposition 3.1.4. *The function $e^{(-)}\colon \mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_n(\mathbb{C})$ is continuous.*

Proof. We can use the bound (3.2) together with the "Weierstraß M-test" to conclude that the exponential series converges uniformly on the set of X with $\|X\| \leq R$ for any $R \geq 0$, which implies continuity on this set, and hence on all of $\mathrm{M}_n(\mathbb{C})$. □

Proposition 3.1.5. *If $XY = YX$, then we have*

$$e^{X+Y} = e^X e^Y = e^Y e^X \tag{3.3}$$

Proof. Since X and Y commute, we have the binomial formula

$$(X+Y)^m = \sum_{k=0}^{m} \binom{m}{k} X^k Y^{m-k},$$

so that same computation as for exponentials in $\mathbb{C}$ goes through:

$$\begin{aligned} e^X e^Y &= \sum_{m=0}^{\infty} \sum_{k=0}^{m} \frac{X^k}{k!} \frac{Y^{m-k}}{(m-k)!} \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} X^k Y^{m-k} \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} (X+Y)^m \\ &= e^{X+Y}. \end{aligned}$$

□

Corollary 3.1.6. *For any matrix X, we have*

$$e^{\lambda X}e^{\mu X} = e^{(\lambda+\mu)X}$$

for scalars $\lambda, \mu \in \mathbb{C}$. In particular, the matrix e^X is invertible with inverse e^{-X}.

Proof. Since λX and μX commute, the first claim follows from Proposition 3.1.5. In particular, we have

$$e^X e^{-X} = e^{-X}e^X = e^{X-X} = e^0 = I,$$

which shows that e^{-X} is inverse to e^X. □

Warning 3.1.7. When the matrices X and Y fail to commute, we generally do *not* have the expression (3.3). Instead, as we will discuss below we have a rather more complicated expression for the product $e^X e^Y$ as an exponential, the Baker–Campbell–Hausdorff formula, when X and Y are sufficiently small.

Lemma 3.1.8. *For any $C \in \mathrm{GL}_n(\mathbb{C})$ and $X \in \mathrm{M}_n(\mathbb{C})$, we have*

$$Ce^X C^{-1} = e^{CXC^{-1}}.$$

Proof. Immediate from the definition, since $(CXC^{-1})^m = CX^mC^{-1}$. □

Remark 3.1.9. We also have $(e^X)^T = e^{X^T}$ and $(e^X)^\dagger = e^{X^\dagger}$.

Exercise 3.4. Suppose $X \in \mathrm{M}_n(\mathbb{C})$ is a diagonalizable matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Show that e^X is diagonalizable with eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_n}$.

As we have seen already, we have to take care when trying to generalize properties of the exponential to matrices. For example, the following is the correct way to extend the formula for the derivative of the function e^{at}:

Proposition 3.1.10. *For any $X \in \mathrm{M}_n(\mathbb{C})$, the function $\mathbb{R} \to \mathrm{M}_n(\mathbb{C})$ given by $t \mapsto e^{tX}$ is smooth, with derivative*

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{tX} = Xe^{tX}.$$

In particular, the derivative at 0 is the matrix X.

Proof. We can differentiate the power series termwise, since for each index the component $(e^{tX})_{ij}$ is a well-behaved power series in t. Hence we have as usual that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}e^{tX} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{n=0}^{\infty}\frac{(tX)^n}{n!}\right)\\ &= \sum_{n=0}^{\infty}\frac{nt^{n-1}X^n}{n!}\\ &= \sum_{n=1}^{\infty}X\frac{(tX)^{n-1}}{(n-1)!}\\ &= \sum_{n=0}^{\infty}X\frac{(tX)^n}{n!}\\ &= Xe^{tX}.\end{aligned}$$

□

Observation 3.1.11. The matrix exponential gives an infinitely differentiable map $\exp\colon \mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_n(\mathbb{C})$: Since the power series defining it converges everywhere, we can differentiate it termwise arbitrarily many times.

Remark 3.1.12. We briefly mention another application of the matrix exponential: consider the system of linear ordinary differential equations

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = A\mathbf{x}(t)$$

for a function $\mathbf{x}\colon \mathbb{R} \to \mathbb{R}^n$ and $A \in \mathrm{M}_n(\mathbb{R})$. For a given initial condition $\mathbf{x}(0) = \mathbf{c}$, this has the unique solution

$$\mathbf{x}(t) = e^{tA}\mathbf{c}.$$

3.2 The Lie algebra of a matrix group

Definition 3.2.1. Let $(G, j\colon G \hookrightarrow \mathrm{GL}_n(\mathbb{R}))$ be an embedded matrix group. We define the *Lie algebra* $\mathfrak{L}(G, j)$ to be the subset of $\mathrm{M}_n(\mathbb{R})$ consisting of matrices X such that e^{tX} is in $j(G)$ for all $t \in \mathbb{R}$.

Variant 3.2.2. If we think of G as a subgroup of $\mathrm{GL}_n(\mathbb{C})$, then it is more natural to think of $\mathfrak{L}(G)$ as a subset of $\mathrm{M}_n(\mathbb{C})$. The compatibility of Remark 3.1.3 means that this does not make a difference to the set of matrices that lie in the Lie algebra.

Remark 3.2.3. Note that the Lie algebra a priori depends on the chosen embedding in a general linear group. (This is a disadvantage of our low-tech approach!) We will see later that the Lie algebras for different embeddings are canonically isomorphic, and so we will often drop the embedding from the notation when this does not cause confusion.

Notation 3.2.4. It is conventional to denote the Lie algebra of a group by the name of the group in lower-case fraktur letters, that is

$$\mathfrak{g} = \mathfrak{L}(G), \quad \mathfrak{h} = \mathfrak{L}(H), \quad \mathfrak{gl}_n(\mathbb{R}) = \mathfrak{L}(\mathrm{GL}_n(\mathbb{R})),$$

etc. Since students tend to dislike fraktur letters at first sight, we will make sure to use them as much as possible.

Remark 3.2.5. We will eventually show that the Lie algebra $\mathfrak{g}$ of G is the tangent space of G at the identity. From Proposition 3.1.10 we know that $t \mapsto e^{tX}$ is a smooth curve in G whose derivative at 0 is X, which at least shows that the elements of the Lie algebra can be thought of as tangent vectors at the identity.

Warning 3.2.6. Physicists like to define the Lie algebra of G to consist of matrices $X \in \mathrm{M}_n(\mathbb{C})$ such that $e^{itX} \in G$ for all t. This means that the Lie algebras we will associate to many groups will differ by a factor of i from the description you might find in a physics book. On the other hand physicists also tend not to distinguish between a group and its Lie algebra at all, so this factor of i might well be the least of your troubles...

Remark 3.2.7. Suppose $G \subseteq \mathrm{GL}_n(\mathbb{R})$ is a matrix group with Lie algebra $\mathfrak{g}$. Then $\mathfrak{g}$ is also the Lie algebra of the identity component $G_0 \subseteq G$: By definition the Lie algebra of G_0 is contained in $\mathfrak{g}$, but conversely if $X \in \mathfrak{g}$ then e^{tX} must lie entirely in G_0 as it is a continuous path through $I = e^{0 \cdot X}$.

We want to show that the Lie algebra of a matrix group is a real vector space. It is easy to see that it is closed under scalar multiplication by $\mathbb{R}$:

Lemma 3.2.8. *If $X \in \mathfrak{g}$ then so is rX for any $r \in \mathbb{R}$.*

Proof. We want to show that $e^{t(rX)} = e^{(tr)X}$ is in G for all $t \in \mathbb{R}$. But by assumption we know that e^{sX} is in G for all $s \in \mathbb{R}$, and so in particular for $s = tr$. □

Warning 3.2.9. Even if we think of G as embedded in $\mathrm{GL}_n(\mathbb{C})$ and the Lie algebra $\mathfrak{g}$ as a subset of $\mathrm{M}_n(\mathbb{C})$, the set $\mathfrak{g}$ is typically *not* closed under multiplication by *complex* scalars. If it is, then G is a *complex* Lie group. This means that it has a canonical structure of a *complex manifold* in the sense that it is locally homeomorphic to open subsets of $\mathbb{C}^n$ with the transition functions being holomorphic, and with respect to this the multiplication and inverse are holomorphic maps; we will not make this more precise here, however.

It takes rather more work to show that $\mathfrak{g}$ is closed under addition. The key input is the following formula for the exponential of a sum:

Theorem 3.2.10 (Lie product formula). *For all $X, Y \in \mathrm{M}_n(\mathbb{C})$, we have*

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

To prove this, we need to introduce the *logarithm* of a matrix, which we postpone until Section 3.5. Here we will instead derive the consequence for Lie algebras we are interested in:

Corollary 3.2.11. *If* $X, Y \in \mathfrak{g}$*, so is* $X + Y$*.*

Proof. We must show that $e^{t(X+Y)}$ lies in G for all $t \in \mathbb{R}$. By Theorem 3.2.10, we have

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

Here $\left(e^{\frac{t}{m}X} e^{\frac{t}{m}Y}\right)^m$ is in G for all m since X and Y lie in $\mathfrak{g}$ and G is closed under multiplication. Since the sequence converges to $e^{t(X+Y)}$ which lies in $\mathrm{GL}_n(\mathbb{R})$, and G is a closed subspace of this, the limit must also lie in G. □

We have seen that the Lie algebra $\mathfrak{g}$ of $G \subseteq \mathrm{M}_n(\mathbb{R})$ is a real vector subspace of $\mathrm{M}_n(\mathbb{R})$. It is generally *not* closed under matrix multiplication, but we will now see that it *is* closed under commutators, which gives the "Lie bracket" — this is the operation that makes $\mathfrak{g}$ deserve to be called an "algebra".

Notation 3.2.12. We write

$$[X, Y] := XY - YX.$$

Proposition 3.2.13. *If* $X, Y \in \mathfrak{g}$*, then* $[X, Y]$ *is also in* $\mathfrak{g}$*.*

We need a few observations first:

Exercise 3.5. Prove the following product rule for matrix multiplication: given differentiable functions $A, B \colon \mathbb{R} \to \mathrm{M}_n(\mathbb{R})$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(A(t)B(t)) = \left(\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right)B(t) + A(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}B(t)\right).$$

Lemma 3.2.14. *If* $X \in \mathfrak{g}$*, then we have* $AXA^{-1} \in \mathfrak{g}$ *for any* $A \in G$*.*

Proof. By Lemma 3.1.8 we have $e^{tAXA^{-1}} = Ae^{tX}A^{-1}$ which lies in G as $X \in \mathfrak{g}$. □

Proof of Proposition 3.2.13. Using Exercise 3.5 and Proposition 3.1.10, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{tX}Ye^{-tX}) = Xe^{tX}Ye^{-tX} + e^{tX}Y(-X)e^{-tX}.$$

At $t = 0$ the right-hand side is $XY - YX$, so we have

$$[X, Y] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}(e^{tX}Ye^{-tX}) = \lim_{t\to 0}\frac{e^{tX}Ye^{-tX} - Y}{t}. \tag{3.4}$$

Here $\frac{1}{t}(e^{tX}Ye^{-tX} - Y)$ lies in $\mathfrak{g}$ for all t, by Corollary 3.2.11, Lemma 3.2.8 and Lemma 3.2.14. Since we know $\mathfrak{g}$ is a vector subspace of $\mathrm{M}_n(\mathbb{C})$ it is topologically a closed subset, and hence the limit $[X, Y]$ must also lie in $\mathfrak{g}$. □

Observation 3.2.15. Let $G \subseteq \mathrm{GL}_n(\mathbb{R})$ be a matrix group with Lie algebra $\mathfrak{g}$. For $A \in \mathrm{GL}_n(\mathbb{R})$, conjugation by A determines a linear map $\mathrm{M}_n(\mathbb{R}) \to \mathrm{M}_n(\mathbb{R})$ given by $X \mapsto AXA^{-1}$, and by Lemma 3.2.14 this restricts to a linear map $\mathfrak{g} \to \mathfrak{g}$ for $A \in G$. Let $\mathrm{GL}(\mathfrak{g})$ denote the group of invertible linear endomorphisms of $\mathfrak{g}$, then it is easy to see that the conjugation maps determine a continuous homomorphism

$$\mathrm{Ad}_G \colon G \to \mathrm{GL}(\mathfrak{g}),$$

known as the *adjoint homomorphism* (or the *adjoint action* of G on its Lie algebra).

3.3 Examples of Lie algebras

Now we will describe the Lie algebras of the matrix groups we introduced above.

Example 3.3.1 (General linear groups)**.** The Lie algebra $\mathfrak{gl}_n(\mathbb{R})$ of $\mathrm{GL}_n(\mathbb{R})$ is all of $\mathrm{M}_n(\mathbb{R})$, since the exponential of a matrix is always invertible. Similarly, $\mathfrak{gl}_n(\mathbb{C})$ is $\mathrm{M}_n(\mathbb{C})$.

Lemma 3.3.2. *For $X \in \mathrm{M}_n(\mathbb{C})$, we have*

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

Proof. Suppose first that X is a diagonalizable matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then by Exercise 3.4 we have

$$\det e^X = \prod_{i=1}^{n} e^{\lambda_i} = e^{\sum_{i=1}^n \lambda_i} = e^{\operatorname{tr} X}.$$

But both $\det e^X$ and $e^{\operatorname{tr} X}$ are continuous functions of X. Since diagonalizable matrices form a dense subset of $\mathrm{M}_n(\mathbb{C})$ the two functions must agree everywhere when they agree on this subset. □

Example 3.3.3 (Special linear groups)**.** The Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ of $\mathrm{SL}_n(\mathbb{R})$ comprises all *traceless* matrices, that is all $X \in \mathrm{M}_n(\mathbb{R})$ with $\operatorname{tr} X = 0$: Since $\det e^{tX} = e^{t \operatorname{tr} X}$ we have that $\det e^{tX} = 0$ for all t if and only if $\operatorname{tr} X = 0$. Similarly, $\mathfrak{sl}_n(\mathbb{C})$ consists of all traceless matrices in $\mathrm{M}_n(\mathbb{C})$. (Note that here we need the condition that $\det e^{tX} = 0$ for all t, since a matrix X with $\operatorname{tr} X = 2\pi i$ would satisfy $\det e^X = 1$.)

Example 3.3.4 (Finite groups)**.** The Lie algebra of any finite group is 0.

Example 3.3.5 (Orthogonal groups)**.** The Lie algebra $\mathfrak{o}_n(\mathbb{R})$ comprises the *skew-symmetric* matrices, that is $X \in \mathrm{M}_n(\mathbb{R})$ such that

$$X^T = -X.$$

To see this, we use Remark 3.1.9 to conclude that X is in $\mathfrak{o}_n(\mathbb{R})$ if and only if

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

for all $t \in \mathbb{R}$. This is certainly true if X is skew-symmetric, but if, conversely, this holds for all t then we also have

$$X^T = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} e^{tX^T} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} e^{-tX} = -X.$$

Since a skew-symmetric matrix is also traceless (as $\operatorname{tr} X = \operatorname{tr} X^T = -\operatorname{tr} X$), we see that $\mathfrak{so}_n(\mathbb{R}) = \mathfrak{o}_n(\mathbb{R})$ (as we must have since $\mathrm{SO}_n(\mathbb{R})$ is the identity component of $\mathrm{O}_n(\mathbb{R})$). Similarly, $\mathfrak{o}_n(\mathbb{C}) = \mathfrak{so}_n(\mathbb{C})$ comprises the skew-symmetric matrices in $\mathrm{M}_n(\mathbb{C})$.

Example 3.3.6 (Unitary groups)**.** By the same argument as in the orthogonal case, the Lie algebra $\mathfrak{u}_n$ of the unitary group U_n consists of the *skew-Hermitian* matrices, meaning those $X \in \mathrm{M}_n(\mathbb{C})$ such that

$$X^\dagger = -X.$$

Note that a skew-Hermitian matrix may not be traceless (since the trace just has to satisfy $\operatorname{tr} X = -\operatorname{tr} X^\dagger = -\overline{\operatorname{tr} X}$), and the Lie algebra $\mathfrak{su}_n$ of SU_n consists of the traceless skew-Hermitian matrices.

Example 3.3.7 (Symplectic groups)**.** The Lie algebra $\mathfrak{sp}_n(\mathbb{R})$ comprises the matrices $X \in \mathrm{M}_{2n}(\mathbb{R})$ such that

$$\Omega X + X^T \Omega = 0.$$

Indeed, for X to lie in $\mathfrak{sp}_n(\mathbb{R})$ we must have

$$e^{-tX} = (e^{tX})^{-1} = -\Omega (e^{tX})^T \Omega = e^{-t\Omega X^T \Omega}$$

since $-\Omega = \Omega^{-1}$. Differentiating, we see that this holds if and only if $-X = -\Omega X^T \Omega$. Similarly, $\mathfrak{sp}_n(\mathbb{C})$ consists of the matrices $X \in \mathrm{M}_{2n}(\mathbb{C})$ such that $\Omega X + X^T \Omega = 0$, while the Lie algebra $\mathfrak{usp}_n$ of the compact symplectic group $\mathrm{USp}_n = \mathrm{Sp}_n(\mathbb{C}) \cap \mathrm{U}_{2n}$ is $\mathfrak{sp}_n(\mathbb{C}) \cap \mathfrak{u}_n$ and so consists of matrices X such that $\Omega X + X^T \Omega = 0$ and $X^\dagger = -X$.

Exercise 3.6. Show that the matrices in $\mathfrak{sp}_n(\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) are precisely the $2n \times 2n$ matrices in $M_{2n}(\mathbb{K})$ of the form

$$\begin{pmatrix} A & B \\ C & -A^T \end{pmatrix}$$

where A, B, C are $n \times n$ blocks such that

$$B^T = B, \quad C^T = C.$$

Exercise 3.7. Find the dimensions of the Lie algebras $\mathfrak{sl}_n(\mathbb{R})$, $(\mathfrak{s})\mathfrak{o}_n(\mathbb{R})$, $\mathfrak{sp}_n(\mathbb{R})$ and $(\mathfrak{s})\mathfrak{u}_n$.

Exercise 3.8. Show that the Lie algebra $\mathfrak{o}_{p,q}(\mathbb{R})$ of $\mathrm{O}_{p,q}(\mathbb{R})$ comprises the matrices $X \in \mathrm{M}_n(\mathbb{R})$ such that $J_{p,q}X + X^T J_{p,q} = 0$. Check that these matrices are all traceless, so that moreover $\mathfrak{so}_{p,q}(\mathbb{R}) = \mathfrak{o}_{p,q}(\mathbb{R})$.

Exercise 3.9. Check that $\mathfrak{gl}_n(\mathbb{C})$, $\mathfrak{sl}_n(\mathbb{C})$, $\mathfrak{so}_n(\mathbb{C})$, and $\mathfrak{sp}_n(\mathbb{C})$ are complex vector spaces, so that $\mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{C})$, $(\mathrm{S})\mathrm{O}_n(\mathbb{C})$, and $\mathrm{Sp}_n(\mathbb{C})$ are complex Lie groups, while $(\mathfrak{s})\mathfrak{u}_n$ and $\mathfrak{usp}_n$ are *not* complex. (Indeed, $(\mathrm{S})\mathrm{U}_n$ and USp_n are *not* complex Lie groups.)

3.4 Abstract Lie algebras

We can formalize the structure we have seen on the Lie algebra of a matrix group into the abstract notion of a Lie algebra. Since we want to consider Lie algebras over both $\mathbb{R}$ and $\mathbb{C}$, let us just say that $\mathbb{K}$ is some field of characteristic 0.

Definition 3.4.1. A *Lie algebra* over $\mathbb{K}$ is a vector space $\mathfrak{g}$ over $\mathbb{K}$ together with a bilinear map $[-,-]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ (called the *Lie bracket*) such that

- the Lie bracket is antisymmetric: for $x, y \in \mathfrak{g}$ we have

$$[x, y] = -[y, x],$$

- the Lie bracket satisifies the *Jacobi identity*

$$[x, [y, z]] = [[x, y], z] + [y, [x, z]]$$

Exercise 3.10. Show that if $[-,-]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is an antisymmetric bilinear map, then the Jacobi identity for $x, y, z \in \mathfrak{g}$ as stated above is equivalent to the identity

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

Example 3.4.2. Suppose A is an associative $\mathbb{K}$-algebra. Then the *commutator bracket* $[x, y] := xy - yx$ makes A a Lie algebra. The bracket is clearly antisymmetric, while the Jacobi identity follows from the calculation

$$\begin{aligned}
[[x, y], z] + [y, [x, z]] &= [(xy - yx), z] + [y, (xz - zx)] \\
&= (xy - yx)z - z(xy - yx) + y(xz - zx) - (xz - zx)y \\
&= xyz - yxz - zxy + zyx + yxz - yzx - xzy + zxy \\
&= xyz + zyx - yzx - xzy \\
&= x(yz - zy) - (yz - zy)x \\
&= [x, [y, z]].
\end{aligned}$$

Example 3.4.3. The same calculation shows that the Jacobi identity holds for the Lie algebra of any matrix group, since this was defined using the commutator bracket for matrix multiplication.

Exercise 3.11. Show that the *cross product* (or vector product) gives a Lie algebra structure on $\mathbb{R}^3$. Prove that this Lie algebra is isomorphic to $\mathfrak{so}_3(\mathbb{R})$.

Definition 3.4.4. If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras over $\mathbb{K}$, then a *homomorphism* $\phi\colon \mathfrak{g} \to \mathfrak{h}$ is a $\mathbb{K}$-linear map such that $[\phi(x), \phi(y)] = \phi([x, y])$ for all $x, y \in \mathfrak{g}$.

Example 3.4.5. For any $\mathbb{K}$-vector space V, the vector space $\mathrm{End}(V)$ of linear endomorphisms of V is an associative $\mathbb{K}$-algebra with composition of endomorphisms as multiplication. As in Example 3.4.2 this also gives it a Lie algebra structure; we denote this Lie algebra by $\mathfrak{gl}(V)$. If $\mathfrak{g}$ is a Lie algebra, consider the linear map $\mathrm{ad}_{\mathfrak{g}}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ defined by $\mathrm{ad}_{\mathfrak{g}}(x) = [x, -]$. We claim that this is a Lie algebra homomorphism; to see this, we compute

$$\begin{aligned}[\mathrm{ad}_{\mathfrak{g}}(x), \mathrm{ad}_{\mathfrak{g}}(y)] &= \mathrm{ad}_{\mathfrak{g}}(x) \circ \mathrm{ad}_{\mathfrak{g}}(y) - \mathrm{ad}_{\mathfrak{g}}(y) \circ \mathrm{ad}_{\mathfrak{g}}(x) \\ &= [x, [y, -]] - [y, [x, -]], \\ \mathrm{ad}_{\mathfrak{g}}([x, y]) &= [[x, y], -].\end{aligned}$$

That these agree for all $x, y \in \mathfrak{g}$ is precisely the Jacobi identity, so that this is in fact equivalent to $\mathrm{ad}_{\mathfrak{g}}$ being a Lie algebra homomorphism. This is one motivation for including the Jacobi identity in the definition of a Lie algebra. (The homomorphism $\mathrm{ad}_{\mathfrak{g}}$ is known as the *adjoint representation* of $\mathfrak{g}$, and this will play an important role later on.)

Definition 3.4.6. Suppose V is an $\mathbb{R}$-vector space. Its *complexification*[1] $V \otimes \mathbb{C}$ or $V \oplus iV$ consists of formal sums $v + iw$ with $v, w \in V$; it is a $\mathbb{C}$-vector space where the operations are defined in the obvious way. We think of V as the $\mathbb{R}$-subspace of $V \otimes \mathbb{C}$ where the coefficient of i is 0.

The complexification has the following universal property:

Lemma 3.4.7. *Let V be an $\mathbb{R}$-vector space. If W is a $\mathbb{C}$-vector space and $\phi\colon V \to W$ is an $\mathbb{R}$-linear map, then ϕ extends uniquely to a $\mathbb{C}$-linear map $\phi_{\mathbb{C}}\colon V \otimes \mathbb{C} \to W$.*

Proof. We must have $\phi_{\mathbb{C}}(v + iw) = \phi_{\mathbb{C}}(v) + i\phi_{\mathbb{C}}(w) = \phi(v) + i\phi(w)$, which proves uniqueness. This equation also clearly does define a $\mathbb{C}$-linear map. □

Proposition 3.4.8. *Suppose $\mathfrak{g}$ is a Lie algebra over $\mathbb{R}$. Then $\mathfrak{g} \otimes \mathbb{C}$ has a unique $\mathbb{C}$-Lie algebra structure such that a $\mathbb{C}$-linear map $\mathfrak{g} \otimes \mathbb{C} \to \mathfrak{h}$ is a $\mathbb{C}$-Lie algebra homomorphism if and only if the restriction $\mathfrak{g} \to \mathfrak{h}$ is an $\mathbb{R}$-Lie algebra homomorphism.*

Proof. The property requires in particular that the inclusion $\mathfrak{g} \to \mathfrak{g} \otimes \mathbb{C}$ must be an $\mathbb{R}$-Lie algebra homomorphism (since this extends to the identity homomorphism of $\mathfrak{g} \otimes \mathbb{C}$). Together with $\mathbb{C}$-bilinearity this forces the bracket on $\mathfrak{g} \otimes \mathbb{C}$ to be given by

$$[v + iw, v' + iw'] = ([v, v'] - [w, w']) + i([w, v'] + [v, w']).$$

[1] As the notation suggests, this is an example of a *tensor product*, which we will consider in more detail later on, but in this simple case we may just as well give a more concrete definition.

It is easy to check that this does indeed define a $\mathbb{C}$-Lie algebra, and that it satisfies the required property. □

Corollary 3.4.9. *Suppose $\mathfrak{g} \subseteq \mathfrak{gl}_n(\mathbb{C})$ is an $\mathbb{R}$-Lie subalgebra with the property that if $X \neq 0$ is in $\mathfrak{g}$, then iX is not in $\mathfrak{g}$. Then $\mathfrak{g} \otimes \mathbb{C}$ is isomorphic to the $\mathbb{C}$-Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ consisting of matrices of the form $X + iY$ with $X, Y \in \mathfrak{g}$.*

Proof. The inclusion of $\mathfrak{g}$ extends uniquely to a $\mathbb{C}$-Lie algebra homomorphism $\mathfrak{g} \otimes \mathbb{C} \to \mathfrak{gl}_n(\mathbb{C})$. This takes the formal linear combination $X + iY$ with $X, Y \in \mathfrak{g}$ to the corresponding linear combination of matrices. To see that $\mathfrak{g} \otimes \mathbb{C}$ is isomorphic to the $\mathbb{C}$-Lie algebra of matrices of this form, we need only prove that this homomorphism if injective. But if $X+iY = 0$ for $X, Y \in \mathfrak{g}$ then $iY = -X$ lies in $\mathfrak{g}$, which is impossible unless $X = Y = 0$. □

We can use Corollary 3.4.9 to describe the complexifications of the Lie algebras we have considered above.

Example 3.4.10. Corollary 3.4.9 applies to $\mathfrak{u}_n$: if $X^\dagger = -X$ then $(iX)^\dagger = iX$, so iX does not lie in $\mathfrak{u}_n$ if $X \neq 0$. Hence $\mathfrak{u}_n \otimes \mathbb{C}$ can be identified with the Lie algebra of matrices on $\mathfrak{gl}_n(\mathbb{C})$ that can be written as $X + iY$ with $X, Y \in \mathfrak{u}_n$. But for any matrix M we can write $M = X + iY$ with $X = \frac{1}{2}(M - M^\dagger)$ and $Y = \frac{1}{2i}(M + M^\dagger)$, which are both skew-Hermitian. Thus we get an isomorphism

$$\mathfrak{u}_n \otimes \mathbb{C} \cong \mathfrak{gl}_n(\mathbb{C}).$$

of $\mathbb{C}$-Lie algebras.

We have further identifications as follows:

$$\begin{aligned}
\mathfrak{gl}_n(\mathbb{R}) \otimes \mathbb{C} &\cong \mathfrak{gl}_n(\mathbb{C}), \\
\mathfrak{su}_n \otimes \mathbb{C} &\cong \mathfrak{sl}_n(\mathbb{C}), \\
\mathfrak{sl}_n(\mathbb{R}) \otimes \mathbb{C} &\cong \mathfrak{sl}_n(\mathbb{C}), \\
(\mathfrak{s})\mathfrak{o}_n(\mathbb{R}) \otimes \mathbb{C} &\cong (\mathfrak{s})\mathfrak{o}_n(\mathbb{C}), \\
\mathfrak{sp}_n(\mathbb{R}) \otimes \mathbb{C} &\cong \mathfrak{sp}_n(\mathbb{C}), \\
\mathfrak{usp}_n \otimes \mathbb{C} &\cong \mathfrak{sp}_n(\mathbb{C}).
\end{aligned}$$

Exercise 3.12. Check as many of these isomorphisms as you can be bothered to.

3.5 Matrix logarithms and the Lie product formula

Our goal in this section is to define a "logarithm" that gives an inverse to the matrix exponential in a neighbourhood of the identity, and use this to prove the Lie product formula. Recall that the Taylor series for the real logarithm at 1 is given by

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

and the same formula defines a complex logarithm in a neighbourhood of 1:

Lemma 3.5.1. *The power series*

$$\log z = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m} \tag{3.5}$$

converges to a holomorphic function on the open disc of radius 1 *around* $1 \in \mathbb{C}$*. For all* $z \in \mathbb{C}$ *with* $|z-1| < 1$ *we have* $e^{\log z} = z$*, and for all* $z \in \mathbb{C}$ *with* $|z| < \log 2$ *we have* $|e^z - 1| < 1$ *and* $\log e^z = z$*.*

Proof. The series has radius of convergence 1 and defines a holomorphic function on $D := \{z \in \mathbb{C} : |z-1| < 1\}$ that agrees with the real logarithm for real z in the interval $0 < z < 2$. Thus $\exp(\log z) = z$ for z in this real interval, but then as both sides are holomorphic functions they must agree on the whole open disc D. On the other hand, if $|u| < \log 2$, then

$$|e^u - 1| = \left| \sum_{n=1}^{\infty} \frac{u^n}{n!} \right| \leq \sum_{n=1}^{\infty} \frac{|u|^n}{n!} = e^{|u|} - 1 < 1,$$

so that $\log e^u$ is defined. Now since $\log e^u = u$ for real u and both sides are holomorphic on the disc $\{u \in \mathbb{C} : |u| < \log 2\}$, we again conclude that they must agree on the entire disc. □

We can extend the same definition to matrices:

Definition 3.5.2. For $A \in \mathrm{M}_n(\mathbb{C})$, we define $\log A$ by

$$\log A := \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m} \tag{3.6}$$

whenever this series converges. (Note that for $A \in \mathrm{M}_n(\mathbb{R})$ we must clearly have $\log A \in \mathrm{M}_n(\mathbb{R})$ if it exists.)

Theorem 3.5.3. *If* $A \in \mathrm{M}_n(\mathbb{C})$ *satisfies* $\|A - I\| < 1$ *then* $\log A$ *exists and satisfies* $e^{\log A} = A$*, and* $\log(-)$ *defines a continuous function on this open set. Moreover, if* $A \in \mathrm{M}_n(\mathbb{C})$ *satisfies* $\|A\| < \log 2$*, then* $\|e^A - I\| < 1$ *and* $\log e^A = A$*.*

Proof. Since $\|(A-I)^m\| \leq \|A - I\|^m$, we have

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

Since the series (3.5) converges absolutely in the disc of radius 1 around 1, we see that the series (3.6) converges absolutely if $\|A - I\| < 1$. It is continuous by the same argument as in the proof of Proposition 3.1.4.

Suppose A is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$, so that

$$A = C^{-1}\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} C$$

for some invertible matrix C. If $\|A - I\| < 1$, then by Exercise 3.3 we have $|\lambda_i - 1| < 1$ for all i. Then the definition gives

$$\log A = C^{-1}\begin{pmatrix} \log \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \log \lambda_n \end{pmatrix} C$$

and so by Lemma 3.1.8 and Lemma 3.5.1 we have

$$e^{\log A} = C^{-1}\begin{pmatrix} e^{\log \lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\log \lambda_n} \end{pmatrix} C = A.$$

Since both sides are continuous in A and diagonalizable matrices are dense, we must have that the equation $e^{\log A} = A$ holds for all A with $\|A - I\| < 1$.

If $\|A\| < \log 2$, then we have

$$\|e^A - I\| = \left\| \sum_{n=1}^{\infty} \frac{A^n}{n!} \right\| \leq \sum_{n=1}^{\infty} \frac{\|A\|^n}{n!} = e^{\|A\|} - 1 < 1.$$

To show that $\log e^A = A$ we may again assume that A is diagonalizable, in which case the eigenvalues must have absolute value smaller than $\log 2$ by Exercise 3.3, and so the equation follows as above by applying Lemma 3.5.1 to the eigenvalues.

□

Warning 3.5.4. It is perfectly possible to have $\log e^X \neq X$ — indeed, this happens already in $\mathbb{C}$, since $\log e^{2\pi i} = \log 1 = 0 \neq 2\pi i$.

Lemma 3.5.5. *There exists $c > 0$ such that for all $B \in \mathrm{M}_n(\mathbb{C})$ with $\|B\| < 1/2$, we have*

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

Proof. We have by definition

$$\log(I + B) - B = \sum_{m=2}^{\infty} (-1)^{m+1} \frac{B^m}{m} = B^2 \sum_{m=0}^{\infty} (-1)^{m+1} \frac{B^m}{m+2},$$

so that $\|B\| < 1/2$ implies

$$\|\log(I + B) - B\| \leq \|B\|^2 \sum_{m=0}^{\infty} \frac{(1/2)^m}{m+2}.$$

□

Notation 3.5.6. In the next proof it is convenient to use the "big-O" notation, where we say that $f(X) = O(g(\|X\|))$ if there exists a constant M such that $\|f(X)\| \leq Mg(\|X\|)$ for X sufficiently close to some limiting value. In this notation, Lemma 3.5.5 says that

$$\log(I + B) = B + O(\|B\|^2)$$

as $\|B\| \to 0$.

We are now ready to prove the Lie product formula:

Proof of Theorem 3.2.10. We want to show that $e^{X+Y} = \lim_{m\to\infty} \left(e^{X/m}e^{Y/m}\right)^m$. From the power series for the matrix exponential we have

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

as $m \to \infty$. Since the exponential is continuous, $e^{X/m}e^{Y/m}$ converges to I as $m \to \infty$. Thus it is in the domain of the matrix logarithm for all sufficiently large m. In this case, Lemma 3.5.5 implies that we have

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

Exponentiating this gives

$$e^{X/m}e^{Y/m} = e^{\frac{X}{m}+\frac{Y}{m}+O(1/m^2)},$$

and so

$$(e^{X/m}e^{Y/m})^m = e^{X+Y+O(1/m)}.$$

Since the exponential is continuous we get

$$\lim_{m\to\infty}(e^{X/m}e^{Y/m})^m = \lim_{m\to\infty} e^{X+Y+O(1/m)} = e^{X+Y},$$

which is what we want to prove. □

3.6 Functoriality of Lie algebras

Our next goal is to show that a homomorphism of Lie groups induces a homomorphism between their Lie algebras. More precisely, we will show:

Theorem 3.6.1. *Let (G, j) and (H, j') be embedded matrix groups with $\mathfrak{g} = \mathfrak{L}(G, j)$ and $\mathfrak{h} = \mathfrak{L}(H, j')$ the corresponding Lie algebras, and let $\Phi\colon G \to H$ be a continuous homomorphism. Then there exists a unique $\mathbb{R}$-linear map $\mathfrak{L}(\Phi)\colon \mathfrak{g} \to \mathfrak{h}$ such that*

$$j'\Phi(j^{-1}e^X) = e^{\mathfrak{L}(\Phi)(X)}$$

for all $X \in \mathfrak{g}$. (We will usually drop the embeddings from the notation and so from now on we write

$$\Phi(e^X) = e^{\mathfrak{L}(\Phi)(X)},$$

and so on.) Moreover, this map satisfies the following:

(1) $\mathfrak{L}(\Phi)(X) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Phi(e^{tX})$ *for all* $X \in \mathfrak{g}$,

(2) $\mathfrak{L}(\Phi)(AXA^{-1}) = \Phi(A)\mathfrak{L}(\Phi)(X)\Phi(A)^{-1}$ *for all* $A \in G, X \in \mathfrak{g}$,

(3) $\mathfrak{L}(\Phi)([X,Y]) = [\mathfrak{L}(\Phi)(X), \mathfrak{L}(\Phi)(Y)]$ *(i.e.* $\mathfrak{L}(\Phi)$ *is a Lie algebra homomorphism).*

Warning 3.6.2. Strictly speaking, the linear map $\mathfrak{L}(\Phi)$ here depends on the chosen embeddings j, j', in that its source and target depend on them. We will occasionally write

$$\mathfrak{L}(\Phi, j, j') : \mathfrak{L}(G, j) \to \mathfrak{L}(H, j')$$

when we want to emphasize this.

A key part in the proof is the classification of one-parameter subgroups of a matrix group:

Definition 3.6.3. A *one-parameter subgroup* of a topological group G G is a continuous homomorphism $A : (\mathbb{R}, +) \to G$, i.e. A is a continuous function such that $A(0) = I$ and $A(s+t) = A(s)A(t)$. (Note that it is the *image* of A that is a subgroup, we don't require the parametrization A to be injective!)

Proposition 3.6.4 (Classification of one-parameter subgroups)**.** *If A is a one-parameter subgroup of $\mathrm{GL}_n(\mathbb{C})$, then there exists a unique $X \in \mathrm{M}_n(\mathbb{C})$ such that $A(t) = e^{tX}$. Moreover, A is differentiable and $X = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} A(t)$.*

We note the following immediate consequence:

Corollary 3.6.5. *Let $G \subseteq \mathrm{GL}_n(\mathbb{C})$ be a matrix group with Lie algebra $\mathfrak{g}$. If $A : \mathbb{R} \to G$ is a one-parameter subgroup, then there exists a unique $X \in \mathfrak{g}$ such that $A(t) = e^{tX}$. Moreover, A is differentiable and $X = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} A(t)$.*

For the proof we need the following observations:

Lemma 3.6.6. *Let B_r be the open ball of radius r around 0 in $\mathrm{M}_n(\mathbb{C})$. For $0 < \epsilon \leq \log 2$, set $U_\epsilon = \exp(B_\epsilon)$. Then U_ϵ is open in $\mathrm{M}_n(\mathbb{C})$.*

Proof. If $X \in U_\epsilon$ then by definition $X = \exp(Y)$ for some $Y \in B_\epsilon$. As $\epsilon \leq \log 2$ we have $\|X - I\| < 1$ and $Y = \log(X)$ by Theorem 3.5.3. Then there exists an open neighbourhood V around X on which log is defined, and by continuity if we choose this small enough then its image must lie in B_ϵ. In this case $V \subseteq U_\epsilon = \exp(B_\epsilon)$, which shows that U_ϵ is open. □

Lemma 3.6.7. *If* $0 < \epsilon \leq \frac{1}{2}\log 2$ *then every* $X \in U_\epsilon$ *has a* unique *square root in* U_ϵ*, given by* $\exp(\frac{1}{2}\log X)$.

Proof. Suppose X is in U_ϵ. If $Y := \exp(\frac{1}{2}\log X)$ then $Y^2 = X$ so this is a square root. If Z is another square root in U_ϵ, then

$$\exp(2\log Z) = Z^2 = X = \exp(\log X).$$

Here we have $2\log Z \in B_{2\epsilon}$ and $\log X \in B_\epsilon$; since exp is injective on $B_{\log 2}$ we get $2\log Z = \log X$ and so $Z = \exp(\frac{1}{2}\log X) = Y$. □

Proof of Proposition 3.6.4. If there exists such an X then we know from Proposition 3.1.10 that $A(t) = e^{tX}$ is differentiable and that we can recover X as $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} A(t)$; this proves uniqueness.

Since A is continuous and $U_{\frac{1}{2}\log 2}$ is open, we can choose $t_0 > 0$ such that $A(t) \in U_{\frac{1}{2}\log 2}$ for all t with $|t| \leq t_0$. Then we can define

$$X := \frac{1}{t_0}\log A(t_0),$$

and this satisfies $e^{t_0 X} = A(t_0)$. Now observe that $A(t_0/2)$ also lies in $U_{\frac{1}{2}\log 2}$ and this satisfies

$$A(t_0/2)^2 = A(t_0).$$

Hence Lemma 3.6.7 implies that $A(t_0/2) = e^{\frac{t_0}{2}X}$. Iterating this argument, we get that $A(t_0/2^k) = e^{\frac{t_0}{2^k}X}$ for all $k \in \mathbb{N}$. Moreover, for any $m \in \mathbb{Z}$ we have

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

The continuous functions $A(t)$ and e^{tX} thus agree on all real numbers of the form $\frac{mt_0}{2^k}$ for $k \in \mathbb{N}$, $m \in \mathbb{Z}$. But such numbers are dense in $\mathbb{R}$, so the functions must agree everywhere. □

Proof of Theorem 3.6.1. For any $X \in \mathfrak{g}$, the map $t \mapsto \Phi(e^{tX})$ is a one-parameter subgroup of H. By Proposition 3.6.4 there is then a unique matrix $\phi(X)$ such that $\Phi(e^{tX}) = e^{t\phi(X)}$ — in particular, $\Phi(e^X) = e^{\phi(X)}$. Moreover, $\phi(X) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi(e^{tX})$. Next we check that ϕ is $\mathbb{R}$-linear: $\phi(sX)$ satisfies

$$e^{t\phi(sX)} = \Phi(e^{tsX}) = e^{ts\phi(X)}$$

for all t, so that $\phi(sX) = s\phi(X)$. For $X, Y \in \mathfrak{g}$ we have from the Lie product formula (Theorem 3.2.10) and the assumption that Φ is a continuous homomorphism that

$$\begin{aligned}
e^{t\phi(X+Y)} &= \Phi(e^{tX+tY}) \\
&= \Phi(\lim_{m\to\infty}(e^{tX/m}e^{tY/m})^m) \\
&= \lim_{m\to\infty}(\Phi(e^{tX/m})\Phi(e^{tY/m}))^m \\
&= \lim_{m\to\infty}(e^{t\phi(X)/m}e^{t\phi(Y)/m})^m \\
&= e^{t(\phi(X)+\phi(Y))},
\end{aligned}$$

which implies that $\phi(X+Y) = \phi(X) + \phi(Y)$.

This proves existence. To prove uniqueness, suppose ψ is an $\mathbb{R}$-linear map such that $\Phi(e^X) = e^{\psi(X)}$ for $X \in \mathfrak{g}$. Since tX is also in $\mathfrak{g}$ for $t \in \mathbb{R}$ we then have $\Phi(e^{tX}) = e^{t\psi(X)}$ for all t by linearity. Hence Proposition 3.6.4 implies that

$$\psi(X) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi(e^{tX}) = \phi(X),$$

as required.

It remains to verify the two remaining properties. For conjugation, we have

$$\begin{aligned}
e^{t\phi(AXA^{-1})} &= \Phi(e^{tAXA^{-1}}) \\
&= \Phi(Ae^{tX}A^{-1}) \\
&= \Phi(A)\Phi(e^{tX})\Phi(A)^{-1} \\
&= \Phi(A)e^{t\phi(X)}\Phi(A)^{-1} \\
&= e^{t\Phi(A)\phi(X)\Phi(A)^{-1}},
\end{aligned}$$

and so $\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A)^{-1}$. For commutators, we have (as in the proof of Proposition 3.2.13) that $[X, Y] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{tX}Ye^{-tX}$. Since the derivative commutes with the linear map ϕ, we get

$$\begin{aligned}
\phi([X,Y]) &= \phi\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{tX}Ye^{-tX}\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \phi(e^{tX}Ye^{-tX}) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi(e^{tX})\phi(Y)\Phi(e^{-tX}) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{t\phi(X)}\phi(Y)e^{-t\phi(X)} \\
&= [\phi(X), \phi(Y)],
\end{aligned}$$

applying at the end the same formula for commutators in $\mathfrak{h}$. □

Example 3.6.8. Recall the adjoint homomorphism $\mathrm{Ad}_G\colon G \to \mathrm{GL}(\mathfrak{g})$ from Observation 3.2.15. We claim $\mathfrak{L}(\mathrm{Ad}_G)$ is the Lie algebra homomorphism $\mathrm{ad}_{\mathfrak{g}}$ from Example 3.4.5. This follows from the computation

$$\mathrm{ad}_{\mathfrak{g}}(X)(Y) = [X,Y] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{tX}Ye^{-tX} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}_G(e^{tX})(Y)$$

using equation (3.4) from the proof of Proposition 3.2.13.

Next, let's look at the map of Lie algebras induced by the 2-to-1 homomorphism from SU_2 to SO_3; we will see that this map of Lie algebras is an isomorphism. For this we need a generators and relations description of $\mathfrak{su}_2$, which we leave as an exercise:

Exercise 3.13. Show that the following is a basis for the Lie algebra $\mathfrak{su}_2$:

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

and check that the Lie bracket is given by the relations

$$[E_1, E_2] = E_3, \quad [E_2, E_3] = E_1, \quad [E_3, E_1] = E_2.$$

Example 3.6.9. Consider the 2-to-1 surjective homomorphism $r\colon \mathrm{SU}_2 \to \mathrm{SO}_3(\mathbb{R})$ we defined in Construction 2.3.4 and Proposition 2.3.6. We defined this by $r(q)(x) = qxq^{-1}$ where q is a unit-length quaternion and x is a purely imaginary quaternion. Thinking of quaternion multiplication as matrix multiplication as in Exercise 2.8, we can express this as the matrix product

$$r(Q)(M) = QMQ^{-1},$$

for $Q \in \mathrm{SU}_2$ and M in the 3-dimensional $\mathbb{R}$-vector space V of matrices

$$\begin{pmatrix} bi & c+di \\ -c+di & -bi, \end{pmatrix} \quad b,c,d \in \mathbb{R}.$$

We then have

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} r(e^{tX})(M) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{tX}Me^{-tX} = XM - MX = [X,M].$$

Note that here V consists precisely of the traceless skew-Hermitian matrices; this means we can actually identify r with the adjoint action of SU_2 on its Lie algebra, and so this computation agrees with Example 3.6.8. Using the commutation relations from Exercise 3.13, we can now immediately write down the 3×3 matrices F_i that represent $\mathfrak{L}(r)(E_i)$ with respect to the basis E_1, E_2, E_3:

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

Since these 3 matrices clearly form a basis for the vector space $\mathfrak{so}_3(\mathbb{R})$ of skew-symmetric matrices in $M_3(\mathbb{R})$, we see that $\mathfrak{L}(r)$ must be a Lie algebra isomorphism.

Exercise 3.14. Verify directly that in terms of the basis F_i, the Lie bracket on $\mathfrak{so}_3(\mathbb{R})$ is given by the relations:

$$[F_1, F_2] = F_3, \quad [F_2, F_3] = F_1, \quad [F_3, F_1] = F_2,$$

and so $\mathfrak{so}_3(\mathbb{R})$ is abstractly isomorphic to $\mathfrak{su}_2$ since it has the same generators-and-relations description.

Proposition 3.6.10. *Suppose $\Phi\colon G \to H$ is a continuous homomorphism of matrix groups. Then $\ker\Phi$ is a closed, normal subgroup of G (and so in particular also a matrix group), and*

$$\mathfrak{L}(\ker\Phi) = \ker\mathfrak{L}(\Phi).$$

Proof. $\ker\Phi$ is a normal subgroup by elementary group theory, and since it is the continuous preimage of the closed set $\{I\} \subseteq H$ it is also closed. If G is isomorphic to a closed subgroup of $GL_n(\mathbb{R})$, then so is $\ker\Phi$ by restriction, so $\ker\Phi$ is a matrix group. Let $\phi := \mathfrak{L}(\Phi)$, then if $X \in \ker\phi$, we have

$$\Phi(e^{tX}) = e^{t\phi(X)} = I,$$

so that $X \in \mathfrak{L}(\ker\Phi)$. Conversely, if e^{tX} is in $\ker\Phi$ for all t then $e^{t\phi(X)} = \Phi(e^{tX}) = I$, and differentiating gives $\phi(X) = 0$. Thus $\ker\phi = \mathfrak{L}(\ker\Phi)$. □

Proposition 3.6.11. *Suppose (G, j), (H, j'), (K, j'') are embedded matrix groups, and $\Phi\colon G \to H$ and $\Psi\colon H \to K$ are continuous homomorphisms. Then we have*

$$\mathfrak{L}(\Psi\circ\Phi, j, j'') = \mathfrak{L}(\Psi, j', j'') \circ \mathfrak{L}(\Phi, j, j').$$

Moreover, $\mathfrak{L}(\mathrm{id}_G, j, j) = \mathrm{id}_{\mathfrak{L}(G,j)}$.

Proof. Let $\psi = \mathfrak{L}(\Psi)$ and $\phi = \mathfrak{L}(\Phi)$. Then for $X \in \mathfrak{g}$, we have

$$e^{t\mathfrak{L}(\Psi\Phi)(X)} = (\Psi\Phi)(e^{tX}) = \Psi(e^{t\phi(X)}) = e^{t\psi(\phi(X))},$$

and hence $\mathfrak{L}(\Psi\Phi)(X) = \psi(\phi(X))$. Moreover,

$$e^{t\mathfrak{L}(\mathrm{id}_G)(X)} = \mathrm{id}_G(e^{tX}) = e^{tX},$$

so that $\mathfrak{L}(\mathrm{id}_G)(X) = X$. □

As a formal consequence of this, we see that the Lie algebra of a matrix group is well-defined up to a canonical isomorphism:

Corollary 3.6.12. *Let G be a topological group, and let $j\colon G \to GL_n(\mathbb{R})$ and $j'\colon G \to GL_m(\mathbb{R})$ be two continuous homomorphisms that identify G as a closed subgroup of the two general linear groups. Then $\mathfrak{L}(\mathrm{id}_G, j, j')\colon \mathfrak{L}(G, j) \to \mathfrak{L}(G, j')$ is a Lie algebra isomorphism with inverse $\mathfrak{L}(\mathrm{id}_G, j', j)$.*

Proof. From Proposition 3.6.11 we have

$$\mathfrak{L}(\mathrm{id}_G, j', j) \circ \mathfrak{L}(\mathrm{id}_G, j, j') = \mathfrak{L}(\mathrm{id}_G, j, j) = \mathrm{id}_{\mathfrak{L}(G,j)},$$

and similarly for the composite in the other order. □

3.7 The exponential map for matrix groups

Suppose $G \subseteq \mathrm{GL}_n(\mathbb{C})$ is a matrix group with Lie algebra $\mathfrak{g}$. Then by definition the matrix exponential $\exp = e^{(-)} : \mathrm{M}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ restricts to a map

$$\exp\colon \mathfrak{g} \to G,$$

which we call the *exponential map for* G.

Remark 3.7.1. It can be shown that every invertible matrix can be written as an exponential, i.e. exp is surjective onto $\mathrm{GL}_n(\mathbb{R})$. However, this need not be the case for the exponential of an arbitrary Lie group. (See [2, Example 3.41].)

Theorem 3.7.2. *Suppose* $G \subseteq \mathrm{GL}_n(\mathbb{C})$ *is a matrix group with Lie algebra* $\mathfrak{g}$*. For* $\epsilon > 0$ *let*

$$U_\epsilon := \{X \in \mathrm{M}_n(\mathbb{C}) : \|X\| < \epsilon\}, \quad V_\epsilon = \exp(U_\epsilon).$$

Then there exists $0 < \epsilon < \log 2$ *such that* $A \in V_\epsilon$ *lies in* G *if and only if* $\log(A)$ *lies in* $\mathfrak{g}$*. In particular,* exp *restricts to a homeomorphism*

$$U_\epsilon \cap \mathfrak{g} \cong V_\epsilon \cap G.$$

Proof. We know that if X is in $\mathfrak{g}$, then e^X is in G. If $\epsilon < \log 2$ then $\log(e^X)$ exists and equals X for all $X \in U_\epsilon$, so what we need to prove is that if $A \in V_\epsilon \cap G$ then $\log A \in \mathfrak{g}$, provided ϵ is sufficiently small.

We know that $\mathfrak{g}$ is a real vector subspace of $\mathrm{M}_n(\mathbb{C}) \cong \mathbb{R}^{2n^2}$; let $\mathfrak{g}^\perp$ denote its orthogonal complement with respect to the standard inner product on $\mathbb{R}^{2n^2}$. Then we can write any $Z \in \mathrm{M}_n(\mathbb{C})$ uniquely as $Z_\mathfrak{g} + Z'$ with $Z_\mathfrak{g} \in \mathfrak{g}$ and $Z' \in \mathfrak{g}^\perp$. Define a map $\Phi\colon \mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_n(\mathbb{C})$ by $\Phi(Z) = e^{Z_\mathfrak{g}} e^{Z'}$. Since the exponential is infinitely differentiable (Observation 3.1.11), so is Φ, and if either $Z \in \mathfrak{g}$ or $Z \in \mathfrak{g}^\perp$ we have

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi(tZ) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{tZ} = Z.$$

From this we conclude that the derivative of Φ at the origin is the identity. As this is invertible, we can apply the inverse function theorem to conclude that Φ has a continuous local inverse defined in some neighbourhood of I.

We now argue by contradiction. Suppose then that for every $0 < \epsilon < \log 2$ there exists $A \in V_\epsilon \cap G$ such that $\log A \notin \mathfrak{g}$. This means that there exists a sequence of $A_m \in G$ such that $A_m \to I$ as $m \to \infty$, but $\log A_m \notin \mathfrak{g}$ for all m. Using the local inverse of Φ, for m sufficiently large we can write A_m as $e^{X_m} e^{Y_m}$ where $X_m \in \mathfrak{g}$, $Y_m \in \mathfrak{g}^\perp$, both converge to 0 as $m \to \infty$, and $Y_m \neq 0$ for all m. Then since A_m and e^{X_m} are both in G, we have that $e^{Y_m} = e^{-X_m} A_m \in G$.

Now consider the sequence $Y_m / \|Y_m\|$; since the unit sphere in $\mathfrak{g}^\perp$ is compact, there exists some convergent subsequence $Y_{m'}/\|Y_{m'}\| \to Y$ with $Y \in \mathfrak{g}^\perp$ and $\|Y\| = 1$. We will show that Y lies in $\mathfrak{g}$, which gives a contradiction since by definition $\mathfrak{g}^\perp$ is the orthogonal complement of $\mathfrak{g}$.

We need to show that e^{tY} is in G for all $t \in \mathbb{R}$. Note that if $B_{m'} := e^{Y_{m'}}$ then $Y_{m'} = \log B_{m'}$ (as $\epsilon < \log 2$). Since $B_{m'} \to I$ we have $\|Y_{m'}\| \to 0$. For any t we can therefore find integers $k_{m'}$ such that $k_{m'}\|Y_{m'}\| \to t$. Then we have

$$e^{k_{m'}Y_{m'}} = e^{(k_{m'}\|Y_{m'}\|)(Y_{m'}/\|Y_{m'}\|)} \to e^{tY}.$$

But here $e^{k_{m'}Y_{m'}} = (B_{m'})^{k_{m'}}$ which lies in G. Since G is closed it follows that the limit e^{tY} also lies in G, as required. □

We can restate Theorem 3.7.2 less precisely as:

Corollary 3.7.3. *Let G be a matrix group with Lie algebra $\mathfrak{g}$. Then there exists a neighbourhood U of $0 \in \mathfrak{g}$ such that the exponential $\mathfrak{g} \to G$ maps U homeomorphically onto a neighbourhood of I in G.*

Theorem 3.7.2 has a number of pleasant consequences:

Corollary 3.7.4. *If G is a matrix group with Lie algebra $\mathfrak{g}$, then G is a smooth manifold of dimension* $\dim \mathfrak{g}$.

Proof. Let $k := \dim \mathfrak{g}$. By Corollary 3.7.3 the identity of G has an open neighbourhood V that is homeomorphic to an open subset U of $\mathfrak{g} \cong \mathbb{R}^k$. Then for any point $A \in G$ the open neighbourhood $A \cdot V$ is also homeomorphic to U, so that G is a topological manifold. Moreover, the transition function $U \to U$ between V and $A \cdot V$ is given by $X \mapsto \log(A^{-1}\exp(X))$, which is smooth, and so G is a smooth k-manifold. □

Corollary 3.7.5. *Suppose $G \subseteq \mathrm{GL}_n(\mathbb{R})$ is a matrix group with Lie algebra $\mathfrak{g}$. Then a matrix $X \in \mathrm{M}_n(\mathbb{R})$ lies in $\mathfrak{g}$ if and only if there exists a smooth curve $\gamma\colon \mathbb{R} \to \mathrm{M}_n(\mathbb{R})$ that lies in G and such that $\gamma(0) = I$ and*

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t) = X.$$

In other words, $\mathfrak{g}$ is the tangent space of G at I.

Proof. If X lies in $\mathfrak{g}$ then e^{tX} is a smooth curve that lies in G by definition and whose derivative at 0 is X. Conversely, if γ is a smooth curve in G with $\gamma(0) = I$, then Theorem 3.7.2 implies that for sufficiently small t if we define $\delta(t) := \log\gamma(t)$ then $\delta(t)$ lies in $\mathfrak{g}$ and $\gamma(t) = e^{\delta(t)}$. Moreover, $\delta(t) = t\delta'(0) + O(t^2)$ so that

$$\gamma'(0) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{\delta(t)} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{t\delta'(0)} = \delta'(0).$$

Since $\delta(t)$ lies in $\mathfrak{g}$ for t sufficiently small, so does the limit $\delta'(0) = \lim_{h\to 0}\frac{\delta(h)}{h}$. □

Corollary 3.7.6. *Every continuous homomorphism between matrix groups is smooth.*

Proof. Suppose $\Phi\colon G \to H$ is a continuous homomorphism between matrix groups G, H with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Let $\phi := \mathfrak{L}(\Phi)$. From Theorem 3.7.2 it follows that we can parametrize a neighbourhood of any element $A \in G$ as Ae^X for X in some neighbourhood of 0 in $\mathfrak{g}$. Then in this neighbourhood Φ as given by

$$\Phi(Ae^X) = \Phi(A)e^{\phi(X)},$$

which is clearly smooth since the exponential and matrix multiplication are smooth, and ϕ is smooth (since it is a linear map between finite-dimensional vector spaces). □

Corollary 3.7.7. *Suppose G is a connected matrix group with Lie algebra $\mathfrak{g}$. Then any element $A \in G$ can be expressed as*

$$A = e^{X_1} \dots e^{X_n}$$

for finitely many $X_1, \dots, X_n \in \mathfrak{g}$.

Proof. Choose $0 < \epsilon < \log 2$ as in Theorem 3.7.2, and take U_ϵ and V_ϵ as defined there. Given $A \in G$, choose a continuous path $\gamma\colon [0,1] \to G$ with $\gamma(0) = I, \gamma(1) = A$. We claim that there then exists some $\delta > 0$ such that if $|t - t'| < \delta$, then we have $\gamma(t)^{-1}\gamma(t') \in V_\epsilon$. Assuming this, choose a positive integer m such that $1/m < \delta$; then for $j = 1, \dots, m$ $\gamma((j-1)/m)^{-1}\gamma(j/m) \in V_\epsilon$ so that $\gamma((j-1)/m)^{-1}\gamma(j/m) = e^{X_j}$ for some $X_j \in \mathfrak{g}$. But then we have $\gamma(j/m) = \gamma((j-1)/m)e^{X_j}$ for all j, and hence $\gamma(j/m) = \gamma(0)e^{X_1} \cdots e^{X_j} = e^{X_1} \cdots e^{X_j}$. In particular,

$$A = \gamma(1) = e^{X_1} \cdots e^{X_m},$$

as required.

We now prove the deferred claim. First recall that V_ϵ is an open neighbourhood of I in $\mathrm{M}_n(\mathbb{C})$, so we can choose $\epsilon' > 0$ so that the open ball of radius ϵ' is contained in V_ϵ. It then suffices to show that we can choose $\delta > 0$ so that $\|\gamma(t)^{-1}\gamma(t') - I\| < \epsilon'$ when $|t - t'| < \delta$. First note that

$$\|\gamma(t)^{-1}\gamma(t') - I\| = \|\gamma(t)^{-1}(\gamma(t') - \gamma(t))\| \le \|\gamma(t)^{-1}\|\|\gamma(t') - \gamma(t)\|.$$

Since $[0,1]$ is compact, the continuous function $\|\gamma(t)^{-1}\|$ is bounded, say $\|\gamma(t)^{-1}\| \le C$ for some $C > 0$. Moreover, again since $[0,1]$ is compact, the function $\gamma(t)$ is uniformly continuous. This means there exists $\delta > 0$ such that if $|t - t'| < \delta$ we have

$$\|\gamma(t') - \gamma(t)\| < \epsilon'/C.$$

Putting these estimates together then gives what we want. □

Corollary 3.7.8. *Suppose G is a connected matrix group and H is an arbitrary matrix group. Suppose $\Phi, \Psi\colon G \to H$ are continuous homomorphisms such that $\mathfrak{L}(\Phi) = \mathfrak{L}(\Psi)$. Then $\Phi = \Psi$.*

Proof. Let $\phi := \mathfrak{L}(\Phi)$ and $\psi := \mathfrak{L}(\Psi)$. By Corollary 3.7.7 if A is any element of G we have an expression $A = e^{X_1} \cdots e^{X_n}$ with $X_i \in \mathfrak{L}(G)$. Then we get

$$\begin{aligned}
\Phi(A) &= \Phi(e^{X_1} \cdots e^{X_n}) \\
&= \Phi(e^{X_1}) \cdots \Phi(e^{X_n}) \\
&= e^{\phi(X_1)} \cdots e^{\phi(X_n)} \\
&= e^{\psi(X_1)} \cdots e^{\psi(X_n)} \\
&= \Psi(e^{X_1}) \cdots \Psi(e^{X_n}) \\
&= \Psi(e^{X_1} \cdots e^{X_n}) \\
&= \Psi(A).
\end{aligned}$$

Since A was arbitrary, this completes the proof. □

Corollary 3.7.9. *Suppose G is a connected matrix group whose Lie algebra $\mathfrak{g}$ is commutative. Then G is also commutative.*

Proof. If $\mathfrak{g}$ is commutative, then for any $X, Y \in \mathfrak{g}$ the exponentials e^X and e^Y also commute. Since any element of G is a product of such exponentials by Corollary 3.7.7, the multiplication in G must also be commutative. □

Warning 3.7.10. Note that Corollary 3.7.9 can fail if G is not assumed to be connected. For example, the Lie algebra of any finite group is 0, and so abelian, but there certainly exist non-abelian finite groups!

Chapter 4

Lie theory for matrix groups

4.1 The Baker–Campbell–Hausdorff formula

Our aim in this chapter is to "lift" information from Lie algebras to matrix groups. In particular, we will investigate when we can lift homomorphisms of Lie algebras to homomorphisms of matrix groups, and Lie subalgebras to subgroups. We have already seen that the exponential map $\mathfrak{g} \to G$ is an isomorphism in a neighbourhood of the identity; the key tool we need to proceed further is a description of the products of two exponentials $e^X e^Y$ in terms of Lie brackets, when X and Y are sufficiently small. This is what the *Baker–Campbell–Hausdorff* formula provides: it says that for sufficiently small $X, Y \in \mathrm{M}_n(\mathbb{C})$, we have

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

where all the terms on the right-hand side are expressed entirely in terms of iterated commutators.

Definition 4.1.1. Let us say that a polynomial over $\mathbb{R}$ in n *non-commuting* variables $x_1, \ldots, x_n$ is a *Lie polynomial* if it is a linear combination of iterated commutators in the variables x_i. In other words, the Lie polynomials form the $\mathbb{R}$-vector subspace $\mathsf{LiePoly}_n$ of $\mathbb{R}\{x_1, \ldots, x_n\}$ spanned by x_i, $[x_i, x_j]$, $[x_i, [x_j, x_k]]$, …. It will be convenient to write

$$P(x_1, \ldots, x_n) \equiv_{\mathsf{Lie}} Q(x_1, \ldots, x_n)$$

if $P - Q$ is a Lie polynomial. Note also that Lie polynomials are closed under substition: if $P_1, \ldots, P_n, Q$ are Lie polynomials, then so is $Q(P_1, \ldots, P_n)$.

Theorem 4.1.2. *If we write*

$$F(X, Y) := \log(e^X e^Y) = \sum_{i=1}^{\infty} F_i(X, Y),$$

where $F_i(X,Y)$ is the sum of the degree-i terms in the power series, then each $F_i(X,Y)$ is a Lie polynomial.

We start with some preliminary observations:

Observation 4.1.3. It follows from the definition of log and exp as power series that $F_1(X,Y) = X + Y$.

Observation 4.1.4. The associativity relation $e^X(e^Y e^Z) = (e^X e^Y)e^Z$ implies

$$e^{F(X,F(Y,Z))} = e^X e^{F(Y,Z)} = e^X(e^Y e^Z) = (e^X e^Y)e^Z = e^{F(X,Y)}e^Z = e^{F(X,F(Y,Z))},$$

so we have

$$F(X,F(Y,Z)) = F(F(X,Y),Z). \tag{4.1}$$

Lemma 4.1.5. *Suppose we know F_i is a Lie polynomial for all $i < n$. Then we have*

$$F_n(X, Y+Z) + F_n(Y,Z) \equiv_{\mathsf{Lie}} F_n(X,Y) + F_n(X+Y,Z). \tag{4.2}$$

Proof. It follows from the hypothesis that all terms of degree $< n$ in both sides of (4.1) are Lie polynomials, and the same goes for the terms of degree n that arise from combining F_i and F_j where both $i, j < n$. That leaves only the terms $F_n(Y,Z)$ (from $F_1(X,F(Y,Z))$) and $F_n(X,Y+Z)$ (from $F(X,F_1(Y,Z))$) on the left, and similarly $F_n(X,Y)$ and $F_n(X+Y,Z)$ on the right. Since the homogeneous terms of degree n in (4.1) must agree, this gives the desired congruence. □

Observation 4.1.6. We must have

$$F_n(rX, sX) = 0 \tag{4.3}$$

for $n > 1$ and $r, s \in \mathbb{R}$ since

$$F(rX,sX) = \log(e^{rX}e^{sX}) = \log(e^{(r+s)X}) = rX + sX = F_1(rX,sX).$$

In particular, $F_n(X,0) = 0$. Note also that

$$F_n(rX, rY) = r^n F_n(X,Y) \tag{4.4}$$

for $r \in \mathbb{R}$ since F_n is a homogeneous polynomial of degree n.

Proof of Theorem 4.1.2. We first apply (4.2) with $Z = -Y$ to get

$$F_n(X,0) + F_n(Y,-Y) \equiv_{\mathsf{Lie}} F_n(X,Y) + F_n(X+Y,-Y).$$

Here the left-hand side is 0 by (4.3), so we get

$$F_n(X,Y) \equiv_{\mathsf{Lie}} -F_n(X+Y,-Y). \tag{4.5}$$

If we instead apply (4.2) with $X = -Y$, we get

$$F_n(-Y, Y+Z) + F_n(Y,Z) \equiv_{\mathsf{Lie}} F_n(-Y,Y) + F_n(0,Z) = 0,$$

and so we also get

$$F_n(X,Y) \equiv_{\mathsf{Lie}} -F_n(-X, X+Y). \tag{4.6}$$

Now we can do the following clever manipulation:

$$\begin{aligned}
F_n(X,Y) &\equiv_{\mathsf{Lie}} -F_n(-X, X+Y) && \text{(by (4.6))}\\
&\equiv_{\mathsf{Lie}} F_n(-X+X+Y, -X-Y) && \text{(by (4.5))}\\
&\equiv_{\mathsf{Lie}} F_n(Y, -X-Y) && \\
&\equiv_{\mathsf{Lie}} -F_n(-Y, -X) && \text{(by (4.6))}\\
&\equiv_{\mathsf{Lie}} -(-1)^n F_n(Y, X), && \text{(by (4.4))}
\end{aligned}$$

so that we have

$$F_n(X,Y) \equiv_{\mathsf{Lie}} (-1)^{n+1} F_n(Y,X). \tag{4.7}$$

Now we return to (4.2) and plug in $Z = -Y/2$ to get

$$F_n(X, Y/2) + F_n(Y, -Y/2) \equiv_{\mathsf{Lie}} F_n(X,Y) + F_n(X+Y, -Y/2).$$

Here $F_n(Y, -Y/2) = 0$ by (4.3), and so

$$F_n(X,Y) \equiv_{\mathsf{Lie}} F_n(X, Y/2) - F_n(X+Y, -Y/2). \tag{4.8}$$

Applying (4.2) with $X = -Y/2$ we get

$$F_n(-Y/2, Y+Z) + F_n(Y,Z) \equiv_{\mathsf{Lie}} F_n(-Y/2, Y) + F_n(Y/2, Z),$$

from which we similarly obtain

$$F_n(X,Y) \equiv_{\mathsf{Lie}} F_n(X/2, Y) - F_n(-X/2, X+Y). \tag{4.9}$$

Now we do two more clever manipulations on the terms on the right-hand side of (4.9):

$$\begin{aligned}
F_n(X/2, Y) &\equiv_{\mathsf{Lie}} F_n(X/2, Y/2) - F_n(X/2+Y, -Y/2) && \text{(by (4.8))}\\
&\equiv_{\mathsf{Lie}} F_n(X/2, Y/2) + F_n(X/2+Y/2, Y/2) && \text{(by (4.5))}\\
&\equiv_{\mathsf{Lie}} 2^{-n} F_n(X,Y) + 2^{-n} F_n(X+Y, Y) && \text{(by (4.4))}\\
F_n(-X/2, X+Y) &\equiv_{\mathsf{Lie}} F_n(-X/2, X/2+Y/2) - F_n(X/2+Y, -X/2-Y/2) && \text{(by (4.8))}\\
&\equiv_{\mathsf{Lie}} -F_n(X/2, Y/2) + F_n(Y/2, X/2+Y/2) && \text{(by (4.6) and (4.5))}\\
&\equiv_{\mathsf{Lie}} -2^{-n} F_n(X,Y) + 2^{-n} F_n(Y, X+Y) && \text{(by (4.4))}
\end{aligned}$$

Now we plug this into (4.9) to get

$$F_n(X,Y) \equiv_{\mathsf{Lie}} 2^{-n+1} F_n(X,Y) + 2^{-n} F_n(X+Y, Y) - 2^{-n} F_n(Y, X+Y).$$

Here we can apply (4.7) to get

$$(1 - 2^{-n+1}) F_n(X,Y) \equiv_{\mathsf{Lie}} 2^{-n}(1 + (-1)^n) F_n(X+Y, Y). \tag{4.10}$$

For $n > 1$ odd, this implies that $F_n(X, Y) \equiv_{\mathsf{Lie}} 0$. If n is even, we instead compute

$$\begin{aligned} 2^{-n+1}F_n(X, Y) &\equiv_{\mathsf{Lie}} (1 - 2^{-n+1})F_n(X - Y, Y) && \text{(by (4.10))} \\ &\equiv_{\mathsf{Lie}} -(1 - 2^{-n+1})F_n(X, -Y) && \text{(by (4.5))}, \end{aligned}$$

and so by applying this twice we get

$$F_n(X, Y) \equiv_{\mathsf{Lie}} -\frac{1 - 2^{-n+1}}{2^{-n+1}} F_n(X, -Y) \equiv_{\mathsf{Lie}} \left(\frac{1 - 2^{-n+1}}{2^{-n+1}}\right)^2 F_n(X, Y).$$

This implies that we must also have $F_n(X, Y) \equiv_{\mathsf{Lie}} 0$ for n even. □

Remark 4.1.7. This proof follows the exposition in [4, Section 7.7] of a proof due to Eichler (1968). The fact that there exists *some* expression in terms of commutators is enough for our purposes, but there are also more explicit forms of the Baker–Campbell–Hausdorff formula; see for example the discussion in [2, 5.3–5.5].

4.2 Lifting homomorphisms of Lie algebras

Theorem 4.2.1. *Let G and H be matrix groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, and suppose that $\phi\colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. If G is simply connected, then there exists a unique continuous homomorphism $\Phi\colon G \to H$ such that $\phi = \mathfrak{L}(\Phi)$.*

Definition 4.2.2. Let G and H be matrix groups and suppose that $U \subseteq G$ is a path-connected neighbourhood of the identity. We say that a continuous map $f\colon U \to H$ is a *local homomorphism* if $f(AB) = f(A)f(B)$ whenever A, B, AB all lie in U.

Proposition 4.2.3. *Let G and H be matrix groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, and suppose that $\phi\colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. Define $S_\epsilon \subseteq G$ to be the subset consisting of $A \in G$ such that $\|A - I\| < 1$ and $\log A < \epsilon$. Then there exists some $\epsilon > 0$ such that the map*

$$f := \exp \circ \phi \circ \log\colon S_\epsilon \to H$$

is a local homomorphism.

Proof. Choose $0 < \epsilon < \log 2$ sufficiently small so that Theorem 3.7.2 holds, so that exp restricts to a homeomorphism between S_ϵ and an open neighbourhood U_ϵ of 0 in $\mathfrak{g}$. Moreover, assume that ϵ is sufficiently small so that if $A, B \in S_\epsilon$ then AB is still in the domain of the logarithm. If $X = \log A$ and $Y = \log B$ for $A, B \in S_\epsilon$, we then have

$$f(AB) = \exp(\phi(\log(e^X e^Y))).$$

We can express $\log(e^X e^Y)$ in terms of iterated commutators of X and Y by the Baker–Campbell–Hausdorff formula, and since ϕ preserves Lie brackets this means that we have

$$\phi(\log(e^X e^Y)) = \log(e^{\phi(X)} e^{\phi(Y)}),$$

and so

$$f(AB) = e^{\phi(\log A)} e^{\phi(\log B)} = f(A)f(B),$$

so that f is a local homomorphism. □

Proposition 4.2.4. *Let G and H be matrix groups, and assume that G is simply connected. If $f\colon U \to H$ is a local homomorphism from a path-connected neighbourhood U of the identity in G, then there exists a unique continuous homomorphism $\Phi\colon G \to H$ such that $\Phi|_U = f$.*

Proof. By assumption, G is path-connected, so for any $A \in G$ there exists a path $\alpha\colon [0,1] \to G$ with $\alpha(0) = I$ and $\alpha(1) = A$. The idea is now to define $\Phi(A)$ by subdividing this path and then check that the result is independent of the partition and of the chosen path, so that $\Phi(A)$ is well-defined.

To that end, let us say that a partition

$$\tau = (0 = t_0 < t_1 < \cdots < t_n = 1)$$

of the interval $[0,1]$ is *U-good* for α if we have $\alpha(t)\alpha(t')^{-1} \in U$ whenever $t_i \leq t \leq t' \leq t_{i+1}$ for some i. The argument at the end of the proof of Corollary 3.7.7 shows that a U-good partition for α always exists.

Given a U-good partition τ for α let us set

$$A_i^\tau = \alpha(t_i)\alpha(t_{i-1})^{-1}.$$

Then we can write

$$A = A_n^\tau A_{n-1}^\tau \cdots A_1^\tau,$$

where each factor A_i^τ lies in U by assumption. We then define

$$\Phi(A, \alpha, \tau) := f(A_n^\tau) f(A_{n-1}^\tau) \cdots f(A_1^\tau).$$

Note that if a homomorphism Φ extending f exists then this *must* be the value $\Phi(A)$, which shows that Φ is unique if it exists.

Now we prove that $\Phi(A, \alpha, \tau)$ is independent of the partition τ. First, note that if τ' is obtained from τ by inserting a single element s between t_i and t_{i+1} then τ' is also a U-good partition for α, and we have

$$\Phi(A, \alpha, \tau') = f(A_n^\tau) \cdots f(A_{i+2}^\tau) f(\alpha(t_{i+1})\alpha(s)^{-1}) f(\alpha(s)\alpha(t_i)^{-1}) f(A_i^\tau) \cdots f(A_1^\tau).$$

Since f is a local homomorphism, it satisfies

$$f(\alpha(t_{i+1})\alpha(s)^{-1}) f(\alpha(s)\alpha(t_i)^{-1}) = f(\alpha(t_{i+1})\alpha(t_i)^{-1}) = f(A_{i+1}^\tau),$$

so that $\Phi(A, \alpha, \tau) = \Phi(A, \alpha, \tau')$. Iterating this argument, we get the same identity when τ' is any refinement of τ. Since any two U-good partitions for α have a common refinement (just take their union), this shows that $\Phi(A, \alpha, \tau)$ is independent of the choice of τ. We write $\Phi(A, \alpha)$ for this common value.

The next step is then to show that $\Phi(A, \alpha)$ is independent of the choice of α. Suppose then that $\alpha' \colon [0, 1] \to G$ is another path from I to A. Since G is simply connected, these two paths are homotopic, i.e. there exists a continuous map $h \colon [0, 1] \times [0, 1] \to G$ such that $h(-, 0) = \alpha$, $h(-, 1) = \alpha'$, $h(0, -)$ is constant at I, and $h(1, -)$ is constant at A. Arguing as in the proof of Corollary 3.7.7, we can choose a positive integer N so that $h(s, t)h(s', t')^{-1} \in U$ for $(s, t), (s', t') \in [0, 1] \times [0, 1]$ such that $|s - s'| < 2/N, |t - t'| < 2/N$. Let us set $\alpha_i := h(-, i/N)$ for $i = 0, \ldots, N$; then $\alpha_0 = \alpha$ and $\alpha_N = \alpha'$, so it suffices to show that $\Phi(A, \alpha_i) = \Phi(A, \alpha_{i+1})$ for all i. To this end, define $\gamma_{i,j} \colon [0, 1] \to [0, 1] \times [0, 1]$ by

$$\gamma_{i,j}(t) := \begin{cases} (t, i/N), & t \leq (j-1)/N, \\ (t, t + (i - j + 1)/N), & (j-1)/N < t < j/N, \\ (t, (i+1)/N), & t \geq j/N \end{cases}$$

If we set $\alpha_{i,j} := h \circ \gamma_{i,j}$ then $\alpha_{i,0} = \alpha_{i+1}$ and $\alpha_{i,N+1} = \alpha_i$; thus it's enough to show $\Phi(A, \alpha_{i,j}) = \Phi(A, \alpha_{i,j+1})$. To do this, let us set $A_k^{i,j} = \alpha_{i,j}(k/N)\alpha_{i,j}((k-1)/N)^{-1}$; then we have $A_k^{i,j+1} = A_k^{i,j}$ for $k \leq j - 1$ and $k \geq j + 2$, so that

$$\Phi(A, \alpha_{i,j+1}) = f(A_N^{i,j}) \cdots f(A_{j+2}^{i,j}) f(A_{j+1}^{i,j+1}) f(A_j^{i,j+1}) f(A_{j-1}^{i,j}) \cdots f(A_1^{i,j}).$$

Now note that here

$$\begin{aligned} A_{j+1}^{i,j+1} A_j^{i,j+1} &= \alpha_{i,j+1}((j+1)/N)\alpha_{i,j+1}(j/N)^{-1}\alpha_{i,j+1}(j/N)\alpha_{i,j+1}((j-1)/N)^{-1} \\ &= \alpha_{i,j+1}((j+1)/N)\alpha_{i,j+1}((j-1)/N)^{-1} \\ &= h((j+1)/N, (i+1)/N)h((j-1)/N, i/N)^{-1}, \end{aligned}$$

which by our assumption on N lies in U. Moreover, this product is the same as $A_{j+1}^{i,j} A_j^{i,j}$, so that we get

$$f(A_{j+1}^{i,j+1}) f(A_j^{i,j+1}) = f(h((j+1)/N, (i+1)/N)h((j-1)/N, i/N)^{-1}) = f(A_{j+1}^{i,j}) f(A_j^{i,j}).$$

Substituting this above, we get

$$\Phi(A, \alpha_{i,j+1}) = \Phi(A, \alpha_{i,j}),$$

as required. This shows $\Phi(A, \alpha) = \Phi(A, \alpha')$; we write $\Phi(A)$ for this common value.

Next, let's check that Φ does indeed restrict to f on U. If $A \in U$, we can choose a path $\alpha(t)$ from I to A that lies entirely in U, since U was by assumption path-connected. If τ is a U-good partition for α, then for every j the partition

$0 = t_0 < \cdots < t_j$ is (after reparametrizing) a U-good partition for $\alpha|_{[0,t_j]}$, so that we have

$$\Phi(\alpha(t_j)) = f(A_j^\tau) \cdots f(A_1^\tau).$$

In particular, $\Phi(\alpha(t_1)) = f(A_1^\tau) = f(\alpha(t_1))$. We now show that we have $\Phi(\alpha(t_j)) = f(\alpha(t_j))$ by induction on j. Indeed, we have

$$\begin{aligned}\Phi(\alpha(t_j)) &= f(A_j^\tau) f(A_{j-1}^\tau) \cdots f(A_1^\tau)\\ &= f(A_j^\tau)\Phi(\alpha(t_{j-1}))\\ &= f(A_j^\tau) f(\alpha(t_{j-1}))\\ &= f(\alpha(t_j)),\end{aligned}$$

where the last equality uses that f was a local homomorphism and $A_j^\tau \alpha(t_{j-1}) = \alpha(t_j)$ lies in U.

Now we check that Φ is a homomorphism. Firstly, we certainly have $\Phi(I) = f(I) = I$. Given $A, B \in G$, choose paths α from I to A and β from I to B. Then define $\gamma\colon [0,2] \to G$ by

$$\gamma(t) := \begin{cases} \alpha(t), & t \leq 1,\\ \beta(t-1)\cdot A, & t \geq 1;\end{cases}$$

this is a path from I to BA. If $\tau = (0 = t_0 < \cdots < t_n = 1)$ is a U-good partition of $[0,1]$ for α and $\tau' = (0 = t_0' < \cdots < t_m' = 1)$ is a U-good partition of $[0,1]$ for β, then

$$\tau'' := \tau \cup (1+\tau') = (0 = t_0'' = t_0 < \cdots < t_n'' = t_n = 1 = 1+t_0' < \cdots < t_{n+m}'' = 1+t_m' = 2)$$

is a U-good partition of $[0,2]$ for γ. Then we get

$$(BA)_i^{\tau''} = \gamma(t_i'')\gamma(t_{i-1}'')^{-1} = \begin{cases} \alpha(t_i)\alpha(t_{i-1})^{-1}, & i \leq n,\\ \beta(t_{i-n}')\beta(t_{i-n-1}')^{-1}, & i > n,\end{cases}$$

since $\gamma(t_{n+i}'')\gamma(t_{n+i-1}'')^{-1} = \beta(t_i')AA^{-1}\beta(t_{i-1}')^{-1}$. Hence we have

$$\begin{aligned}\Phi(BA) &= f((BA)_{n+m}^{\tau''}) \cdots f((BA)_1^{\tau''})\\ &= f(B_m^{\tau'}) \cdots f(B_1^{\tau'}) f(A_n^\tau) \cdots f(A_1^\tau)\\ &= \Phi(B)\Phi(A).\end{aligned}$$

Finally, to see that Φ is continuous, it suffices to show it is continuous on an open neighbourhood of every $A \in G$, which we can take to be $A \cdot U$. For $B \in A \cdot U$ we have $\Phi(B) = \Phi(A \cdot A^{-1}B) = \Phi(A)\Phi(A^{-1}B) = \Phi(A) f(A^{-1}B)$, i.e. $\Phi|_{A\cdot U} = \Phi(A)\cdot f(A^{-1}\cdot -)$, which is continuous. □

Proof of Theorem 4.2.1. By Proposition 4.2.3, we can lift ϕ to a local homomorphism $f: U \to H$, and then by Proposition 4.2.4 we can extend f to a unique homomorphism $\Phi: G \to H$. By definition, we then have $\Phi(e^X) = e^{\phi(X)}$ for X in a neighbourhood V of 0 in $\mathfrak{g}$. For an arbitrary $X \in \mathfrak{g}$ we have $X/m \in V$ for $m \in \mathbb{N}$ sufficiently large, and then

$$\Phi(e^X) = \Phi(e^{X/m})^m = (e^{\phi(X/m)})^m = e^{\phi(X)}.$$

Thus we get

$$\mathfrak{L}(\Phi)(X) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi(e^{tX}) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{t\phi(X)} = \phi(X).$$

Uniqueness follows from Corollary 3.7.8. □

Exercise 4.1. Suppose G and G' are simply connected matrix group such that the Lie algebras $\mathfrak{L}(G)$ and $\mathfrak{L}(G')$ are isomorphic. Show that then the matrix groups G and G' are also isomorphic.

Definition 4.2.5. If G is a connected matrix group, then a *universal cover* of G is a simply connected matrix group $\widetilde{G}$ with a continuous homomorphism $\Phi: \widetilde{G} \to G$ such that $\mathfrak{L}(\Phi): \mathfrak{L}(\widetilde{G}) \to G$ is an isomorphism.

Example 4.2.6. SU_2 is a universal cover of $SO_3(\mathbb{R})$ via the homomorphism r from Construction 2.3.4 by Example 3.6.9.

Exercise 4.2. Show that if $\Phi: \widetilde{G} \to G$ and $\Phi': \widetilde{G'} \to G$ are both universal covers of G, then there exists a unique isomorphism $\Psi: \widetilde{G} \xrightarrow{\sim} \widetilde{G'}$ such that $\Phi' \circ \Psi = \Phi$.

One can show that the universal cover (in the topological sense) of a Lie group G is again a Lie group, which has the defining property of the universal cover of G in the preceding sense. However, even if G is a matrix group, its universal cover may not be. Here is an example:

Proposition 4.2.7. *The group* $SL_2(\mathbb{R})$ *does not have a universal cover that is a matrix group.*

Proof. We assume (we have not proved it in the course) that $SL_2(\mathbb{R})$ is not simply connected, but $SL_2(\mathbb{C})$ *is*. The key point is then to observe that, even though $SL_2(\mathbb{R})$ is *not* simply connected, any Lie algebra homomorphism $\phi: \mathfrak{sl}_2(\mathbb{R}) \to \mathfrak{gl}_n(\mathbb{C})$ lifts to a continuous homomorphism $SL_2(\mathbb{R}) \to GL_n(\mathbb{C})$. This is because $SL_2(\mathbb{C})$ *is* simply connected, and so we can lift $\phi_{\mathbb{C}}: \mathfrak{sl}_2(\mathbb{C}) \cong \mathfrak{sl}_2(\mathbb{R}) \otimes \mathbb{C} \to \mathfrak{gl}_n(\mathbb{C})$ uniquely to a continuous homomorphism $\Phi_{\mathbb{C}}: SL_2(\mathbb{C}) \to GL_n(\mathbb{C})$ by Theorem 4.2.1. Restricting $\Phi_{\mathbb{C}}$ to $SL_2(\mathbb{R})$ then gives the required lift of ϕ.

Suppose then that $G \subseteq GL_n(\mathbb{C})$ is a simply connected matrix group, and $\Phi: G \to SL_2(\mathbb{R})$ is a continuous homomorphism such that $\phi := \mathfrak{L}(\Phi): \mathfrak{L}(\mathfrak{g}) \to \mathfrak{sl}_2(\mathbb{R})$ is an isomorphism. Then the Lie algebra homomorphism $\phi^{-1}: \mathfrak{sl}_2(\mathbb{R}) \to \mathfrak{L}(G) \hookrightarrow \mathfrak{gl}_n(\mathbb{C})$ can be lifted to a continuous homomorphism $SL_2(\mathbb{R}) \to$

$GL_n(\mathbb{C})$. This factors through a continuous homomorphism $\Phi' \colon SL_2(\mathbb{R}) \to G$: we have $\Phi'(e^X) = e^{\phi^{-1}(X)} \in G$, and since $SL_2(\mathbb{R})$ is connected all elements are products of exponentials.

But then Proposition 3.6.11 and Corollary 3.7.8 imply that $\Phi \circ \Phi' = \mathrm{id}_{SL_2(\mathbb{R})}$ and $\Phi' \circ \Phi = \mathrm{id}_G$, so that Φ is an isomorphism. But this is impossible since $SL_2(\mathbb{R})$ is *not* simply connected. □

Since universal covers always exist as Lie groups, this gives an example of a Lie group that is not a matrix group:

Example 4.2.8. The universal cover $\widetilde{SL}_2(\mathbb{R})$ of $SL_2(\mathbb{R})$ is a Lie group that is not a matrix group.

4.3 Lifting Lie subalgebras

Any matrix group $G \subseteq GL_n(\mathbb{C})$ gives us a subalgebra $\mathfrak{g} = \mathfrak{L}(G) \subseteq \mathfrak{gl}_n(\mathbb{C})$. We've seen that the Lie algebra $\mathfrak{g}$ controls much of the behaviour of the group G, so it's natural to wonder whether the subalgebra $\mathfrak{g}$ determines G. For this to have any chance of being true we certainly need to assume that G is connected, since the Lie algebra of G is the same as that of the connected component of the identity. We will see that this is in fact the only obstruction: as a subgroup of $GL_n(\mathbb{C})$, G is completely determined by $\mathfrak{g}$ as a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$.

Given this, we might then ask which subalgebras of $\mathfrak{gl}_n(\mathbb{C})$ correspond to matrix groups. This is not so easy to answer, since there may be no *closed* subgroup that corresponds to a given Lie subalgebra:

Example 4.3.1. Let $\mathfrak{g}$ be the subalgebra of $\mathfrak{gl}_2(\mathbb{C})$ given by

$$\mathfrak{g} = \left\{ \begin{pmatrix} it & 0 \\ 0 & ita \end{pmatrix} : t \in \mathbb{R} \right\},$$

where a is *irrational*. If there existed a closed subgroup G of $GL_2(\mathbb{C})$ with $\mathfrak{g}$ as its Lie algebra, then G certainly has to contain the group

$$H := \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{ita} \end{pmatrix} : t \in \mathbb{R} \right\}$$

for all $t \in \mathbb{R}$, and hence also the *closure* $\overline{H}$ of this set in $GL_2(\mathbb{C})$. But it is not hard to check that

$$\overline{H} := \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{it'} \end{pmatrix} : t, t' \in \mathbb{R} \right\}.$$

This means that the Lie algebra of G must contain all matrices

$$\begin{pmatrix} it & 0 \\ 0 & it' \end{pmatrix}, \qquad t, t' \in \mathbb{R},$$

and so is bigger than $\mathfrak{g}$.

If we are willing to drop the condition that the subgroup must be closed, however, it turns out that there *is* an exact correspondence between subalgebras of $\mathfrak{gl}_n(\mathbb{C})$ and a certain class of subgroups of $\mathrm{GL}_n(\mathbb{C})$ (and similarly for subalgebras in the Lie algebra of any matrix group).

Definition 4.3.2. If G is any subgroup of $\mathrm{GL}_n(\mathbb{C})$, we write $\mathfrak{L}(G)$ for the set of matrices $X \in \mathrm{M}_n(\mathbb{C})$ such that $e^{tX} \in G$ for all $t \in \mathbb{R}$. We say that G is a *connected Lie subgroup* of $\mathrm{GL}_n(\mathbb{C})$ if $\mathfrak{L}(G)$ is a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ (i.e. $\mathfrak{L}(G)$ is a vector subspace closed under commutators) and every element of G can be written in the form $e^{X_1} \cdots e^{X_n}$ for some $X_1, \ldots, X_n \in \mathfrak{L}(G)$. If $H \subseteq \mathrm{GL}_n(\mathbb{C})$ is a matrix group and $G \subseteq H$ is a subgroup that is a connected Lie subgroup of $\mathrm{GL}_n(\mathbb{C})$, then we also say that G is a *connected Lie subgroup of* H.

Remark 4.3.3. Note that a connected Lie subgroup is necessarily path-connected, since an element of the form $e^{X_1} \cdots e^{X_n}$ is connected to I by the path $e^{tX_1} \cdots e^{tX_n}$ which lies in G as $X_1, \ldots, X_n \in \mathfrak{L}(G)$. Note also that any connected matrix group $G \subseteq \mathrm{GL}_n(\mathbb{C})$ is a connected Lie subgroup by Corollary 3.7.7.

Theorem 4.3.4. *If $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$, then there exists a unique connected Lie subgroup $G \subseteq \mathrm{GL}_n(\mathbb{C})$ such that $\mathfrak{g} = \mathfrak{L}(G)$.*

For the proof we need the following technical observation, whose proof (which is where we will use the Baker–Campbell–Hausdorff formula) we defer until the end of this section:

Lemma 4.3.5. *Let $\mathfrak{g}$ be a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$. If B is a basis for $\mathfrak{g}$, we say that an element of $\mathfrak{g}$ is B-*rational *if it is a $\mathbb{Q}$-linear combination of the basis elements. Suppose $A \in \mathrm{GL}_n(\mathbb{C})$ can be written as $e^{X_1} \cdots e^{X_n}$ with $X_i \in \mathfrak{g}$. Then for every $\delta > 0$ there exist B-rational elements $R_1, \ldots, R_m$ of $\mathfrak{g}$ such that we can write*

$$A = e^{R_1} \cdots e^{R_m} e^Y$$

where $Y \in \mathfrak{g}$ and $\|Y\| < \delta$.

Proof of Theorem 4.3.4. Let us define

$$G := \{e^{X_1} \cdots e^{X_n} \in \mathrm{GL}_n(\mathbb{C}) : X_i \in \mathfrak{g}, n \in \mathbb{N}\}.$$

This is a subgroup of $\mathrm{GL}_n(\mathbb{C})$. If we can show that $\mathfrak{L}(G)$ equals $\mathfrak{g}$, then we will know that G is a connected Lie subgroup. Moreover, it must be the *unique* such connected Lie subgroup: if G' is a connected Lie subgroup of $\mathrm{GL}_n(\mathbb{C})$ with $\mathfrak{L}(G') = \mathfrak{g}$, then in particular $e^X \in G$ for $X \in \mathfrak{g}$, and hence $G \subseteq G'$ sine G' is a subgroup. On the other hand, every element of G' is a product of such exponentials and so is also contained in G, so that $G = G'$.

If $X \in \mathfrak{g}$, then $tX \in \mathfrak{g}$ for all t, and so by definition we have we have $e^{tX} \in G$ for all t; thus $\mathfrak{g} \subseteq \mathfrak{L}(G)$. To show the converse inclusion, we proceed as in the proof of Theorem 3.7.2: thinking of $\mathfrak{gl}_n(\mathbb{C})$ as $\mathbb{R}^{2n^2}$, we let $\mathfrak{g}^\perp$ be the orthogonal complement of $\mathfrak{g}$; then there exist neighbourhoods U and V of 0 in $\mathfrak{g}$ and $\mathfrak{g}^\perp$, respectively, and a neighbourhood W of I in $\mathrm{GL}_n(\mathbb{C})$, such that every $A \in W$ can be written uniquely as

$$A = e^X e^Y, \qquad X \in U \subseteq \mathfrak{g}, \quad Y \in V \subseteq \mathfrak{g}^\perp,$$

in such a way that X and Y depend continuously on A. If we have $Z \in \mathfrak{L}(G)$, this means that for t sufficiently small we can write

$$e^{tZ} = e^{X(t)} e^{Y(t)}$$

with $X(t) \in U \subseteq \mathfrak{g}$ and $Y(t) \in V$, with $X(t)$ and $Y(t)$ being continuous in t. Here by assumption $e^{tZ} \in G$ for all t, and $e^{X(t)} \in G$ since $\mathfrak{g} \subseteq \mathfrak{L}(G)$. Thus $e^{Y(t)} = e^{-X(t)} e^{tZ}$ must also lie in G for small t.

Our goal is now to show that $Y(t)$ must be constant. This will complete the proof, since we have $Y(0) = 0$ and so we must have $Y(t) = 0$ for all sufficiently small t, and then $e^{tZ} = e^{X(t)}$ so that $tZ = X(t)$ for small t. But then $tZ \in \mathfrak{g}$ for small t, and hence also $Z \in \mathfrak{g}$ as required.

Now observe that since $Y(t)$ is continuous, if it is not constant it must take on uncountably many values. We will show that the set

$$E := \{Y \in V : e^Y \in G\}$$

is at most countable, so that this is impossible.

Choose $\delta > 0$ so that $C(X, Y) = \log(e^X e^Y)$ is defined and contained in U for $X, Y \in \mathfrak{g}$ with $\|X\|, \|Y\| < \delta$. Pick a basis B for $\mathfrak{g}$; we then claim that for every choice $R_1, \dots, R_m$ of B-rational elements of $\mathfrak{g}$, there is at most one $X \in \mathfrak{g}$ with $\|X\| < \delta$ such that the element $e^{R_1} \cdots e^{R_m} e^X$ belongs to $\exp(V)$. Indeed, if

$$e^{R_1} \cdots e^{R_m} e^X = e^Y, \qquad e^{R_1} \cdots e^{R_m} e^{X'} = e^{Y'}$$

with $Y, Y' \in V$, then

$$e^{-Y} e^{Y'} = e^{-X} e^{X'},$$

so that

$$e^{-Y} = e^{-X} e^{X'} e^{-Y'} = e^{C(-X, X')} e^{-Y'}$$

with $C(-X, X') \in U$. But then the *uniqueness* of the representation of this element in terms of U and V implies that $Y = Y'$, and hence also $X = X'$. Now we use that by Lemma 4.3.5, any element of G can be expressed in the form $e^{R_1} \cdots e^{R_m} e^X$ with R_i B-rational and $\|X\| < \delta$. Since there are only countably many lists $(R_1, \dots, R_n)$ of B-rational elements, and each produces at most one element $e^{R_1} \cdots e^{R_m} e^X$ with $\|X\| < \delta$ that lies in $\exp(V)$, we see that E must have at most countably many elements. □

Corollary 4.3.6. *Let $G \subseteq \mathrm{GL}_n(\mathbb{C})$ be a matrix group with Lie algebra $\mathfrak{g}$. If $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$, then there exists a unique connected Lie subgroup $H \subseteq G$ such that $\mathfrak{h} = \mathfrak{L}(H)$.*

Proof. If we regard $\mathfrak{h}$ as a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$, then there exists a unique connected Lie subgroup $H \subseteq \mathrm{GL}_n(\mathbb{C})$ such that $\mathfrak{h} = \mathfrak{L}(H)$ by Theorem 4.3.4. Moreover, H consists of products of the elements e^X for $X \in \mathfrak{h}$, which lie in G since $\mathfrak{h}$ is by assumption a subalgebra of $\mathfrak{g}$, so that $H \subseteq G$ as required. □

Remark 4.3.7. Suppose $G \subseteq \mathrm{GL}_n(\mathbb{C})$ is a connected Lie subgroup with Lie algebra $\mathfrak{g}$. If G is not closed in $\mathrm{GL}_n(\mathbb{C})$, then G may well *not* be a manifold with respect to the subspace topology from $\mathrm{GL}_n(\mathbb{C})$. However, we can define a better-behaved topology on G as follows: For $A \in G$ and $\epsilon > 0$, let $U(A, \epsilon)$ be the subset of G consisting of elements of the form Ae^X where $X \in \mathfrak{g}$ and $\|X\| < \epsilon$; we then say a subset $U \subseteq G$ is open in the new topology if it contains a neighbourhood of the form $U(A, \epsilon)$ for every $A \in U$. This topology is finer in the subspace topology, and it is not hard to show (just as we did for closed subgroups before) that with this topology G has a natural structure of a smooth manifold such that the group operations are smooth, i.e. G is a Lie group. (See [2, Theorem 5.23] for more details.)

We can ask, more generally, which finite-dimensional abstract Lie algebras $\mathfrak{g}$ are the Lie algebras of matrix groups, or more generally of Lie groups. This in fact turns out to be all of them, a result often known as *Lie's Third Theorem*. We will not prove this here, but we can show a weaker result in this direction if we assume the following algebraic result:

Fact 4.3.8 (Ado's Theorem)**.** *Every finite-dimensional Lie algebra (over $\mathbb{R}$ or $\mathbb{C}$) is isomorphic to a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ for some n.*

Combining this with Theorem 4.3.4, we have:

Corollary 4.3.9. *If $\mathfrak{g}$ is a finite-dimensional $\mathbb{R}$-Lie algebra, then there exists a connected Lie subgroup G of $\mathrm{GL}_n(\mathbb{C})$ for some n such that the Lie algebra $\mathfrak{L}(G)$ is isomorphic to $\mathfrak{g}$.*

Remark 4.3.10. It can be shown that if G is a connected Lie subgroup of $\mathrm{GL}_n(\mathbb{C})$, then G (with its improved topology) can always be embedded as a *closed* subgroup of some $\mathrm{GL}_{n'}(\mathbb{C})$. This means that every finite-dimensional Lie algebra over $\mathbb{R}$ can be realized as the Lie algebra of some matrix group.

Finally, let us come back to our technical lemma:

Proof of Lemma 4.3.5. Choose $\epsilon > 0$ sufficiently small so that for $X, Y \in \mathfrak{g}$ with $\|X\|, \|Y\| < \epsilon$, the function $C(X, Y) := \log(e^X e^Y)$ is defined. Then $C(X, Y)$ is continuous in X and Y. Without loss of generality we may assume that $\delta < \epsilon$ and that for $\|X\|, \|Y\| < \delta$ we have $\|C(X, Y)\| < \epsilon$. If $A = e^{X_1} \cdots e^{X_n}$ with X_i in $\mathfrak{g}$ then we can also write $A = e^{Y_1} \cdots e^{Y_m}$ where $Y_i \in \mathfrak{g}$ satisfies $\|Y_i\| < \delta$ — just replace e^{X_i} by $(e^{X_i/k})^k$ for k sufficiently large. We can now proceed by induction on m in such an expression.

If $m = 0$, then $A = I = e^0$ and there is nothing to prove. Assuming the lemma is proved for expressions of length $< m$ and $A = e^{Y_1} \cdots e^{Y_m}$, we can apply the lemma to $e^{Y_1} \cdots e^{Y_{m-1}}$ to obtain

$$\begin{aligned} A &= e^{R_1} \cdots e^{R_k} e^X e^{Y_m} \\ &= e^{R_1} \cdots e^{R_k} e^{C(X,Y_m)}, \end{aligned}$$

where the R_i's are B-rational and $\|C(X, Y_m)\| < \epsilon$. Now we apply the Baker–Campbell–Hausdorff formula to conclude that since $\mathfrak{g}$ is a Lie subalgebra, $C(X, Y_m)$ must also lie in $\mathfrak{g}$, being an infinite sum of terms given by iterated brackets of X and Y_m. (Note that here we crucially use that $\mathfrak{g}$ is a Lie subalgebra, and that this is the only point where we apply the Baker–Campbell–Hausdorff formula in this section!)

Now we can choose a B-rational element R_{k+1} close to $C(X, Y_m)$ and such that $\|R_{k+1}\| < \epsilon$. Then we have

$$\begin{aligned} A &= e^{R_1} \cdots e^{R_k} e^{R_{k+1}} e^{-R_{k+1}} e^{C(X,Y_m)} \\ &= e^{R_1} \cdots e^{R_k} e^{R_{k+1}} e^{X'}, \end{aligned}$$

where $X' = C(-R_{k+1}, C(X, Y_m))$. Then we again have that $X' \in \mathfrak{g}$, and we claim that if we choose R_{k+1} sufficiently close to $C(X, Y_m)$ then we will have $\|X'\| < \delta$. Indeed, since $C(-Z, Z) = \log(e^{-Z}e^Z) = 0$ for any Z, continuity implies that if Z' is sufficiently close to Z', then $C(-Z', Z)$ can be made arbitrarily small. □

Part II
Representation theory

Chapter 5

Basic representation theory

5.1 Representations of Lie groups and Lie algebras

Notation 5.1.1. If V is a finite-dimensional $\mathbb{K}$-vector space (for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$), we write $\mathrm{GL}(V)$ for the group of invertible linear endomorphisms of V; we have $\mathrm{GL}(V) \cong \mathrm{GL}_n(\mathbb{K})$ if $n = \dim_{\mathbb{K}} V$, so that $\mathrm{GL}(V)$ is a matrix group. Similarly, we write $\mathfrak{gl}(V)$ for the Lie algebra of all $\mathbb{K}$-linear endomorphisms of V, which we can identify with $\mathfrak{gl}_n(\mathbb{K})$; thus $\mathfrak{gl}(V)$ is the Lie algebra of $\mathrm{GL}(V)$.

Definition 5.1.2. Let G be a matrix group. If V is a finite-dimensional $\mathbb{K}$-vector space, a *representation* of G on V is a continuous homomorphism $\rho\colon G \to \mathrm{GL}(V)$. For $\mathbb{K} = \mathbb{R}$ we say that this is a *real* representation of G and for $\mathbb{K} = \mathbb{C}$ that it is a *complex* representation.

Exercise 5.1. Show that the data of a representation of G on V is equivalent to that of a continuous $\mathbb{K}$-linear *action* of G on V, meaning a continuous map

$$\alpha\colon G \times V \to V$$

such that $\alpha(gg', v) = \alpha(g, \alpha(g', v))$ and $\alpha(g, -)$ is $\mathbb{K}$-linear for all $g \in G$.

Notation 5.1.3. If $\rho\colon G \to \mathrm{GL}(V)$ is a representation of G, we might denote the representation as (V, ρ) or just as V when this does not create any confusion. In the latter case we will typically denote the action of $g \in G$ on $v \in V$ as $g \cdot v$ instead of $\rho(g)(v)$.

Definition 5.1.4. Let $\mathfrak{g}$ be an $\mathbb{R}$-Lie algebra. If V is a finite-dimensional $\mathbb{K}$-vector space, a *representation* of $\mathfrak{g}$ on V is an $\mathbb{R}$-Lie algebra homomorphism $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$. For $\mathbb{K} = \mathbb{R}$ we say that this is a *real* representation of $\mathfrak{g}$ and for $\mathbb{K} = \mathbb{C}$ that it is a *complex* representation.

Exercise 5.2. Show that the data of a real representation of a Lie algebra $\mathfrak{g}$ amounts to an $\mathbb{R}$-bilinear map $\beta\colon \mathfrak{g} \times V \to V$ which satisfies

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

Moreover, if V is a $\mathbb{C}$-vector space, then asking for this to be a complex representation amounts to $\beta(X, -)$ being $\mathbb{C}$-linear for all $X \in \mathfrak{g}$.

Definition 5.1.5. Let $\mathfrak{g}$ be a $\mathbb{C}$-Lie algebra. If V is a finite-dimensional $\mathbb{C}$-vector space, a (complex) *representation* of $\mathfrak{g}$ on V is a $\mathbb{C}$-Lie algebra homomorphism $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$.

Remark 5.1.6. From Proposition 3.4.8 we see that if $\mathfrak{g}$ is an $\mathbb{R}$-Lie algebra and V is a $\mathbb{C}$-vector space, then a complex representation of $\mathfrak{g}$ on V is equivalent to a representation of the $\mathbb{C}$-Lie algebra $\mathfrak{g} \otimes \mathbb{C}$ on V.

Remark 5.1.7. From Theorem 3.6.1 we see that a (real or complex) representation ρ of a matrix group G on V induces a (real or complex) representation $\mathfrak{L}(\rho)$ of the Lie algebra $\mathfrak{L}(G)$ on the same vector space V. Moreover, if G is simply connected then Theorem 4.2.1 implies that a representation of G is uniquely determined by the corresponding representation of $\mathfrak{g}$. Combining this with the previous remark, we see that *complex* representations of a simply connected matrix group G are equivalent to representations of the $\mathbb{C}$-Lie algebra $\mathfrak{L}(G) \otimes \mathbb{C}$.

Definition 5.1.8. If V and W are $\mathbb{K}$-vector spaces and $\rho\colon G \to \mathrm{GL}(V)$, $\rho'\colon G \to \mathrm{GL}(W)$ are $\mathbb{K}$-representations of the same matrix group G, then an *intertwining map*[1] from (V, ρ) to (W, ρ') is a $\mathbb{K}$-linear map $\phi\colon V \to W$ such that for $g \in G$ we have

$$\phi(\rho(g)v) = \rho'(g)(\phi(v)).$$

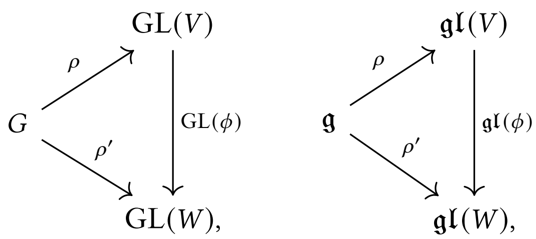
We say that an intertwining map is an *isomorphism* of representations if the underlying $\mathbb{K}$-linear map is an isomorphism of vector spaces.

Definition 5.1.9. If V and W are $\mathbb{K}$-vector spaces and $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$, $\rho'\colon \mathfrak{g} \to \mathfrak{gl}(W)$ are $\mathbb{K}$-representations of the same Lie algebra $\mathfrak{g}$, then an *intertwining map* from (V, ρ) to (W, ρ') is a $\mathbb{K}$-linear map $\phi\colon V \to W$ such that for $X \in \mathfrak{g}$ we have

$$\phi(\rho(X)v) = \rho'(X)(\phi(v)).$$

We say that an intertwining map is an *isomorphism* of representations if the underlying $\mathbb{K}$-linear map is an isomorphism of vector spaces.

Remark 5.1.10. An isomorphism of representations of a group G or a Lie algebra $\mathfrak{g}$ can also be described as a commutative triangle



where the vertical map comes from an isomorphism $\phi\colon V \xrightarrow{\sim} W$.

[1] Also called an *equivariant* map, at least by algebraic topologists.

Exercise 5.3. Show that an isomorphism between the representations (V, ρ) and (W, ρ') of G or $\mathfrak{g}$ can equivalently be defined as a pair of intertwining maps $\phi\colon (V, \rho) \to (W, \rho')$, $\psi\colon (W, \rho') \to (V, \rho)$ such that $\phi\psi = \mathrm{id}_W$, $\psi\phi = \mathrm{id}_V$.

Proposition 5.1.11. *Suppose G is a connected matrix group with Lie algebra $\mathfrak{g}$. Then two representations $\rho\colon G \to \mathrm{GL}(V)$ and $\rho'\colon G \to \mathrm{GL}(W)$ are isomorphic if and only if the induced Lie algebra representations $\mathfrak{L}(\rho)$ and $\mathfrak{L}(\rho')$ are isomorphic.*

Proof. From the formulation of isomorphisms in Remark 5.1.10 the "only if" direction is obvious from Proposition 3.6.11. Conversely, suppose $\phi\colon V \to W$ is an isomorphism of Lie algebra representations. This means that ϕ is an isomorphism of vector spaces that is an intertwining map, which means $\mathfrak{gl}(\phi) \circ \mathfrak{L}(\rho) = \mathfrak{L}(\rho')$. But from Proposition 3.6.11 the left-hand side is $\mathfrak{L}(\mathrm{GL}(\phi) \circ \rho)$. Since G is connected this identity between Lie algebra homomorphisms then implies $\mathrm{GL}(\phi) \circ \rho = \rho'$ by Corollary 3.7.8, i.e. ϕ is an isomorphism between ρ and ρ'. □

Examples 5.1.12. Let's give a few initial examples of representations:

(i) For any vector space V and matrix group G, we always have the *trivial representation* where $g \cdot v = v$ for all $g \in G$, $v \in V$. As a group homomorphism, this is the composite
$$G \to 1 \to \mathrm{GL}(V)$$
through the trivial group. This corresponds to the trivial Lie algebra representation $\mathfrak{g} \to 0 \to \mathfrak{gl}(V)$ where $X \cdot v = 0$ for $X \in \mathfrak{g}, v \in V$.

(ii) If $(G, j\colon G \hookrightarrow \mathrm{GL}_n(\mathbb{R}))$ is an embedded matrix group, then we call the representation of G corresponding to the inclusion j the *standard representation* of G (and similarly if we instead embed G in $\mathrm{GL}_n(\mathbb{C})$). Less formally, if we define the matrix group G as a subgroup of $\mathrm{GL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{C})$, then the standard representation is the corresponding matrix action on $\mathbb{R}^n$ or $\mathbb{C}^n$. This corresponds to the standard representation of the Lie algebra $\mathfrak{g} \subseteq \mathfrak{gl}_n(\mathbb{R})$ or $\mathfrak{gl}_n(\mathbb{C})$ acting on $\mathbb{R}^n$ or $\mathbb{C}^n$.

(iii) For any matrix group G, we have the *adjoint representation* $\mathrm{Ad}_G\colon G \to \mathrm{GL}(\mathfrak{g})$ of G on its Lie algebra, given by $\mathrm{Ad}_G(A)(X) = A^{-1}XA$. The corresponding Lie algebra representation is $\mathrm{ad}_{\mathfrak{g}}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$, given by $\mathrm{ad}_{\mathfrak{g}}(X)(Y) = [X, Y]$.

Finally, we mention an important special class of representations:

Definition 5.1.13. Let $(V, \langle -, - \rangle)$ be a finite-dimensional *complex* inner product space; we can then define the subgroup $\mathrm{U}(V) \subseteq \mathrm{GL}(V)$ of *unitary* automorphisms of V, i.e. those automorphisms ϕ such that $\langle \phi v, \phi w \rangle = \langle v, w \rangle$ for all $v, w \in V$. (Of course, this is isomorphic to U_n if we identify V with $\mathbb{C}^n$ with its standard inner product by picking an orthonormal basis.) We then say a representation $\rho\colon G \to \mathrm{GL}(V)$ is *unitary* if it factors through the subgroup $\mathrm{U}(V)$, that is if the automorphism $\rho(g)\colon V \xrightarrow{\sim} V$ is unitary for all $g \in G$.

Definition 5.1.14. For V as above, the Lie algebra $\mathfrak{u}(V)$ of $\mathrm{U}(V)$ consists of *skew-Hermitian* (or *skew-self-adjoint*) endomorphisms of V, meaning $\phi\colon V \to V$ such that $\langle \phi(v), w\rangle = -\langle v, \phi(w)\rangle$ for $v, w \in V$. We therefore say that a Lie algebra representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is *unitary* if ρ factors through the subalgebra $\mathfrak{u}(V)$, that is if the endomorphism $\rho(X)$ is skew-Hermitian for all $X \in \mathfrak{g}$.

Proposition 5.1.15. *Let $(V, \langle -, -\rangle)$ be a finite-dimensional complex inner product space, suppose G is a matrix group with Lie algebra $\mathfrak{g}$, and let $\rho\colon G \to \mathrm{GL}(V)$ be a representation. If ρ is unitary, then $\mathfrak{L}(\rho)$ is a unitary represenation of $\mathfrak{g}$. If G is connected, then we have conversely that if $\mathfrak{L}(\rho)$ is unitary, then so is ρ.*

Proof. If ρ is unitary, i.e. factors through $\mathrm{U}(V)$, then $\mathfrak{L}(\rho)$ will factor through the corresponding Lie subalgebra $\mathfrak{u}(V)$, so the “if” direction is clear. Conversely, if $\mathfrak{L}(\rho)(X)$ is skew-Hermitian, then $\rho(e^X) = e^{\mathfrak{L}(\rho)(X)}$ is unitary. If G is connected, then every element of G is a product of such exponentials by Corollary 3.7.7 and so $\rho(A)$ must be unitary for all $A \in G$. □

5.2 Irreducible representations and Schur’s Lemma

In this section we will define *irreducible* representations, which are the smallest pieces we might hope to break representations up into, and prove a key property thereof, known as Schur’s Lemma.

Definition 5.2.1. Let V be a $\mathbb{K}$-representation of a matrix group G. We say that a $\mathbb{K}$-vector subspace W of V is *invariant* if for every $g \in G$ and $w \in W$, the vector $g \cdot w$ is also in W. Similarly, if V is a representation of a Lie algebra $\mathfrak{g}$ we say that a subspace W is *invariant* if $X(w)$ is in W for all $X \in \mathfrak{g}$ and $w \in W$. An invariant subspace W of V is called *trivial* if $W = 0$ or $W = V$, and otherwise *non-trivial*.

Definition 5.2.2. A representation V of a matrix group G or a Lie algebra $\mathfrak{g}$ is *irreducible* if V has *no* non-trivial invariant subspaces.

Proposition 5.2.3. *Suppose G is a connected matrix group with Lie algebra $\mathfrak{g}$, and let (V, ρ) be a representation of G. Then a subspace $W \subseteq V$ is G-invariant if and only if W is $\mathfrak{g}$-invariant for the induced representation $\mathfrak{L}(\rho)$.*

Proof. First suppose W is a subspace of V that is invariant under $\mathfrak{L}(\rho)(X)$ for all $X \in \mathfrak{g}$. Then for $w \in W$ we have that

$$\rho(e^X)(w) = e^{\mathfrak{L}(\rho)(X)}(w) = \sum_n \frac{1}{n!}\mathfrak{L}(\rho)(X)^n(w),$$

which must also be in W since this is a closed subspace of V and each term in the series lies in W. Thus W is invariant under $\rho(e^X)$ for all $X \in \mathfrak{g}$. But since G is connected, Corollary 3.7.7 tells us that every element of G is a product of such exponentials, so that W must be invariant under the action of G.

For the converse, suppose W is an invariant subspace of V for ρ. Then for $X \in \mathfrak{g}$ and $w \in W$ we have that $\rho(e^{tX})(w)$ is in W for all t. Since W is closed, this meansp

$$\mathfrak{L}(\rho)(X)(w) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho(e^{tX})(w) = \lim_{t \to 0} \frac{\rho(e^{tX})(w) - w}{t}$$

also lies in W. Thus W is an invariant subspace for $\mathfrak{L}(\rho)$. □

Corollary 5.2.4. *Suppose G is a connected matrix group with Lie algebra $\mathfrak{g}$. Then a representation (V, ρ) of G is irreducible if and only if the induced representation $(V, \mathfrak{L}(\rho))$ of $\mathfrak{g}$ is irreducible.* □

Proposition 5.2.5. *Suppose $\mathfrak{g}$ is an $\mathbb{R}$-Lie algebra and $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a complex representation of $\mathfrak{g}$. Then a $\mathbb{C}$-subspace $W \subseteq V$ is $\mathfrak{g}$-invariant if and only if it is invariant for the induced representation $\rho_{\mathbb{C}}\colon \mathfrak{g} \otimes \mathbb{C} \to \mathfrak{gl}(V)$ of $\mathfrak{g} \otimes \mathbb{C}$.*

Proof. For a $\mathbb{C}$-vector subspace W of V to be invariant for $\rho_{\mathbb{C}}$ means that it is invariant under $\rho_{\mathbb{C}}(X + iY) = \rho(X) + i\rho(Y)$ for all $X, Y \in \mathfrak{g}$. Since W is by assumption invariant under multiplication by i, this holds if and only if W is invariant under $\rho(X)$ and $\rho(Y)$. Hence the $\rho_{\mathbb{C}}$-invariant subspaces of V are precisely the ρ-invariant ones. □

Corollary 5.2.6. *Suppose $\mathfrak{g}$ is an $\mathbb{R}$-Lie algebra and $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a complex representation of $\mathfrak{g}$. Then the induced representation $\rho_{\mathbb{C}}\colon \mathfrak{g} \otimes \mathbb{C} \to \mathfrak{gl}(V)$ is irreducible if and only if ρ is irreducible.* □

Proposition 5.2.7. *Suppose V and W are irreducible representations of a matrix group or a Lie algebra. If $\phi\colon V \to W$ is an intertwining map, then ϕ is either 0 or an isomorphism.*

Proof. We consider the group case; the proof for Lie algebras is the same. If $x \in \ker\phi$ then $\phi(g \cdot x) = g \cdot \phi(x) = 0$ for all g in the group, so $\ker\phi$ is an invariant subspace of V. Hence either $\ker\phi = 0$ or $\ker\phi = V$, i.e. ϕ is either 0 or injective.

On the other hand, if $x \in \operatorname{im}\phi$ so that $x = \phi(y)$, then $g \cdot x = \phi(g \cdot y)$, so $\operatorname{im}\phi$ is an invariant subspace of W. Then $\operatorname{im}\phi = W$ or $\operatorname{im}\phi = 0$, i.e. ϕ is either 0 or surjective. Thus if $\phi \neq 0$ it must be both injective and surjective, that is to say an isomorphism. □

Corollary 5.2.8 (Schur's Lemma)**.** *Suppose V is an irreducible* complex *representation of a matrix group or Lie algebra, and that $\phi\colon V \to V$ is an intertwining map. Then $\phi = \lambda \cdot \mathrm{id}$ for some $\lambda \in \mathbb{C}$.*

Proof. We again consider the group case; the proof in the other case is the same. Since we are working over $\mathbb{C}$, the endomorphism ϕ has an eigenvalue λ. Let $W \subseteq V$ be the eigenspace for λ. Then if $w \in W$ and g is in the group, we have

$$\phi(g \cdot w) = g \cdot \phi(w) = g \cdot \lambda w = \lambda(g \cdot w).$$

Thus $g \cdot w$ is also an eigenvector for λ, and so W is an invariant subspace. By assumption $W \neq 0$, so since V is irreducible we must have $W = V$. This means that $\phi(v) = \lambda v$ for all $v \in V$, i.e. $\phi = \lambda \mathrm{id}$. □

Let us now note a few consequences of Schur's lemma:

Corollary 5.2.9. *Suppose V and W are irreducible* complex *representations of a matrix group or a Lie algebra. If $\phi, \phi' \colon V \to W$ are two intertwining maps and $\phi' \neq 0$, then we must have $\phi = \lambda\phi'$ for some $\lambda \in \mathbb{C}$.*

Proof. Since $\phi' \neq 0$, it must be an isomorphism. Then $\phi'^{-1}\phi \colon V \to V$ is an intertwining map, which means $\phi'^{-1}\phi = \lambda \mathrm{id}_V$ for some $\lambda \in \mathbb{C}$. Then $\phi = \lambda\phi'$, as required. □

Corollary 5.2.10. *Let (V, ρ) be an irreducible complex representation of a matrix group G. If A is in the centre of G (i.e. $AB = BA$ for all $B \in G$) then $\rho(A) = \lambda \mathrm{id}_V$ for some $\lambda \in \mathbb{C}$. Similarly, if (V, ρ) is an irreducible complex representation of a Lie algebra $\mathfrak{g}$ and X is in the centre of $\mathfrak{g}$ (i.e. $[X, Y] = 0$ for all $Y \in \mathfrak{g}$) then $\rho(X) = \lambda \mathrm{id}_V$ for some $\lambda \in \mathbb{C}$.*

Proof. We again just prove the group case. If A is in the centre of G, then for all $B \in G$ we have

$$\rho(A)\rho(B) = \rho(AB) = \rho(BA) = \rho(B)\rho(A).$$

This means that $\rho(A) \colon V \to V$ is an intertwining map, and so we must have $\rho(A) = \lambda \mathrm{id}_V$ by Corollary 5.2.8. □

Corollary 5.2.11. *An irreducible complex representation of a commutative matrix group or Lie algebra is 1-dimensional.*

Proof. We consider the case of an irreducible representation (V, ρ) of a commutative group G. Then Corollary 5.2.10 says that for every $A \in G$ we have $\rho(A) = \lambda(A)\mathrm{id}_V$ for some $\lambda(A) \in \mathbb{C}$. This means that *every* subspace of V is invariant, so the only way V can have no non-trivial invariant subspaces is if it is 1-dimensional. □

5.3 Completely reducible representations

Definition 5.3.1. Suppose (V, ρ) and (W, σ) are representations of a matrix group G. Then their *direct sum* is the representation on $V \oplus W$ where $g \cdot (v, w) = (\rho(g)v, \sigma(g)w)$. The direct sum of Lie algebra representations is defined similarly.

Remark 5.3.2. As a group homomorphism, the direct sum is the representation

$$G \to G \times G \to \mathrm{GL}(V) \times \mathrm{GL}(W) \to \mathrm{GL}(V \oplus W),$$

where the first map is the diagonal and the last is the inclusion of the automorphisms of $V \oplus W$ that act on V and W separately, as in Example 2.2.14. Similarly, the direct sum of Lie algebra representations is the composite

$$\mathfrak{g} \to \mathfrak{g} \oplus \mathfrak{g} \to \mathfrak{gl}(V) \oplus \mathfrak{gl}(W) \to \mathfrak{gl}(V \oplus W);$$

this shows in particular that the Lie algebra representation corresponding to a direct sum of group representations is also the direct sum.

We would like to decompose representations into direct sums of irreducible ones. Unfortunately it turns out that this is not always true, so we instead make it a definition:

Definition 5.3.3. A finite-dimensional representation (of a matrix group or Lie algebra) is *completely reducible* if it is isomorphic to a direct sum of fintely many irreducible representatioins.

Example 5.3.4. Consider the representation ρ of $(\mathbb{R}, +)$ on $\mathbb{R}^2$ (or $\mathbb{C}^2$) where $x \in \mathbb{R}$ acts by

$$\rho(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

We claim this representation is *not* completely reducible. The 1-dimensional subspace spanned by $e_1 = (1, 0)$ is clearly invariant and irreducible, but we claim this is the *only* non-trivial invariant subspace; in particular, it does not have any invariant complement. To see this, suppose the subspace spanned by some vector $v = (a, b)$ is invariant. That means we can write

$$\rho(x)v = \lambda(x)v \qquad \Longleftrightarrow \qquad (a + bx, b) = (\lambda(x)a, \lambda(x)b)$$

for all x. This is impossible if $b \neq 0$ since we must have $\lambda(x) = 1$ and so $a = a+bx$ for all x.

Proposition 5.3.5. *Let V be a finite-dimensional representation (of a matrix group or Lie algebra). Then V is completely reducible if and only if for every invariant subspace $W \subseteq V$ there exists another invariant subspace W' such that V decomposes as the direct sum representation $W \oplus W'$.*

Proof. Suppose first that V has the given property; we want to show that V is a direct sum of irreducible representations. Let V_1 be an invariant subspace of V of minimal dimension; then V_1 is necessarily irreducible. If $V = V_1$ then we are done; otherwise we can choose an invariant complement U_1 so that $V = V_1 \oplus U_1$. We can then repeat the process by choosing an invariant subspace $V_2 \subseteq U_1$ of minimal dimension. Then $V_1 \cap V_2 \subseteq V_1 \cap U_1 = 0$, so the sum of V_1 and V_2 in V is direct. Now either $V = V_1 \oplus V_2$, or we can find a non-zero invariant complement U_2. Iterating this process, we can write V as a sum $V_1 \oplus \cdots \oplus V_n \oplus U_n$ where the V_i are irreducible. Since V is finite-dimensional we must eventually get $U_n = 0$, at which point we have written V as a direct sum of irreducibles, as desired.

To prove the converse, suppose $V = V_1 \oplus \cdots \oplus V_n$ is a direct sum of irreducibles V_i, and $U \subseteq V$ is an invariant subspace. We want to show that U has an invariant complement in V. For each i, $U \cap V_i$ is an invariant subspace of V_i, and so is either 0 or all of V_i. If $U \neq V$ there must exist some i_1 such that $V_{i_1} \cap U = 0$; then the sum U_1 of U and V_{i_1} in V is direct. We can now repeat the process with U_1, so that we obtain a sequence of subspaces $U_n = V_{i_1} \oplus \cdots \oplus V_{i_n} \oplus U$. Since V is finite-dimensional we must eventually get $U_n = V$, in which case we have found an invariant complement of U, namely $V_{i_1} \oplus \cdots \oplus V_{i_n}$. □

Corollary 5.3.6.

(i) Let G be a connected matrix group with Lie algebra $\mathfrak{g}$. Then a representation (V, ρ) of G is completely reducible if and only if the induced representation $(V, \mathfrak{L}(\rho))$ of $\mathfrak{g}$ is completely reducible.

(ii) Let $\mathfrak{g}$ be an $\mathbb{R}$-Lie algebra. Then a complex representation (V, ρ) of $\mathfrak{g}$ is completely reducible if and only if the induced representation $(V, \rho_{\mathbb{C}})$ of $\mathfrak{g} \otimes \mathbb{C}$ is completely reducible.

Proof. In both cases, (V, ρ) is completely reducible if and only if every invariant subspace has an invariant complement. But in (i) the G-invariant subspaces for ρ are precisely the $\mathfrak{g}$-invariant subspaces for $\mathfrak{L}(\rho)$ by Proposition 5.2.3, while (ii) follows similarly from Proposition 5.2.5. □

Lemma 5.3.7. *Any invariant subspace of a completely reducible representation is again completely reducible.*

Proof. Let V be a completely reducible representation, and let $U \subseteq V$ be an invariant subspace. We will show that U has the invariant complement property from Proposition 5.3.5. To that end, let $W \subseteq U$ be an invariant subspace. Since V is completely reducible, we can find an invariant complement W' of W in V. Set $W'' := W' \cap U$; we claim that this gives the desired invariant complement of W' in U. Certainly $W'' \cap W \subseteq W' \cap W = 0$ so that we have a direct sum $W \oplus W'' \subseteq U$. To see that this sum is all of U, we observe that any $u \in U$ can be written as $u = w + w'$ with $w \in W$ and $w' \in W'$. But since $W \subseteq U$ we must have $w' = u - w \in U$ so that $U \subseteq W \oplus W'$. □

Proposition 5.3.8. *Every finite-dimensional* unitary *complex representation of a matrix group or Lie algebra is completely reducible.*

Proof. Let (V, ρ) be a finite-dimensional unitary complex representation of a matrix group G. Suppose $W \subseteq V$ is an invariant subspace, and let $W^\perp$ be its orthogonal complement. We claim that then $W^\perp$ is also invariant, which will imply that V is completely reducible by Proposition 5.3.5.

Indeed, for $w \in W$, $w' \in W^\perp$, and $A \in G$, we have

$$\langle w, \rho(A)(w') \rangle = \langle \rho(A)^\dagger w, w' \rangle = \langle \rho(A^{-1}) w, w' \rangle = 0$$

since $\rho(A)^\dagger = \rho(A)^{-1} = \rho(A^{-1})$ as the representation is unitary, and $\rho(A^{-1})w$ is in W by invariance. Thus $\rho(A)(w')$ is in $W^\perp$ for all $A \in G$, as required.

The same argument works if (V, ρ) is a unitary representation of a Lie algebra $\mathfrak{g}$, the only difference being that in the inner product calculation we now use $\rho(X)^\dagger = -\rho(X)$ instead. □

Corollary 5.3.9. *Every finite-dimensional complex representation of a* finite *group is completely reducible.*

Proof. Let (V, ρ) be a finite-dimensional complex representation of a finite group G. We are going to construct an inner product on V for which the representation is unitary. We start by picking an arbitrary inner product $\langle -, - \rangle$ on V. Now we define

$$\langle v, w \rangle_G := \sum_{g \in G} \langle \rho(g)v, \rho(g)w \rangle.$$

We claim that then $\langle -, - \rangle_G$ is again an inner product. Sesquilinearity is immediate and conjugate symmetry are immediate from the same properties of $\langle -, - \rangle$, so we only need to check positive definiteness:

$$\langle v, v \rangle_G = \sum_{g \in G} \|\rho(g)v\|^2 \geq 0,$$

and this equals 0 if and only if $\|\rho(g)v\|^2$ is zero for all g, which can only happen for $v = 0$ since $\langle -, - \rangle$ was positive definite.

To see that ρ is unitary with respect to $\langle -, - \rangle_G$, we compute

$$\langle v, \rho(h)w \rangle_G = \sum_{g \in G} \langle \rho(g)v, \rho(gh)w \rangle = \sum_{g' \in G} \langle \rho(g'h^{-1})v, \rho(g')w \rangle = \langle \rho(h)^{-1}v, w \rangle_G,$$

where we reindex the sum over $g' = gh$. This says $\rho(h)^\dagger = \rho(h)^{-1}$, i.e. $\rho(h)$ is unitary, as required. Applying Proposition 5.3.8, this implies that (V, ρ) is completely reducible. □

The key idea here was to define a new inner product on V by averaging over the group G. This makes sense not just for finite groups, but for any *compact* matrix group G, if we replace the sum over the group by a suitable integral to define

$$\langle v, w \rangle_G := \int_G \langle \rho(g)v, \rho(g)w \rangle \, dg.$$

Making this precise is not hard, but requires some tools beyond the scope of this course (we need to consider a G-invariant differential form on G and integrate the inner product with respect to this). Assuming that this works, we can use exactly the same argument to show that any finite-dimensional complex representation of a compact matrix group can be made unitary, and apply Proposition 5.3.8 to conclude:

Theorem 5.3.10. *Any finite-dimensional complex representation of a* compact *matrix group is completely reducible.* □

Corollary 5.3.11. *Suppose $\mathfrak{g}$ is a $\mathbb{C}$-Lie algebra such that there exists a simply connected compact matrix group K whose Lie algebra $\mathfrak{k}$ satisfies $\mathfrak{k} \otimes \mathbb{C} \cong \mathfrak{g}$. Then every finite-dimensional complex representation of $\mathfrak{g}$ is completely reducible.*

Proof. Suppose (V, ρ) is a finite-dimensional complex representation of $\mathfrak{g}$. Then we can lift V first to a complex representation of $\mathfrak{k}$ by Remark 5.1.6, and then to a complex representation of K by Remark 5.1.7. Since K is compact, V splits as a finite sum of irreducible K-representations by Theorem 5.3.10. But then these give irreducible representations of $\mathfrak{k}$ by Corollary 5.2.4 and so of $\mathfrak{g}$ by Corollary 5.2.6. Thus V is also completely reducible as a representation of $\mathfrak{g}$. □

Example 5.3.12 (Representations of $U_1 \cong SO_2(\mathbb{R})$)**.** The group U_1 consists of the norm-1 complex numbers under multiplication. Since it is compact, its finite-dimensional complex representations are completely reducible by Theorem 5.3.10. In this simple case this is also easy to deduce directly from Proposition 5.3.8: if (V, ρ) is a representation of U_1 and $\langle -, - \rangle$ is an inner product on V, then it is easy to check that

$$\langle v, w \rangle_{U_1} := \int_0^{2\pi} \langle \rho(e^{it})(v), \rho(e^{it})(w) \rangle \, \mathrm{d}t,$$

is an inner product, and by changing variables in the integral we get

$$\begin{aligned} \langle v, \rho(e^{is})w \rangle_{U_1} &= \int_0^{2\pi} \langle \rho(e^{it})(v), \rho(e^{i(t+s)})(w) \rangle \, \mathrm{d}t \\ &= \int_0^{2\pi} \langle \rho(e^{i(t-s)})(v), \rho(e^{it})(w) \rangle \, \mathrm{d}t \\ &= \langle \rho(e^{-is})v, w \rangle_{U_1}, \end{aligned}$$

i.e. (V, ρ) is unitary with respect to this inner product. Since U_1 is commutative we also know that all of its irreducible representations are 1-dimensional. Such a 1-dimensional representation is a continuous homomorphism $\rho \colon U_1 \to GL_1(\mathbb{C}) \cong (\mathbb{C} \backslash \{0\}, \times)$. If we define a function $f \colon \mathbb{R} \to GL_1(\mathbb{C})$ by $f(t) := \rho(e^{it})$ then this is a one-parameter subgroup, so that $f(t) = e^{t\lambda}$ for a unique $\lambda \in \mathbb{C}$ by Proposition 3.6.4 (or we can compute directly that

$$f'(t) = \lim_{h \to 0} \frac{f(t+h) - f(t)}{h} = \lim_{h \to 0} \frac{f(t)(f(h) - 1)}{h} = f(t)f'(0),$$

where the only solution with $f(0) = 1$ is $f(t) = e^{f'(0)t}$.) Since $f(2\pi) = f(0) = 1$ we must have $e^{2\pi\lambda} = 1$, which means that $\lambda = in$ for some integer $n \in \mathbb{Z}$. In terms of ρ, this means $\rho(z) = z^n$ for some integer $n \in \mathbb{Z}$. Thus every finite-dimensional complex representation of U_1 is a direct sum of 1-dimensional representations of this form.

Remark 5.3.13. The classification of 1-parameter subgroups also shows that 1-dimensional complex representations of $(\mathbb{R}, +)$ are of the form $\rho(t) = e^{t\lambda}$ for some unique $\lambda \in \mathbb{C}$. More generally, representations on a vector space V are of the form $\rho(t) = e^{tA}$ where A is a linear endomorphism of V. On the other hand, representations of the commutative Lie algebra $\mathbb{R}$ (which is the Lie algebra of both U_1 and $(\mathbb{R}, +)$) on V are the maps of the form $\rho(t) = tA$ for some endomorphism A. In particular, the 1-dimensional complex representations are given by $\rho(t) = t\lambda$ for $\lambda \in \mathbb{C}$. We thus see that, as expected, representations of the Lie algebra $\mathbb{R}$ correspond bijectively to representations of the simply connected group $(\mathbb{R}, +)$. Only some of these representations lift to U_1, which is not simply connected, but the lifts are unique when they exist, as they should be since U_1 is connected. (On the other hand, we saw earlier that representations of $(\mathbb{R}, +)$, and so also of the Lie algebra $\mathbb{R}$, need not be completely reducible.)

Chapter 6

Representations of $\mathfrak{sl}_2$ and $\mathfrak{sl}_3$

6.1 Representations of $\mathfrak{sl}_2$

From now on we will be considering complex representations of complex Lie algebras, so we take the complex versions of Lie algebras as the default ones. That is, we will write

$$\mathfrak{sl}_n := \mathfrak{sl}_n(\mathbb{C}), \qquad \mathfrak{so}_n := \mathfrak{so}_n(\mathbb{C}), \qquad \mathfrak{sp}_n := \mathfrak{sp}_n(\mathbb{C}).$$

In this section we will analyze the representations of $\mathfrak{sl}_2 = \mathfrak{sl}_2(\mathbb{C})$. This is not only a simple but interesting example in its own right (for example, the irreducible representations of $\mathfrak{sl}_2$ correspond to the possible "spins" of particles such as electrons), but also plays a central role in understanding the representations of more complex Lie algebras.

Recall that $\mathfrak{sl}_2$ is the $\mathbb{C}$-vector space of traceless 2×2 complex matrices. This has the basis

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

Exercise 6.1. Check that we have the following commutation relations for $\mathfrak{sl}_2$:

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

Let (V, ρ) be a finite-dimensional complex representation of $\mathfrak{sl}_2$. We will make a series of elementary observations about the linear maps $\rho(H), \rho(X), \rho(Y)$:

(1) Since we're working over $\mathbb{C}$, the endomorphism $\rho(H)$ has at least one eigenvector $v \in V$ with eigenvalue λ. Then the commutation relations in Exercise 6.1 give

$$\rho(H)\rho(X)v - \rho(X)\rho(H)v = 2\rho(X)v,$$

so that

$$\rho(H)\rho(X)v = (\lambda + 2)\rho(X)v,$$

so $\rho(X)v$ is an eigenvector of $\rho(H)$ with eigenvalue $\lambda + 2$. Similarly, $\rho(Y)v$ is an eigenvector with eigenvalue $\lambda - 2$.

(2) Iterating, we get that $\rho(X)^n v$ and $\rho(Y)^n v$ are either 0 or eigenvectors of $\rho(H)$ with eigenvalues $\lambda + 2n$ and $\lambda - 2n$, respectively.

(3) Since V is finite-dimensional, $\rho(H)$ has only finitely many eigenvalues. Thus there exists some m such that $\rho(X)^m v \neq 0$ but $\rho(X)^{m+1} v = 0$.

(4) If we set $u_0 := \rho(X)^m v$ and $\alpha = \lambda + 2m$, then $\rho(H)u_0 = \alpha u_0$ and $\rho(X)u_0 = 0$. If we define $u_k := \rho(Y)^k u_0$, then $\rho(H)u_k = (\alpha - 2k)u_k$. Now the third commutation relation $[X, Y] = H$ gives by induction on k that

$$\rho(X)u_k = k(\alpha - k + 1)u_{k-1}$$

for $k \geq 0$: This holds trivially for $k = 0$, and if it holds for k then

$$\begin{aligned} \rho(H)u_k &= \rho(X)\rho(Y)u_k - \rho(Y)\rho(X)u_k \\ &= \rho(X)u_{k+1} - k(\alpha - k + 1)u_k, \\ \rho(X)u_{k+1} &= (\alpha - 2k)u_k + k(\alpha - k + 1)u_k \\ &= (k + 1)(\alpha - k)u_k. \end{aligned}$$

(5) Since $\rho(H)$ still has only finitely many eigenvalues, the u_k must eventually be 0. In other words, there is an $m \in \mathbb{N}$ such that $u_k \neq 0$ for $k \leq m$, but $u_{m+1} = \rho(Y)^{m+1} u_0 = 0$. But then the formula above gives

$$0 = \rho(X)u_{m+1} = (m + 1)(\alpha - m)u_m.$$

As $u_m \neq 0$, this can only happen if $\alpha = m$.

(6) The non-zero vectors $u_0, \ldots, u_m$ are eigenvectors of $\rho(H)$ with distinct eigenvalues, and so they are linearly independent. Moreover, the $(m+1)$-dimensional subspace of V spanned by these vectors is clearly invariant under the action of $\rho(H)$, $\rho(X)$, and $\rho(Y)$, and so under all of $\mathfrak{sl}_2$.

In particular, if we assume the representation V is irreducible, it must be spanned by our vectors $u_0, \ldots, u_m$, so that we have completely described the possible structure of an irreducible $\mathfrak{sl}_2$-representation:

Proposition 6.1.1. *Suppose (V, ρ) is an irreducible finite-dimensional complex representation of $\mathfrak{sl}_2$ of dimension $m + 1$. Then there exists a basis $u_0, \ldots, u_m$ of V for which the action of $\mathfrak{sl}_2$ is described by*

$$\begin{aligned} \rho(H)u_k &= (m - 2k)u_k, \\ \rho(Y)u_k &= \begin{cases} u_{k+1}, & k < m \\ 0, & k = m \end{cases} \\ \rho(X)u_k &= \begin{cases} k(m - k + 1)u_{k-1}, & k > 0 \\ 0, & k = 0. \end{cases} \end{aligned} \tag{6.1}$$

Next, we will show the converse: the equations (6.1) define an irreducible representation of $\mathfrak{sl}_2$ for all m.

Exercise 6.2. Show that the equations in (6.1) always define a representation of $\mathfrak{sl}_2$, i.e. check that the required commutation relations hold for the operations $\rho(H), \rho(X), \rho(Y)$.

Exercise 6.3. Show that the adjoint representation of $\mathfrak{sl}_2$ on itself and the standard representation on $\mathbb{C}^2$ are irreducible of dimension 3 and 2, respectively.

Lemma 6.1.2. *The $\mathfrak{sl}_2$-representation given by (6.1) is irreducible for every m.*

Proof. Let V be the given representation of dimension $m+1$, and suppose $W \subseteq V$ is a non-zero invariant subspace; we then want to show that we must have $W = V$. Let $w \neq 0$ be an element of W. Then we can write

$$w = a_0 u_0 + \cdots + a_m u_m,$$

for $a_i \in \mathbb{C}$. Let i_0 be the smallest value of i such that $a_i \neq 0$, and consider $\rho(Y)^{m-i_0} w$. From (6.1) we see that $\rho(Y)^{m-i_0} u_i = 0$ for $i > i_0$, so that $\rho(Y)^{m-i_0} w = a_{i_0} u_m$. Since $a_{i_0} \neq 0$, we see that u_m lies in W. But then $\rho(X)^k u_m$ lies in W, and this is a non-zero multiple of u_{m-k}. Thus all the basis vectors $u_0, \ldots, u_m$ lie in W, i.e. $W = V$ as required. ☐

We have thus classified the irreducible representations of $\mathfrak{sl}_2$: for every $m \in \mathbb{N}$ there is a unique irreducible representation of dimension $m+1$ described by the equations (6.1). For future reference, we also derive some further consequences of our observations above for a general representation of $\mathfrak{sl}_2$:

Proposition 6.1.3. *Let (V, ρ) be a finite-dimensional complex representation of $\mathfrak{sl}_2$. Then we have:*

(i) Every eigenvalue of $\rho(H)$ is an integer.

(ii) If v is an eigenvector of $\rho(H)$ with eigenvalue k and $\rho(X)v = 0$, then k is non-negative.

(iii) The operators $\rho(X)$ and $\rho(Y)$ are nilpotent.

(iv) If we define $S\colon V \to V$ by $S = e^{\rho(X)} e^{-\rho(Y)} e^{-\rho(X)}$, then S satisfies

$$S\rho(H)S^{-1} = -\rho(H).$$

(v) If $k \in \mathbb{Z}$ is an eigenvalue for $\rho(H)$, then so is each of the integers $-|k|, -|k|+2, \ldots, |k|-2, |k|$.

Proof. Points (i) and (ii) we saw in the discussion above. For (iii), we know that $\rho(H)$ has a basis of generalized eigenvectors, i.e. vectors v such that $(\rho(H) -$

$\lambda I)^k v = 0$ for some $k > 0$ and λ an eigenvalue of $\rho(H)$. Using the commutation relations we get by induction that

$$(\rho(H) - (\lambda + 2)I)^k \rho(X) = \rho(X)(\rho(H) - \lambda I)^k.$$

Thus if v is a generalized eigenvector of $\rho(H)$ with eigenvalue λ, then $\rho(X)v$ is either 0 or a generalized eigenvector with eigenvalue $\lambda + 2$. Since $\rho(H)$ has finitely many eigenvalues, this means that $\rho(X)^k v$ must be 0 for k sufficiently large. Since this is true for each of the finitely many vectors in a basis of V, we see that $\rho(X)^k = 0$ for k sufficiently large, i.e. $\rho(X)$ is nilpotent. Applying the commutation relation for $[H, Y]$ similarly, we see that $\rho(Y)$ is also nilpotent.

To prove (iv), we use that

$$e^{\rho(X)}\rho(H)e^{-\rho(X)} = \mathrm{Ad}(e^{\rho(X)})(\rho(H)) = e^{\mathrm{ad}(\rho(X))}(\rho(H)) = \sum_{n=0}^{\infty} \frac{1}{n!}\mathrm{ad}(\rho(X))^n(\rho(H)).$$

Now here $\mathrm{ad}(\rho(X))(\rho(H)) = [\rho(X), \rho(H)] = \rho([X, H]) = -2\rho(X)$, so that $\mathrm{ad}(\rho(X))^2(\rho(H)) = \rho([X, -2X]) = 0$, and only the first two terms in this series are non-zero. This means that

$$e^{\rho(X)}\rho(H)e^{-\rho(X)} = \rho(H) - 2\rho(X).$$

Now we can similarly compute that

$$\begin{aligned}
e^{-\rho(Y)}e^{\rho(X)}\rho(H)e^{-\rho(X)}e^{\rho(Y)} &= \mathrm{Ad}(e^{\rho(Y)})(\rho(H) - 2\rho(X)) \\
&= -\rho(H) - 2\rho(X), \\
S\rho(H)S^{-1} &= e^{\rho(X)}e^{-\rho(Y)}e^{\rho(X)}\rho(H)e^{-\rho(X)}e^{\rho(Y)}e^{-\rho(X)} \\
&= \mathrm{Ad}(e^{\rho(X)})(-\rho(H) - 2\rho(X)) \\
&= -\rho(H) - 2\rho(X) + 2\rho(X) \\
&= -\rho(H).
\end{aligned}$$

Note that if v is an eigenvector of $\rho(H)$ with eigenvalue λ, then we get

$$S\rho(H)S^{-1}v = -\rho(H)v = -\lambda v,$$

so that $S^{-1}v$ is an eigenvector for $\rho(H)$ with eigenvalue $-\lambda$. From this the final point follows: if k is an eigenvalue of $\rho(H)$, then so is $-k$, so we may assume without loss of generality that k is a non-negative integer. Then we know there exists another eigenvector v_0 with eigenvalue $m := k + 2N$ for some non-negative N, such that $\rho(X)v_0 = 0$, and from this we obtain a chain of eigenvectors with eigenvalues ranging from $-m$ to m in increments of 2. In particular, all of $-k, -k + 2, \ldots, k - 2, k$ are indeed eigenvalues. □

Remark 6.1.4. $\mathfrak{sl}_2$ is the complexification of both $\mathfrak{sl}_2(\mathbb{R})$ and $\mathfrak{su}_2$, which we saw in Example 3.6.9 is isomorphic to $\mathfrak{so}_3(\mathbb{R})$. Thus $\mathfrak{sl}_2$ is not only the Lie algebra of the complex matrix group $SL_2(\mathbb{C})$, but the complexified Lie algebra of the groups $SL_2(\mathbb{R})$, SU_2, and $SO_3(\mathbb{R})$. Since SU_2 is compact and simply connected, we know from Corollary 5.3.11 that all finite-dimensional complex representations of $\mathfrak{sl}_2$ are completely reducible. It is also not hard to show this directly.

We know from Remark 5.1.7 and Remark 5.1.6 that all the irreducible representations of $\mathfrak{sl}_2$ can be lifted to representations of SU_2, since this is simply connected. We can also give an explicit construction of these representations of SU_2:

Exercise 6.4. Let V_m be the vector space of degree-m homogeneous polynomials in two complex variables z, w, so that V_m is spanned by the basis $w^m, zw^{m-1}, \dots, z^{m-1}w, z^m$. For $U \in SU_2$ define an operation $\rho(U)$ on V_m by

$$\rho(U)(f(z,w)) = f(U^{-1}(z,w)) = f\left((U^{-1})_{11}z + (U^{-1})_{12}w, (U^{-1})_{21}z + (U^{-1})_{22}w\right).$$

(i) Check that $\rho(U)(f(z,w))$ is again a homogeneous polynomial of degree m, so that $\rho(U)$ is a $\mathbb{C}$-linear endomorphism of V_m.

(ii) Check that ρ is then a representation of SU_2.

Exercise 6.5. Identify the representation of $\mathfrak{su}_2$ associated to the SU_2-representation in the previous exercise, as follows:

(i) The action of $M \in \mathfrak{su}_2$ is defined by

$$\mathfrak{L}(\rho)(M)(f(z,w)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(e^{-tM}(z,w)).$$

Use the chain rule to show that if

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then this gives

$$\mathfrak{L}(M)(f(z,w)) = -(az+bw)\frac{\partial f}{\partial z} - (cz+dw)\frac{\partial f}{\partial w}.$$

(ii) The same formula for $M \in \mathfrak{sl}_2$ gives the induced representation of $\mathfrak{sl}_2 \simeq \mathfrak{su}_2 \otimes \mathbb{C}$, so that we have

$$\mathfrak{L}(\rho)(H) = -z\frac{\partial}{\partial z} + w\frac{\partial}{\partial w}, \qquad \mathfrak{L}(\rho)(X) = -w\frac{\partial}{\partial z}, \qquad \mathfrak{L}(\rho)(Y) = -z\frac{\partial}{\partial w}.$$

Compute that on the basis vector $z^{m-k}w^k$ we get

$$\begin{aligned} \mathfrak{L}(\rho)(H)(z^{m-k}w^k) &= (-m+2k)z^{m-k}w^k, \\ \mathfrak{L}(\rho)(X)(z^{m-k}w^k) &= -(m-k)z^{m-k-1}w^{k+1}, \\ \mathfrak{L}(\rho)(Y)(z^{m-k}w^k) &= -kz^{m-k+1}w^{k-1}. \end{aligned}$$

(iii) Conclude that $\mathfrak{L}(\rho)$ is isomorphic to the $(m+1)$-dimensional irreducible representation of $\mathfrak{sl}_2$ we found above.

The group $\mathrm{SO}_3(\mathbb{R})$ also has $\mathfrak{sl}_2$ as its complexified Lie algebra. This is compact, so its complex representations are completely reducible by Theorem 5.3.10, and it is connected, so any representation of $\mathfrak{sl}_2$ lifts to at most one representation of $\mathrm{SO}_3(\mathbb{R})$ by Proposition 5.1.11. Moreover, the irreducible representations of $\mathrm{SO}_3(\mathbb{R})$ are precisely those that are lifts of irreducible $\mathfrak{sl}_2$-representations, by Corollary 5.2.4. To understand the complex representations of $\mathrm{SO}_3(\mathbb{R})$, we therefore only need to know exactly *which* of the irreducible representations of $\mathfrak{sl}_2$ arise from $\mathrm{SO}_3(\mathbb{R})$-representations:

Proposition 6.1.5. *The $(m+1)$-dimensional irreducible representation of $\mathfrak{sl}_2$ lifts to an $\mathrm{SO}_3(\mathbb{R})$-representation if and only if m is even (i.e. the dimension of the representation is odd).*

Proof. First suppose that m is odd; we want to show that the $(m+1)$-dimensional representation (V,ρ) of $\mathfrak{sl}_2$ with basis $u_0,\dots,u_m$ as in Proposition 6.1.1 does not lift to $\mathrm{SO}_3(\mathbb{R})$. Let us write σ for the corresponding representation of $\mathfrak{so}_3(\mathbb{R})$ (so that $\rho = \sigma_{\mathbb{C}}$). We use the basis F_1, F_2, F_3 for $\mathfrak{so}_3(\mathbb{R})$ from Example 3.6.9, corresponding to the basis E_1, E_2, E_3 for $\mathfrak{su}_2$. In terms of our basis H, X, Y for $\mathfrak{sl}_2$ we then have $E_1 = \frac{i}{2}H$, so that

$$\sigma(F_1)u_k = \rho(E_1)u_k = \frac{i}{2}\rho(H)u_k = \frac{i(m-2k)}{2}u_k$$

In terms of the basis $u_0,\dots,u_m$ the linear map $\sigma(F_1)$ is represented by the diagonal matrix

$$\begin{pmatrix} \frac{im}{2} & & & 0 \\ & \frac{i(m-2)}{2} & & \\ & & \ddots & \\ 0 & & & \frac{-im}{2} \end{pmatrix},$$

so that in the same basis $e^{\sigma(2\pi F_1)} = -I$, since it is diagonal with entries $e^{(m-2k)\pi i}$ where $(m-2k)$ is by assumption an odd integer. Thus a lift Σ of σ to $\mathrm{SO}_3(\mathbb{R})$ must satisfy

$$\Sigma(e^{2\pi F_1}) = e^{\sigma(2\pi F_1)} = -I.$$

On the other hand, we can compute that in $\mathrm{SO}_3(\mathbb{R})$ we have

$$e^{2\pi F_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\pi & -\sin 2\pi \\ 0 & \sin 2\pi & \cos 2\pi \end{pmatrix} = I$$

(for example, we can diagonalize the matrix F_1 over $\mathbb{C}$ and use this to exponentiate it). But then $\Sigma(e^{2\pi F_1}) = \Sigma(I) = I$, a contradiction.

Now suppose m is even. Then the representation (V_m, ρ) of SU_2 from Exercise 6.4 satisfies $\rho(-I) = I$ since for a homogeneous polynomial $f(z,w)$ of degree m we have $f(-z,-w) = (-1)^m f(z,w) = f(z,w)$. Thinking of ρ as a continuous homomorphism $SU_2 \to GL(V_m)$ it follows that ρ factors through the quotient by the subgroup $\{\pm I\}$. But we identified this quotient group with $SO_3(\mathbb{R})$ in the proof of Corollary 2.3.7, so this means that ρ factors through the homomorphism $r\colon SU_2 \to SO_3(\mathbb{R})$; as this gives an isomorphism $\mathfrak{su}_2 \xrightarrow{\sim} \mathfrak{so}_3(\mathbb{R})$ on Lie algebras, this means that the $(m+1)$-dimensional irreducible representation of $\mathfrak{sl}_2$ comes from a representation of $SO_3(\mathbb{R})$, as required. □

Remark 6.1.6. For any m, the representation V_m of SU_2 gives a *projective* representation of $SO_3(\mathbb{R})$, meaning a continuous homomorphism

$$SO_3(\mathbb{R}) \to PGL(V) := GL(V)/\{\pm I\}.$$

In quantum mechanics, symmetries of a physical system give a projective representation of the symmetry group on the Hilbert space of quantum states. Thus a system with rotational symmetry in 3 dimensions gives a projective representation of $SO_3(\mathbb{R})$. The projective representation of $SO_3(\mathbb{R})$ coming from V_m describes the states of a particle with *spin* $m/2$.

6.2 Roots and weights for $\mathfrak{sl}_3$

Our next goal is to analyze the representations of the Lie algebra $\mathfrak{sl}_3$, consisting of traceless 3×3 complex matrices. Before we start, we will pick a basis for $\mathfrak{sl}_3$:[1]

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$X_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad X_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$Y_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad Y_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

We also note the following commutation relations (which follow from the relations for $\mathfrak{sl}_2$ by inserting a row and column of 0's):

$$\begin{aligned} [H_1, X_1] &= 2X_1, & [H_2, X_2] &= 2X_2, \\ [H_1, Y_1] &= -2Y_1, & [H_2, Y_2] &= -2Y_2, \\ [X_1, Y_1] &= H_1, & [X_2, Y_2] &= H_2. \end{aligned}$$

[1]While this hopefully looks like a reasonable basis in any case, we will soon see (cf. Exercise 6.6) that there is a good reason for picking this particular basis.

Thus the subalgebra spanned by H_i, X_i, Y_i is a copy of $\mathfrak{sl}_2$ for $i = 1, 2$. Also note that we have

$$[H_1, H_2] = 0,$$

so the subalgebra $\mathfrak{h}$ spanned by H_1 and H_2 (which consists of all traceless diagonal matrices) is commutative.

When we analyzed representations ρ of $\mathfrak{sl}_2$, the operator $\rho(H)$ and its eigenvalues played a key role. For $\mathfrak{sl}_3$, we must instead look at *simultaneous* eigenvectors for the operators $\rho(H_1)$ and $\rho(H_2)$.

Definition 6.2.1. Let (V, ρ) be a finite-dimensional complex representation of $\mathfrak{sl}_3$. A *weight vector* for ρ is a non-zero vector $v \in V$ such that we have

$$\rho(H_1)(v) = a_1 v, \qquad \rho(H_2)(v) = a_2 v$$

for some $a_1, a_2 \in \mathbb{C}$. (In other words, v is an eigenvector of both $\rho(H_1)$ and $\rho(H_2)$.) The pair of eigenvalues $\mu = (a_1, a_2)$ is then called a *weight* of ρ, and we call the subspace of V of weight vectors with this weight the *weight space* of μ. The *multiplicity* of the weight μ is the dimension of its weight space.

Remark 6.2.2. It is sometimes convenient to think of the weight $\mu = (a_1, a_2)$ as a linear functional $\mu \in \mathfrak{h}^*$ on the commutative subalgebra $\mathfrak{h}$, determined by

$$\mu(H_i) = a_i$$

on the basis vectors. Then a non-zero $v \in V$ is a weight vector for μ if it satisfies

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

for $H \in \mathfrak{h}$.

Lemma 6.2.3. *Let (V, ρ) be a finite-dimensional complex representation of $\mathfrak{sl}_3$.*

(i) The representation ρ has at least one weight.

(ii) If $\mu = (a_1, a_2)$ is a weight for ρ, then both a_1 and a_2 are integers.

Proof. The operator $\rho(H_1)$ has at least one eigenvalue $a_1 \in \mathbb{C}$; let $W \subseteq V$ be the corresponding eigenspace. Then for $w \in W$ we have (as $[\rho(H_1), \rho(H_2)] = \rho([H_1, H_2]) = 0$)

$$\rho(H_1)\rho(H_2)w = \rho(H_2)\rho(H_1)w = a_1\rho(H_2)w,$$

so that $\rho(H_2)$ restricts to a linear map $W \to W$. This restriction again has at least one eigenvalue a_2 with eigenvector $w \in W$, but then by construction w is a weight vector with weight (a_1, a_2).

By restricting ρ to the two copies of $\mathfrak{sl}_2$ spanned by $\{H_i, X_i, Y_i\}$ we see from Proposition 6.1.3 that a_i must be an integer for $i = 1, 2$. □

Exercise 6.6. Find (or look up, e.g. in [2, Section 6.2]) all the commutation relations for our basis of $\mathfrak{sl}_3$. Conclude from this that the adjoint representation $\mathrm{ad}_{\mathfrak{sl}_3}$ has the following 7 weights, with corresponding weight vectors:

Weight	Weight vector
$(0,0)$	H_1, H_2
$(2,-1)$	X_1
$(-1,2)$	X_2
$(1,1)$	X_3
$(-2,1)$	Y_1
$(1,-2)$	Y_2
$(-1,-1)$	Y_3

(Note that since the weight vectors form a basis for $\mathfrak{sl}_3$, there can be no other weights.)

Definition 6.2.4. A *root* of $\mathfrak{sl}_3$ is a *non-zero* weight of the adjoint representation $\mathrm{ad}_{\mathfrak{sl}_3}$. In other words, a root is a pair (a_1, a_2) of integers, not both zero, such that there exists $Z \neq 0$ in $\mathfrak{sl}_3$ with $[H_i, Z] = a_i Z$ for $i = 1, 2$. We see from Exercise 6.6 that $\mathfrak{sl}_3$ has exactly 6 roots.

Exercise 6.7. Let

$$\alpha_1 = (2,-1), \qquad \alpha_2 = (-1,2).$$

Check that all the roots of $\mathfrak{sl}_3$ can be expressed as integer linear combinations of α_1 and α_2, with coefficients that are either both positive or both negative.

Definition 6.2.5. We will call α_1 and α_2 the *positive simple roots* of $\mathfrak{sl}_3$. (This is somewhat arbitrary: What we need is a choice of roots with the property in Exercise 6.7.)

The following is the key computation for describing representations of $\mathfrak{sl}_3$:

Lemma 6.2.6. *Let (V, ρ) be a finite-dimensional complex representation of $\mathfrak{sl}_3$, and let μ be a weight of ρ with weight vector v. If α is a root of $\mathfrak{sl}_3$ with root vector Z_α, then for $H \in \mathfrak{h}$ we have*

$$\rho(H)\rho(Z_\alpha)(v) = (\mu + \alpha)(H)\rho(Z_\alpha)v.$$

Thus either $\rho(Z_\alpha)(v) = 0$ or $\rho(Z_\alpha)(v)$ is a weight vector with weight $\mu + \alpha$.

Proof. By definition, we have $[H, Z_\alpha] = \alpha(H)Z_\alpha$, so that

$$\rho(H)\rho(Z_\alpha)(v) = \rho(Z_\alpha)\rho(H)(v) + \alpha(H)\rho(Z_\alpha)(v) = (\mu(H) + \alpha(H))\rho(Z_\alpha)(v). \quad \square$$

If we start with a weight vector we can thus apply $\rho(X_i)$ and $\rho(Y_i)$ to it to get new weight vectors. Since ρ is a finite-dimensional representation there can only be finitely many weights, so iterating this process must always eventually produce 0's. For $\mathfrak{sl}_2$ our strategy was to iterate X to find a maximal eigenvalue for H. We now want to do something similar for $\mathfrak{sl}_3$, which requires a notion of one weight being "higher" than another:

Definition 6.2.7. A weight μ is *higher* than μ' (and μ' is *lower* than μ) if we can write

$$\mu - \mu' = a_1\alpha_1 + a_2\alpha_2$$

with $a_i \geq 0$ for $i = 1, 2$. We write $\mu \geq \mu'$ when μ is higher than μ'.

Definition 6.2.8. If (V, ρ) is a representation of $\mathfrak{sl}_3$, then a weight μ for ρ is a *highest weight* if for all weights μ' of ρ we have $\mu \geq \mu'$.

Definition 6.2.9. We say that a pair $\mu = (a_1, a_2)$ in $\mathbb{C}^2$ is *integral* if the a_i are integers, and *dominant* if the a_i are real and non-negative.

Remark 6.2.10. We make some simple observations about the notion of "higher" weights:

- The relation $\geq$ is only a *partial* order. For example $\alpha_1 - \alpha_2$ is neither lower nor higher than 0.
- The coefficients a_i when we write $\mu - \mu' = a_1\alpha_1 + a_2\alpha_2$ are not required to be integers. For example, $(1, 0) \geq (0, 0)$ since $(1, 0) = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$.
- The partial ordering $\geq$ is somewhat arbitrary, since it depends on our choice of positive simple roots; the results we will prove about highest weights would hold equally for any other choice.

6.3 The theorem of the highest weight

We can now state the following theorem, which summarizes the results we are going to prove about the irreducible representations of $\mathfrak{sl}_3$:

Theorem 6.3.1 ("Theorem of the highest weight"). *Let V be a finite-dimensional irreducible complex representation of $\mathfrak{sl}_3$. Then:*

(i) As a vector space, V is the direct sum of its weight spaces.

(ii) The representation V has a unique *highest weight, which must be a* dominant integral *weight.*

(iii) If V' is another irreducible representation with the same highest weight as V, then $V' \cong V$.

Our goal for the rest of this section is to prove each part of this theorem in turn. In the next section we will show that for every dominant integral pair there exists an irreducible representation with that as its highest weight.

Lemma 6.3.2. *Let (V, ρ) be a finite-dimensional irreducible complex representation of $\mathfrak{sl}_3$. Then there exists a basis of V with respect to which which the operators $\rho(H_1)$ and $\rho(H_2)$ are both diagonal. In other words, V is the direct sum of its weight spaces.*

Proof. Let W be the subspace of V spanned by the simultaneous eigenvectors of $\rho(H_1)$ and $\rho(H_2)$, i.e. the sum of the weight spaces in V. Since weight vectors for distinct weights are linearly independent, this sum is direct. Moreover, $W \neq 0$ since V has at least one weight by Lemma 6.2.3. On the other hand, Lemma 6.2.6 shows that if $Z \in \mathfrak{sl}_3$ is a root vector, then $\rho(Z)$ takes W to itself. Since $\mathfrak{sl}_3$ is spanned by the root vectors together with H_1 and H_2 by Exercise 6.6, it follows that W is an invariant subspace of V. Since V is by assumption irreducible, we must then have $W = V$, as required. □

Definition 6.3.3. A finite-dimensional complex representation (V, ρ) of $\mathfrak{sl}_3$ is *highest weight cyclic* with weight μ if there exists a weight vector $v \neq 0$ for μ such that $\rho(X_i)v = 0$ for $i = 1, 2$, and the only invariant subspace of V that contains v is V itself.

Lemma 6.3.4 (Reordering lemma)**.** *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and (V, ρ) a representation of $\mathfrak{g}$. If $X_1, \ldots, X_m$ is an ordered basis for $\mathfrak{g}$ as a vector space, then any expression of the form $\rho(X_{i_1}) \cdots \rho(X_{i_n})$ can be rewritten as a linear combination of terms of the form $\rho(X_m)^{k_m} \cdots \rho(X_1)^{k_1}$ where $k_i \in \mathbb{N}$ and $k_1 + \cdots + k_m \leq n$.*

Proof. Since the commutator $[\rho(X_j), \rho(X_k)]$ is $\rho([X_j, X_k])$, which is some linear combination of the $\rho(X_i)$'s, we can replace $\rho(X_j)\rho(X_k)$ by $\rho(X_k)\rho(X_j)$ at the expense of introducing new terms with one fewer factor. Now we just need to induct on the length n of the sequence. □

Proposition 6.3.5. *Let (V, ρ) be a highest weight cyclic representation of $\mathfrak{sl}_3$ with weight μ. Then μ is the unique highest weight of ρ, and the weight space of μ is 1-dimensional.*

Proof. Let v be a weight vector for μ that exhibits ρ as highest weight cyclic, and define W to be the subspace of V spanned by the elements of the form $w = \rho(Y_{i_1}) \cdots \rho(Y_{i_n})(v)$. We then claim that W is invariant. Indeed, if we apply $\rho(Z)$ to w, where Z is some basis element, we can use Lemma 6.3.4 to express this as a sum of terms of the form

$$\rho(Y_1)^{\eta_1}\rho(Y_2)^{\eta_2}\rho(Y_3)^{\eta_3}\rho(H_1)^{i_1}\rho(H_2)^{i_2}\rho(X_1)^{\xi_1}\rho(X_2)^{\xi_2}\rho(X_3)^{\xi_3}v.$$

If any of the ξ_i are non-zero, we get 0, so we may without loss of generality assume $\xi_i = 0$ for all three i; then since v is an eigenvector for $\rho(H_i)$, we see that this term is some multiple of $\rho(Y_i)$'s applied to v and so lies in W, as required. Thus W is an invariant subspace of V that contains v, and so by assumption we have $W = V$.

Moreover, by Lemma 6.2.6 each element w as above with $n > 0$ is a weight vector with weight lower than μ, since Y_1, Y_2, Y_3 are root vectors for the roots $-\alpha_1, -\alpha_2, -\alpha_1 - \alpha_2$, respectively. Thus the only weight vectors in $W = V$ with weight μ are the multiples of v, and all other weights that occur are strictly lower than μ. □

Proposition 6.3.6. *Every finite-dimensional irreducible complex representation of $\mathfrak{sl}_3$ is a highest weight cyclic representation.*

Proof. Let (V, ρ) be a finite-dimensional irreducible complex representation of $\mathfrak{sl}_3$. By Lemma 6.3.2 the vector space V is the direct sum of its weight spaces. Since V is finite-dimensional, there can be only finitely many weights, and so we can find a *maximal* weight μ, meaning that if μ' is a weight such that $\mu' \geq \mu$ then $\mu' = \mu$. By Lemma 6.2.6 this means that if $v \neq 0$ is weight vector with weight μ then $\rho(X_i)(v) = 0$ for $i = 1, 2$. Since V is irreducible, it follows that it is highest weight cyclic. □

Combining Proposition 6.3.6 and Proposition 6.3.5 we get:

Corollary 6.3.7. *Every finite-dimensional irreducible complex representation of $\mathfrak{sl}_3$ has a unique highest weight, and the weight space for this weight is one-dimensional.* □

Proposition 6.3.8. *Suppose (V, ρ) is a finite-dimensional highest weight cyclic representation of $\mathfrak{sl}_3$ that is completely reducible*[2]*. Then V is irreducible.*

Proof. Suppose $v \in V$ exhibits ρ as highest weight cyclic with weight μ. Since V is completely reducible, we have $V \cong \bigoplus V_i$ where each V_i is irreducible. By Lemma 6.3.2 each V_i is a direct sum of its weight spaces. It can be shown (see [2, Proposition A.17]) that since the weight μ occurs in V, it must occur in some V_i. But by Proposition 6.3.5 the weight space for μ in V is 1-dimensional, so this means that v must lie in some V_i. Then V_i is an invariant subspace that contains v, whence $V = V_i$ as required. □

Proposition 6.3.9. *Two irreducible representations of $\mathfrak{sl}_3$ with the same highest weight are isomorphic.*

Proof. Suppose (V, ρ) and (W, σ) are two irreducible representations of $\mathfrak{sl}_3$ with the same highest weight μ, and let $v \in V$ and $w \in W$ be highest weight vectors. Let U be the smallest subrepresentation of $V \oplus W$ that contains (v, w). Then (v, w) exhibits U as highest weight cyclic, and since the direct sum $V \oplus W$ is trivially a completely reducible representation, so is the subspace U by Lemma 5.3.7. It follows from Proposition 6.3.8 that U is irreducible. The projections $V \oplus W \to V, W$ are intertwining maps, so their restrictions to U are either isomorphisms or 0 by Proposition 5.2.7. But since the image contains v or w, respectively, they cannot be 0, and so we have isomorphisms

$$V \xleftarrow{\sim} U \xrightarrow{\sim} W,$$

which in particular implies that $V \cong W$. □

[2] In fact, since $\mathfrak{sl}_3 \cong \mathfrak{su}_3 \otimes \mathbb{C}$ and SU_3 is compact and simply connected, it follows from Theorem 5.3.10 that *all* finite-dimensional complex representations of $\mathfrak{sl}_3$ are completely reducible, but we do not need this assumption.

Proposition 6.3.10. *Let (V, ρ) be an irreducible representation of $\mathfrak{sl}_3$ with highest weight $\mu = (a_1, a_2)$. Then μ is dominant integral, i.e. a_i is a non-negative integer for $i = 1, 2$.*

Proof. We already saw in Lemma 6.2.3 that the a_i are integers. If v is a highest weight vector with weight μ then we must have $\rho(X_i)v = 0$ for $i = 1, 2$. Applying Proposition 6.1.3 to the two copies of $\mathfrak{sl}_2$ in $\mathfrak{sl}_3$ we conclude that a_i must be non-negative. □

Remark 6.3.11. Note that in a sense we have hardly used anything specific about $\mathfrak{sl}_3$ in this section. One of our goals in the next chapter will be to identify a general class of complex Lie algebras for which exactly the same arguments go through to give a classification of their irreducible representations in terms of "highest weights".

Exercise 6.8. Show that the adjoint representation of $\mathfrak{sl}_3$ on itself is irreducible with highest weight $(1, 1)$.

Exercise 6.9. Show that the standard representation of $\mathfrak{sl}_3$ on $\mathbb{C}^3$ is irreducible with highest weight $(1, 0)$.

6.4 Tensor products and irreducible $\mathfrak{sl}_3$-representations

Our goal in this section is to show that for every dominant integral weight, there exists a (necessarily unique) irreducible $\mathfrak{sl}_3$-representation with that as its highest weight. To do this, we will show that for any such weight μ there exists a completely reducible representation V with μ as its highest weight; the smallest invariant subspace of V that contains a highest weight vector for μ is then highest weight cyclic and completely reducible, and so irreducible by Proposition 6.3.8[3]. In order to construct these representations we need to introduce two general algebraic constructions for making new representations, namely tensor products and duals.

We begin by defining tensor products of group representations (see §A.3 for a review of tensor products of vector spaces):

Definition 6.4.1. Let G and H be matrix groups. Suppose (U, ρ) and (V, σ) are representations of G and H, respectively. Then we define the representation $(U \otimes V, \rho \boxtimes \sigma)$ of $G \times H$ by setting

$$(\rho \boxtimes \sigma)(g, h) := \rho(g) \otimes \sigma(h)$$

for $g \in G$, $h \in H$.

[3] One can also give completely explicit descriptions of these representations.

Exercise 6.10. Use Exercise A.5 to show that for $\mathbb{K}$-vector spaces U and V, taking tensor products of linear maps yields a bilinear map

$$\mathrm{End}(U) \times \mathrm{End}(V) \to \mathrm{End}(U \otimes V),$$

given by $(f, g) \mapsto f \otimes g$, and this is compatible with composition. Conclude that (for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) this restricts to a continuous homomorphism

$$\mathrm{GL}(U) \times \mathrm{GL}(V) \to \mathrm{GL}(U \otimes V),$$

and that for representations $\rho\colon G \to \mathrm{GL}(U)$, $\sigma\colon H \to \mathrm{GL}(V)$, our definition of $\rho \boxtimes \sigma$ corresponds to the composite

$$G \times H \to \mathrm{GL}(U) \times \mathrm{GL}(V) \to \mathrm{GL}(U \otimes V);$$

in particular, $\rho \boxtimes \sigma$ is indeed a representation.

Definition 6.4.2. Let G be a matrix group. Suppose (U, ρ) and (V, σ) are two representations of G. Then we define the representation $(U \otimes V, \rho \otimes \sigma)$ of G by setting

$$(\rho \otimes \sigma)(g) := \rho(g) \otimes \sigma(g)$$

for $g \in G$.

Remark 6.4.3. The representation $\rho \otimes \sigma$ of G above is by definition obtained by restricting the representation $\rho \boxtimes \sigma$ of $G \times G$ along the diagonal $G \to G \times G$. In particular, $\rho \otimes \sigma$ is indeed a representation.

Lemma 6.4.4. *Suppose U and V are $\mathbb{K}$-vector spaces for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and let $\mu\colon \mathrm{GL}(U) \times \mathrm{GL}(V) \to \mathrm{GL}(U \otimes V)$ be the continuous homomorphism given by $\mu(f, g) = f \otimes g$. Then $\mathfrak{L}(\mu)\colon \mathfrak{gl}(U) \oplus \mathfrak{gl}(V) \to \mathfrak{gl}(U \otimes V)$ is given by*

$$\mathfrak{L}(\mu)(X, Y) = X \otimes I + I \otimes Y.$$

Proof. Suppose we have a smooth curve $u(t)$ in U and $v(t)$ in V. By calculating in a basis, we get the following product rule for differentiation:

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t) \otimes v(t) = u'(t) \otimes v(t) + u(t) \otimes v'(t).$$

Applying this to $e^{tX}(u)$ and $e^{tY}(v)$ for $X \in \mathfrak{gl}(U), Y \in \mathfrak{gl}(V), u \in U$ and $v \in V$, we get

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{tX}(u) \otimes e^{tY}(v) = Xe^{tX}(u) \otimes e^{tY}(v) + e^{tX}(u) \otimes Ye^{tY}(v),$$

and so

$$\mathfrak{L}(\mu)(X, Y)(u, v) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{tX}(u) \otimes e^{tY}(v) = Xu \otimes v + u \otimes Yv,$$

as required. □

Lemma 6.4.4 suggests the following definitions of tensor products of Lie algebra representations:

Definition 6.4.5. Let $\mathfrak{g}$ and $\mathfrak{h}$ be $\mathbb{K}$-Lie algebras (for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$), and let (U, ρ) and (V, σ) be representations of $\mathfrak{g}$ and $\mathfrak{h}$, respectively. The representation $(U \otimes V, \rho \boxtimes \sigma)$ of $\mathfrak{g} \oplus \mathfrak{h}$ is defined by

$$(\rho \boxtimes \sigma)(X, Y) = \rho(X) \otimes I + I \otimes \sigma(Y).$$

Similarly, if (U, ρ) and (V, σ) are both representations of $\mathfrak{g}$, then the representation $(U \otimes V, \rho \otimes \sigma)$ of $\mathfrak{g}$ is defined by

$$(\rho \otimes \sigma)(X) = \rho(X) \otimes I + I \otimes \sigma(X).$$

We then have the following immediate consequence of Lemma 6.4.4:

Lemma 6.4.6. *Let G and H be matrix groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$.*

(i) If (U, ρ) and (V, σ) are representations of G and H, respectively, then the induced Lie algebra representation of $\mathfrak{g} \oplus \mathfrak{h}$ from $\rho \boxtimes \sigma$ is

$$\mathfrak{L}(\rho \boxtimes \sigma) = \mathfrak{L}(\rho) \boxtimes \mathfrak{L}(\sigma).$$

(ii) If (U, ρ) and (V, σ) are representations of G, then the induced Lie algebra representation of $\mathfrak{g}$ from $\rho \otimes \sigma$ is

$$\mathfrak{L}(\rho \otimes \sigma) = \mathfrak{L}(\rho) \otimes \mathfrak{L}(\sigma).$$

We next want to define dual representations, for which we make use of the following definition and calculation:

Exercise 6.11. Show that for (finite-dimensional) $\mathbb{K}$-vector spaces U and V (with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) there is a continuous homomorphism

$$\alpha\colon \mathrm{GL}(U) \times \mathrm{GL}(V) \to \mathrm{GL}(\mathrm{Hom}(U, V))$$

given by

$$\alpha(f, g)(\phi) = g \circ \phi \circ f^{-1}$$

for $f \in \mathrm{GL}(U)$, $g \in \mathrm{GL}(V)$, $\phi \in \mathrm{Hom}(U, V)$. (Note that we do not get a homomorphism if we don't take the inverse of f.)

Lemma 6.4.7. *For U, V and α as above, the Lie algebra homomorphism*

$$\mathfrak{L}(\alpha)\colon \mathfrak{gl}(U) \oplus \mathfrak{gl}(V) \to \mathfrak{gl}(\mathrm{Hom}(U, V))$$

is given by

$$\mathfrak{L}(\alpha)(X, Y)(\phi) = Y \circ \phi - \phi \circ X.$$

Proof. From the chain rule, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{tY} \circ \phi \circ e^{-tX} = Ye^{tY} \circ \phi \circ e^{-tX} + e^{tY} \circ \phi \circ (-X)e^{-tX},$$

so that

$$\mathfrak{L}(\alpha)(X,Y)(\phi) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} e^{tY} \circ \phi \circ e^{-tX} = Y \circ \phi - \phi \circ X.$$

□

Exercise 6.12. If (U, ρ) and (V, σ) are representations of G and H, respectively, show that there is a representation $\mathrm{Hom}(\rho, \sigma)$ of $G \times H$ on $\mathrm{Hom}(U, V)$ given by

$$\mathrm{Hom}(\rho, \sigma)(g, h)(\phi) = \sigma(h) \circ \phi \circ \rho(g)^{-1}.$$

Find the induced Lie algebra representation.

Definition 6.4.8. Given a $\mathbb{K}$-representation (U, ρ) of a matrix group G, the *dual* representation ρ^* on $U^* := \mathrm{Hom}(U, \mathbb{K})$ is defined by

$$\rho^*(g)(\phi) = \phi \circ \rho(g)^{-1} = \left(\rho(g)^{-1}\right)^* \phi.$$

Similarly, if (U, ρ) is a representation of a Lie algebra $\mathfrak{g}$, the *dual* representation ρ^* on U^* is defined by

$$\rho^*(X)(\phi) = -\phi \circ X = -X^* \phi.$$

Then $\mathfrak{L}(\rho^*) = \mathfrak{L}(\rho)^*$ for ρ a matrix group representation.

Remark 6.4.9. Suppose (V, ρ) is a $\mathbb{K}$-representation of a matrix group G. If we pick a basis for V, we can identify ρ with a homomorphism $R \colon G \to \mathrm{GL}_n(\mathbb{K})$ (where $n = \dim_{\mathbb{K}} V$). In the dual basis of V^*, the representation ρ^* corresponds to $g \mapsto \left(R(g)^{-1}\right)^T$. Similarly, if ρ is instead a representation of a Lie algebra $\mathfrak{g}$, then ρ^* corresponds to $X \mapsto -R(X)^T$.

Proposition 6.4.10. *Suppose (U, ρ) and (V, σ) are unitary representations of a matrix group G or a Lie algebra $\mathfrak{g}$.*

1. *The tensor product representation $(U \otimes V, \rho \otimes \sigma)$ is also unitary.*

2. *The dual representation (U^*, ρ^*) is also unitary.*

Exercise 6.13. Suppose U and V are inner product spaces over $\mathbb{K}$, for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Show that if we set

$$\langle u \otimes v, u' \otimes v' \rangle := \langle u, u' \rangle \langle v, v' \rangle$$

for $u, u' \in U$, $v, v' \in V$ and extend this linearly, then we get an inner product on $U \otimes V$.

Proof of Proposition 6.4.10. We consider the group case; the Lie algebra case is proved similarly. Using the inner product on $U \otimes V$ from Exercise 6.13, we have for $u, u' \in U$, $v, v' \in V$, $g \in G$ that

$$\begin{aligned} \langle (\rho \otimes \sigma)(g)(u \otimes v), (\rho \otimes \sigma)(g)(u' \otimes v') \rangle &= \langle \rho(g)(u) \otimes \sigma(g)(v), \rho(g)(u') \otimes \sigma(g)(v') \rangle \\ &= \langle \rho(g)(u), \rho(g)(u') \rangle \langle \sigma(g)v, \sigma(g)(v') \rangle \\ &= \langle u, u' \rangle \langle v, v' \rangle \\ &= \langle u \otimes v, u' \otimes v' \rangle. \end{aligned}$$

Extending this linearly, we see that $\rho \otimes \sigma$ is unitary.

In the case of the dual, suppose we pick a basis for U so that $\rho(g)$ is represented by a matrix $R(g)$; the representation ρ is unitary if and only if the matrices $R(g)$ are unitary. If we take the dual basis on U^*, then $\rho^*(g)$ is represented by the matrix $R(g^{-1})^T$. It is clear that the transpose of a unitary matrix is again unitary (and in fact we have $R(g^{-1})^T = \overline{R(g)}$), so this is again a unitary representation. □

We are now ready to return to our study of $\mathfrak{sl}_3$. Recall from Exercise 6.9 that the standard representation of $\mathfrak{sl}_3$ on $\mathbb{C}^3$ is irreducible with highest weight $(1,0)$.

Exercise 6.14. Show that the dual of the standard representation is irreducible with highest weight $(0,1)$. (Use that the dual representation is given by $X \in \mathfrak{sl}_3$ acting on $\mathbb{C}^3$ as $-X^T$.)

Exercise 6.15. Check that both the standard representation of $\mathfrak{sl}_3$ and its dual are unitary when restricted to representations of the real Lie algebra $\mathfrak{su}_3$.

Exercise 6.16. Suppose U and V are representations of $\mathfrak{sl}_3$, and that $u \in U$ and $v \in V$ are weight vectors for weights μ, ν, respectively. Show that $u \otimes v$ is a weight vector for $\mu + \nu$ in $U \otimes V$.

Proposition 6.4.11. *Let μ be a dominant integral weight for $\mathfrak{sl}_3$, i.e.$\mu = (a,b)$ where a and b are non-negative integers. Then there exists a finite-dimensional irreducible complex $\mathfrak{sl}_3$-representation whose highest weight is μ.*

Proof. Let V denote the standard representation of $\mathfrak{sl}_3$ and v a highest weight vector with weight $(1,0)$. By Exercise 6.14 the dual representation V^* is irreducible with highest weight $(0,1)$; let w be a corresponding weight vector. Now consider the iterated tensor product $V^{\otimes a} \otimes (V^*)^{\otimes b}$. By Exercise 6.16 it contains a weight vector $v_{a,b} := v^{\otimes a} \otimes w^{\otimes b}$ with weight (a,b). It is also clear from the formulas for tensor products of Lie algebra representations that $X_i(v_{a,b}) = 0$ for $i = 1,2$, so if we let $V_{a,b}$ be the smallest invariant subspace containing $v_{a,b}$ then this is a highest weight cyclic representation for μ. By Proposition 6.3.8, this representation is irreducible provided $V_{a,b}$ is completely reducible. This does follows from Corollary 5.3.11 since SU_3 is compact and simply connected, but we do not need to appeal to this result: Instead, first observe that it suffices by Lemma 5.3.7 to show that the tensor product $V^{\otimes a} \otimes (V^*)^{\otimes b}$ is completely reducible, and since $\mathfrak{sl}_3 \cong \mathfrak{su}_3 \otimes \mathbb{C}$ we can equivalently show that this is completely reducible as an $\mathfrak{su}_3$-representation. But V and V^* are unitary as representations of $\mathfrak{su}_3$ by Exercise 6.15, hence so is the iterated tensor product by Proposition 6.4.10, and it is then completely reducible by Proposition 5.3.8. □

Remark 6.4.12. We have shown that for any $a, b \in \mathbb{N}$ there exists a unique irreducible representation $V_{a,b}$ of $\mathfrak{sl}_3$ with highest weight (a, b). There are some obvious questions we might ask about these representations:

- What is the dimension of $V_{a,b}$?
- Which other weights occur in $V_{a,b}$?
- What are the multiplicities of these weights?

The answers to all three questions are known, but we will not say anything further at this point.

Chapter 7

Complex semisimple Lie algebras and their representations

7.1 Simple and semisimple Lie algebras

Our goal is to extend the classification of representations we worked out for $\mathfrak{sl}_3$ to a more general class of Lie algebras: the complex *semisimple* Lie algebras. In this section we introduce this notion from an algebraic perspective and look at some examples.

Definition 7.1.1. Let $\mathfrak{g}$ be a finite-dimensional $\mathbb{K}$-Lie algebra ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$). An *ideal* of $\mathfrak{g}$ is a $\mathbb{K}$-subspace $\mathfrak{h} \subseteq \mathfrak{g}$ such that $[X, H] \in \mathfrak{h}$ for all $X \in \mathfrak{g}$ and $H \in \mathfrak{h}$. (Note that if $\mathfrak{h} \subseteq \mathfrak{g}$ is an ideal, then $\mathfrak{h}$ is in particular closed under Lie brackets, and so is a Lie subalgebra.) We say that $\mathfrak{g}$ is *irreducible* if its only ideals are 0 and $\mathfrak{g}$, and *simple* if it is irreducible and not abelian (or equivalently, if $\dim_{\mathbb{K}} \mathfrak{g} > 1$).

Remark 7.1.2. An ideal in $\mathfrak{g}$ is precisely a $\mathfrak{g}$-invariant subspace for the adjoint representation of $\mathfrak{g}$ on itself. Thus $\mathfrak{g}$ is an irreducible Lie algebra if and only if the adjoint representation is irreducible.

Example 7.1.3. The adjoint representations of $\mathfrak{sl}_2$ and $\mathfrak{sl}_3$ are irreducible by Exercises 6.3 and 6.8, respectively. Thus the Lie algebras $\mathfrak{sl}_2$ and $\mathfrak{sl}_3$ are both simple, since their dimensions are greater than 1.

Definition 7.1.4. A Lie algebra $\mathfrak{g}$ is *reductive* if there is a an isomorphism of Lie algebras

$$\mathfrak{g} \cong \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$$

where each $\mathfrak{g}_i$ is irreducible. If each $\mathfrak{g}_i$ is in fact simple, we say that $\mathfrak{g}$ is *semisimple*.

Exercise 7.1. Show that $\mathfrak{gl}_n$ is the direct sum of Lie algebras $\mathbb{C} \cdot I \oplus \mathfrak{sl}_n$. Conclude (using Proposition 7.1.11) that $\mathfrak{gl}_n$ is reductive, but not semisimple.

Lemma 7.1.5. *A Lie algebra $\mathfrak{g}$ is reductive if and only if the adjoint representation is completely reducible.*

Proof. First suppose $\mathfrak{g}$ is reductive, so that $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$ where $\mathfrak{g}_i$ is irreducible. Then each $\mathfrak{g}_i$ is an ideal: The direct sum decomposition of $\mathfrak{g}$ as a Lie algebra means that $[X, Y] = 0$ if $X \in \mathfrak{g}_i$ and $Y \in \mathfrak{g}_j$ with $i \neq j$. Thus if $X = \sum X_i$ with $X_i \in \mathfrak{g}_i$ and $Y \in \mathfrak{g}_j$ then $[X, Y] = [X_j, Y] \in \mathfrak{g}_j$. The direct sum decomposition therefore gives a decomposition of the adjoint representation as a sum of representations. Moreover, each $\mathfrak{g}_i$ is irreducible as a $\mathfrak{g}$-representation since an invariant subspace would in particular be an ideal of $\mathfrak{g}_i$.

Now suppose the adjoint representation is completely reducible, so that

$$\mathfrak{g} = \mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_n$$

where $\mathfrak{h}_i$ is an irreducible representation of $\mathfrak{g}$. We claim this gives a decomposition of $\mathfrak{g}$ as a direct sum of Lie algebras. Indeed, if $X \in \mathfrak{h}_i$ and $Y \in \mathfrak{h}_j$ then $[X, Y] \in \mathfrak{h}_j$ since $\mathfrak{h}_j$ is an ideal, but we also have $[X, Y] = -[Y, X]$ which lies in $\mathfrak{h}_i$ as this is an ideal. Thus $[X, Y] \in \mathfrak{h}_i \cap \mathfrak{h}_j$, which is 0 if $i \neq j$. Moreover, if $X = \sum X_i$ with $X_i \in \mathfrak{h}_i$ and $Y \in \mathfrak{h}_j$, then $[X, Y] = [X_j, Y]$, which means that any ideal of $\mathfrak{h}_j$ is also $\mathfrak{g}$-invariant. Thus $\mathfrak{h}_j$ must be an irreducible Lie algebra as it is an irreducible $\mathfrak{g}$-representation. □

Definition 7.1.6. If $\mathfrak{g}$ is a Lie algebra, its *centre* $\mathfrak{z}(\mathfrak{g})$ is the subspace of $\mathfrak{g}$ consisting of elements Z such that $[X, Z] = 0$ for all $X \in \mathfrak{g}$. Note that $\mathfrak{z}(\mathfrak{g})$ is always an ideal of $\mathfrak{g}$.

Observation 7.1.7. If $\mathfrak{g}$ is an $\mathbb{R}$-Lie algebra then $\mathfrak{z}(\mathfrak{g}) \otimes \mathbb{C} \cong \mathfrak{z}(\mathfrak{g} \otimes \mathbb{C})$: If $Z \in \mathfrak{z}(\mathfrak{g})$ then $[Z, X + iY] = [Z, X] + i[Z, Y] = 0$, so $\mathfrak{z}(\mathfrak{g}) \otimes \mathbb{C} \subseteq \mathfrak{z}(\mathfrak{g} \otimes \mathbb{C})$. Conversely if $Z \in \mathfrak{z}(\mathfrak{g} \otimes \mathbb{C})$ and $Z = X + iY$ with $X, Y \in \mathfrak{g}$, then for $V \in \mathfrak{g}$ we have $[Z, V] = [X, V] + i[Y, V] = 0$ where $[X, V]$ and $[Y, V]$ lie in $\mathfrak{g}$. Since we have a direct sum decomposition of $\mathfrak{g} \otimes \mathbb{C}$ as $\mathfrak{g} \oplus i\mathfrak{g}$, this can only happen if $[X, V] = [Y, V] = 0$ for all $V \in \mathfrak{g}$, so we must have $X, Y \in \mathfrak{z}(\mathfrak{g})$, and hence $Z \in \mathfrak{z}(\mathfrak{g}) \otimes \mathbb{C}$.

Lemma 7.1.8. *A reductive Lie algebra $\mathfrak{g}$ is semisimple if and only if $\mathfrak{z}(\mathfrak{g}) = 0$.*

Proof. Suppose $\mathfrak{g}$ is reductive. Since $\mathfrak{z}(\mathfrak{g})$ is an ideal, we can write $\mathfrak{g}$ as $\mathfrak{z}(\mathfrak{g}) \oplus \mathfrak{g}'$ where $\mathfrak{g}'$ decomposes as a sum of irreducible Lie algebras. If $\mathfrak{g}$ is semisimple then we must have $\mathfrak{z}(\mathfrak{g}) = 0$, or we would have a decomposition as a direct sum of irreducible Lie algebras that are *not* all simple. Conversely, if $\mathfrak{z}(\mathfrak{g}) = 0$ then no irreducible piece in the direct sum decomposition can be abelian, as otherwise it would lie in $\mathfrak{z}(\mathfrak{g})$. □

Proposition 7.1.9. *If $\mathfrak{g}$ is a complex semisimple Lie algebra, then the decomposition*

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$$

of $\mathfrak{g}$ as a direct sum of simple Lie algebras $\mathfrak{g}_i$ is unique up to reordering. (This means the simple subalgebras $\mathfrak{g}_i$ are unique up to equality*, not just isomorphism!)*

Proof. We first note that as representations of $\mathfrak{g}$, the ideals $\mathfrak{g}_j$ are non-isomorphic: the action of $\mathfrak{g}_k$ on $\mathfrak{g}_j$ is trivial for $k \neq j$, but the action of $\mathfrak{g}_j$ on itself is non-trivial, since it is not commutative. Suppose $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ that is a simple Lie algebra, and so in particular irreducible as a $\mathfrak{g}$-representation. For any j, the projection map $\pi_j \colon \mathfrak{g} \to \mathfrak{g}_j$ is an intertwining map for the adjoint representation. By Schur's Lemma, it follows that $\pi_j|_{\mathfrak{h}} \colon \mathfrak{h} \to \mathfrak{g}_j$ must be 0 or an isomorphism. Given that $\mathfrak{h} \neq 0$ there must be some index j such that $\pi_j|_{\mathfrak{h}} \neq 0$. But the representations $\mathfrak{g}_j$ are non-isomorphic for different values of j, so $\pi_j|_{\mathfrak{h}}$ can't give an isomorphism for more than one index j. Thus $\pi_j|_{\mathfrak{h}}$ is an isomorphism for a single j and zero for all other indices, which means that $\mathfrak{h} = \mathfrak{g}_j$. □

Remark 7.1.10. The conclusion of Proposition 7.1.9 is false if we consider a complex Lie algebra $\mathfrak{g}$ that is reductive, but not semisimple. For example, a 2-dimensional commutative Lie algebra can be decomposed as a sum of 1-dimensional subalgebras in many ways (since this just amounts to choosing a basis).

Proposition 7.1.11. *The complex Lie algebra $\mathfrak{sl}_n$ is simple for all n.*

Proof. Let E_{ij} denote the $n \times n$ matrix with 1 in position (i, j) as its single non-zero entry. Then the matrices E_{ij} for $i \neq j$ and $E_{ii} - E_{nn}$ for $i = 1, \ldots, n-1$ form a basis for the vector space $\mathfrak{sl}_n$, which consists of traceless $n \times n$ matrices. If X is any $n \times n$ matrix, the matrix product XE_{ij} has all columns 0 except the jth one, which equals the ith column of X, while $E_{ij}X$ has all rows zero except the ith one, which equals the jth row of X. Thus the commutator $[X, E_{ij}] = XE_{ij} - E_{ij}X$ is the matrix

$$\begin{pmatrix} & & & X_{1i} & & & \\ & 0 & & \vdots & & 0 & \\ & & & X_{(i-1)i} & & & \\ -X_{j1} & \cdots & -X_{j(j-1)} & X_{ii} - X_{jj} & -X_{j(j+1)} & \cdots & -X_{jn} \\ & & & X_{(i+1)i} & & & \\ & 0 & & \vdots & & 0 & \\ & & & X_{ni} & & & \end{pmatrix}$$

As a special case, we have

$$[E_{ij}, E_{jk}] = \begin{cases} E_{ik}, & i \neq k, \\ E_{ii} - E_{jj}, & i = k. \end{cases}$$

From this we get for $i \neq j$ that

$$[[E_{ij}, E_{ji}], E_{ji}] = [E_{ii} - E_{jj}, E_{ji}] = -[E_{ji}, E_{ii}] - [E_{jj}, E_{ji}] = -2E_{ji}.$$

More generally, we can apply the calculation above twice to compute that for $i \neq j$ we have

$$[[X, E_{ij}], E_{ij}] = -2X_{ji}E_{ij}.$$

Now suppose $\mathfrak{h}$ is an ideal of $\mathfrak{sl}_n$ and that $X \neq 0$ lies in $\mathfrak{h}$. If the entry X_{ij} is non-zero for some $i \neq j$, then these computations imply:

(1) $E_{ij} \in \mathfrak{h}$ since $[[X, E_{ij}], E_{ij} = -2X_{ji}E_{ij}$ is in $\mathfrak{h}$,

(2) $E_{ik} \in \mathfrak{h}$ for all $k \neq i$ since $[E_{ij}, E_{jk}] = E_{ik}$,

(3) $E_{ki} \in \mathfrak{h}$ for all $k \neq i$ since $[[E_{ik}, E_{ki}], E_{ki}] = -2E_{ki}$,

(4) $E_{kl} \in \mathfrak{h}$ for all $k \neq l$ since if $k, j \neq l$ we have $[E_{ki}, E_{il}] = E_{kl}$.

(5) $E_{kk} - E_{nn} \in \mathfrak{h}$ for all $1 \leq k < n$ since $[E_{kn}, E_{nk}] = E_{kk} - E_{nn}$.

Thus all the basis vectors lie in $\mathfrak{h}$, which means $\mathfrak{h} = \mathfrak{sl}_n$.

The remaining case is where $X_{ij} = 0$ for all $i \neq j$, so that X is diagonal. There must be indices $i \neq j$ such that $X_{ii} \neq X_{jj}$, since if they were all equal the trace of $X \neq 0$ would be non-zero. For such an X the computation above gives

$$[X, E_{ij}] = (X_{ii} - X_{jj})E_{ij},$$

which implies that $E_{ij} \in \mathfrak{h}$. Now the same argument as before shows that $\mathfrak{h} = \mathfrak{sl}_n$. □

Exercise 7.2.

(i) Prove that a 4×4 skew-symmetric matrix (over $\mathbb{K} = \mathbb{R}, \mathbb{C}$) can be uniquely written as a sum of the form

$$\begin{pmatrix} 0 & -a & -b & -c \\ a & 0 & -c & b \\ b & c & 0 & -a \\ c & -b & a & 0 \end{pmatrix} + \begin{pmatrix} 0 & -x & -y & -z \\ x & 0 & z & -y \\ y & -z & 0 & x \\ z & y & -x & 0 \end{pmatrix}.$$

(ii) Use this to write down a natural basis for $\mathfrak{so}_4(\mathbb{K})$ and compute the commutators.

(iii) Conclude that as a Lie algebra $\mathfrak{so}_4(\mathbb{K})$ is the direct sum $\mathfrak{so}_3(\mathbb{K}) \oplus \mathfrak{so}_3(\mathbb{K})$. In particular, we have $\mathfrak{so}_4 \cong \mathfrak{sl}_2 \oplus \mathfrak{sl}_2$ (as $\mathfrak{so}_3 \cong \mathfrak{so}_3(\mathbb{R}) \otimes \mathbb{C} \cong \mathfrak{su}_2 \otimes \mathbb{C} \cong \mathfrak{sl}_2$), and so $\mathfrak{so}_4$ is semisimple, but not simple.

Remark 7.1.12. The isomorphism of $\mathfrak{so}_4(\mathbb{R})$ with $\mathfrak{su}_2 \oplus \mathfrak{su}_2$ comes from a double cover of $SO_4(\mathbb{R})$ by $SU_2 \times SU_2$. This can be described in terms of quaternions: if we identify $\mathbb{R}^4$ with $\mathbb{H}$ as a vector space, then every rotation of $\mathbb{R}^4$ is of the form $v \mapsto q^{-1}vq'$ where q, q' are unit quaternions (corresponding to elements of SU_2). See [4, §2.7] for more details.

Remark 7.1.13. One can also show by explicit matrix calculations that the Lie algebras $\mathfrak{so}_n$ (for $n > 4$) and $\mathfrak{sp}_n$ are simple. This is a bit more involved than for $\mathfrak{sl}_n$, however; see [4, §6.5–6] for the details.

7.2 Semisimplicity and compact groups

In this section we look at the relation between reductive (and semisimple) Lie algebras and (simply connected) compact groups.

Proposition 7.2.1. *Suppose K is a compact matrix group with Lie algebra $\mathfrak{k}$. Then the complex Lie algebra $\mathfrak{g} := \mathfrak{k} \otimes \mathbb{C}$ is reductive.*

Proof. The adjoint representation of $\mathfrak{g}$ lifts to a representation of K, namely the complexification of the adjoint action of K on $\mathfrak{k}$. This representation of K is completely reducible by Theorem 5.3.10. We may assume without loss of generality that K is connected (since the identity component of a compact group is again compact, and its Lie algebra is the same). Then Corollary 5.3.6 shows that $\mathfrak{g}$ is completely reducible as a representation of K if and only if it is completely reducible as a representation of $\mathfrak{g}$. □

In fact, the converse is also true:

Theorem 7.2.2. *A complex Lie algebra $\mathfrak{g}$ is reductive if and only if there exists a compact matrix group K with Lie algebra $\mathfrak{k}$ such that $\mathfrak{g} \cong \mathfrak{k} \otimes \mathbb{C}$.* □

In the situation of this theorem, the real Lie algebra $\mathfrak{k}$ is called a *compact real form* of the complex Lie algebra $\mathfrak{g}$. Proving the existence of such a compact real form is beyond the scope of this course, but we will feel free to make use of this characterization when it simplifies definitions and arguments. In particular, we will make frequent use of the following consequence:

Proposition 7.2.3. *Let $\mathfrak{g}$ be a reductive complex Lie algebra, and suppose $\mathfrak{g} \cong \mathfrak{k} \otimes \mathbb{C}$ where $\mathfrak{k}$ is the Lie algebra of a compact group K. Then there exists an inner product on $\mathfrak{g}$ that is real-valued on $\mathfrak{k}$ and such that the adjoint action of $\mathfrak{k}$ on $\mathfrak{g}$ is unitary in the sense that*

$$\langle [X,Y], Z\rangle = -\langle Y, [X,Z]\rangle$$

for $X \in \mathfrak{k}$ and $Y, Z \in \mathfrak{g}$.

Proof. By the real analogue of Theorem 5.3.10, we can define a (real) inner product on $\mathfrak{k}$ that is invariant under the adjoint action of K. On the Lie algebra level, this means that the adjoint representation of $\mathfrak{k}$ on itself is skew-symmetric, in the sense that

$$\langle [X,Y], Z\rangle = -\langle Y, [X,Z]\rangle$$

for $X, Y, Z \in \mathfrak{k}$. This inner product extends to a complex inner product on $\mathfrak{g}$ for which the adjoint action of $\mathfrak{k}$ is unitary. □

Observation 7.2.4. Suppose $\mathfrak{g}$ and $\mathfrak{k}$ are as in Proposition 7.2.3, and assume that $\mathfrak{g}$ has an inner product for which $[X, -]$ is unitary when $X \in \mathfrak{k}$. For $X \in \mathfrak{g}$ given as $X_1 + iX_2$ with $X_1, X_2 \in \mathfrak{k}$, we define $X^* := -X_1 + iX_2$. Then

$$\langle [X,Y], Z\rangle = \langle Y, [X^*, Z]\rangle$$

for all $Y, Z \in \mathfrak{g}$, since we have[1]

$$\begin{aligned}\langle [X,Y], Z\rangle &= \langle [X_1, Y], Z\rangle - i\langle [X_2, Y], Z\rangle \\ &= -\langle Y, [X_1, Z]\rangle + i\langle Y, [X_2, Z]\rangle \\ &= \langle Y, [-X_1 + iX_2, Z]\rangle.\end{aligned}$$

Observation 7.2.5. From Observation 7.1.7 and Lemma 7.1.8 it follows that if K is a compact matrix group with Lie algebra $\mathfrak{k}$, then $\mathfrak{g} := \mathfrak{k} \otimes \mathbb{C}$ is semisimple if and only if $\mathfrak{z}(\mathfrak{k}) = 0$.

Proposition 7.2.6. *Suppose K is a simply connected compact matrix group with Lie algebra $\mathfrak{k}$. Then the complex Lie algebra $\mathfrak{g} := \mathfrak{k} \otimes \mathbb{C}$ is semisimple.*

For the proof we need the following consequence of Theorem 4.2.1:

Proposition 7.2.7. *Suppose G is a simply connected matrix group and that the Lie algebra $\mathfrak{g}$ of G is a direct sum $\mathfrak{g} \cong \mathfrak{g}_1 \oplus \mathfrak{g}_2$. Then there exist closed, simply connected subgroups G_1 and G_2 of G with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, such that*

$$G \cong G_1 \times G_2.$$

Proof. Let $\pi_1 : \mathfrak{g} \to \mathfrak{g}$ be Lie algebra homomorphism given by projection to the factor $\mathfrak{g}_1$. Since G is simply connected, there is a Lie algebra homomorphism $\Pi_1 : G \to G$ lifting π_1 by Theorem 4.2.1. Then $\ker \Pi_1$ is a closed subgroup of G with Lie algebra $\ker \pi_1 = \mathfrak{g}_2$ by Proposition 3.6.10. Let G_2 be the identity component of $\ker \Pi_1$; this is a closed, connected subgroup of G with Lie algebra $\mathfrak{g}_2$. Considering the projection to $\mathfrak{g}_2$, we similarly find a closed, connected subgroup G_1 with Lie algebra $\mathfrak{g}_1$.

Note that since π_1 is the identity on $\mathfrak{g}_1$ and G_1 is connected, the restriction $\Pi_1|_{G_1}$ must agree with the inclusion $G_1 \hookrightarrow G$ by Corollary 3.7.8. Moreover, since G is connected it also follows from Corollary 3.7.7 that the image of Π_1 is contained in G_1.

To show that G_1 is simply connected, consider a loop $\gamma : [0,1] \to G_1$. Since G is simply connected, there is a homotopy $h: [0,1] \times [0,1] \to G$ connecting γ to a constant loop. If we define $h' := \Pi_1 \circ h$, then this is a homotopy from γ to a constant loop, but this lies in G_1. Thus G_1 is simply connected, as is G_2 by the same argument.

By assumption elements of $\mathfrak{g}_1$ commute with those in $\mathfrak{g}_2$. It follows from Corollary 3.7.7 that the elements of G_1 commute with those in G_2, so that we have a continuous homomorphism $\Phi: G_1 \times G_2 \to G$ given by $\Phi(A, B) = AB$, whose associated Lie algebra homomorphism is an isomorphism. Since G is simply connected, there is a continuous homomorphism $\Psi: G \to G_1 \times G_2$ whose Lie algebra homomorphism is $\mathfrak{L}(\Phi)^{-1}$. Then Corollary 3.7.8 implies that Ψ is inverse to Φ, and so Φ is an isomorphism, as required. ◻

[1] Assuming $\langle -, - \rangle$ is linear in the *second* variable — otherwise some signs change, but the end result is the same.

Proof of Proposition 7.2.6. Let $\mathfrak{z} := \mathfrak{z}(\mathfrak{k})$. The adjoint representation of $\mathfrak{k}$ on itself is completely reducible (by the real version of Proposition 5.3.8), so we have a direct sum decomposition $\mathfrak{k} \cong \mathfrak{z} \oplus \mathfrak{k}'$.

By Proposition 7.2.7 we have that $K \cong Z \times K'$ where Z and K' are simply connected closed subgroups with Lie algebras $\mathfrak{z}$ and $\mathfrak{k}'$. Thus if $n = \dim \mathfrak{z}$, the groups Z and $\mathbb{R}^n$ are both simply connected matrix groups with Lie algebras isomorphic to $\mathfrak{z}$. Then Exercise 4.1 implies that there is a continuous isomorphism $Z \cong \mathbb{R}^n$. On the other hand, the group Z is compact, since it is a closed subset of the compact group K. This is a contradiction unless $n = 0$, so we must have $\mathfrak{z} = 0$.

Thus $\mathfrak{g}$ is reductive by Proposition 7.2.1 and by Observation 7.1.7 we have $\mathfrak{z}(\mathfrak{g}) \cong \mathfrak{z} \otimes \mathbb{C} = 0$. Thus $\mathfrak{g}$ is semisimple by Lemma 7.1.8. □

The converse is also true here, though in the simply connected case we have to extend from matrix groups to the more general setting of Lie groups:

Theorem 7.2.8. *The following are equivalent for a $\mathbb{C}$-Lie algebra $\mathfrak{g}$:*

(1) $\mathfrak{g}$ is semisimple.

(2) There exists a compact Lie group K whose Lie algebra $\mathfrak{k}$ has trivial centre, such that $\mathfrak{g} \cong \mathfrak{k} \otimes \mathbb{C}$.

(3) There exists a simply connected compact Lie group K with Lie algebra $\mathfrak{k}$, such that $\mathfrak{g} \cong \mathfrak{k} \otimes \mathbb{C}$.

Examples 7.2.9. From Proposition 7.2.6 we get several families of examples of complex semisimple Lie algebras:

(i) $\mathfrak{sl}_n \cong \mathfrak{su}_n \otimes \mathbb{C}$ and SU_n is simply connected and compact. Therefore $\mathfrak{sl}_n$ is semisimple for all n.

(ii) $\mathfrak{sp}_n \cong \mathfrak{usp}_n \otimes \mathbb{C}$ and USp_n is simply connected and compact. Therefore $\mathfrak{sp}_n$ is semisimple for all n.

(iii) $\mathfrak{so}_n \cong \mathfrak{so}_n(\mathbb{R}) \otimes \mathbb{C}$ and $\mathrm{SO}_n(\mathbb{R})$ is compact. Therefore $\mathfrak{so}_n$ is reductive for all n. In fact, $\mathfrak{so}_n$ is semisimple for $n \geq 3$, which can be shown by checking that its centre is trivial. (Alternatively, we can find the universal cover of $\mathrm{SO}_n(\mathbb{R})$, the *spin group* Spin_n, and prove that this is still compact.)

7.3 Cartan subalgebras

In our analysis of representations of $\mathfrak{sl}_3$, the subalgebra of diagonal matrices played a key role. Here we introduce the notion of a *Cartan subalgebra*, which formalizes the key properties this subalgebra has, and show that every semisimple complex Lie algebra has one.

Definition 7.3.1. Let $\mathfrak{g}$ be a complex Lie algebra. A *Cartan subalgebra* of $\mathfrak{g}$ is a $\mathbb{C}$-subspace $\mathfrak{h}$ such that:

(1) For all H, H' in $\mathfrak{h}$, we have $[H, H'] = 0$. (Thus $\mathfrak{h}$ is an abelian subalgebra of $\mathfrak{g}$).

(2) If we have $[H, X] = 0$ for all $H \in \mathfrak{h}$, then $X \in \mathfrak{h}$. (Thus $\mathfrak{h}$ is a *maximal* abelian subalgebra.)

(3) The endomorphism $\mathrm{ad}_{\mathfrak{g}}(H) = [H, -]$ of $\mathfrak{g}$ is diagonalizable for all $H \in \mathfrak{h}$.

Observation 7.3.2. Since the endomorphisms $\mathrm{ad}(H)$ for $H \in \mathfrak{h}$ commute, they are in fact *simultaneously* diagonalizable.

Remark 7.3.3. Any finite-dimensional Lie algebra has a (non-zero) maximal commutative subalgebra: any 1-dimensional subspace is a commutative subalgebra, and an increasing sequence of commutative subalgebras must necessarily terminate since their dimensions will increase. The key distinguishing property of a Cartan subalgebra is thus that the endomorphisms $\mathrm{ad}(H)$ are diagonalizable. In general, a finite-dimensional complex Lie algebra $\mathfrak{g}$ may not have any Cartan subalgebra, but we will see that $\mathfrak{g}$ has one provided it is semisimple.

Proposition 7.3.4. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra, and suppose K is a compact Lie group with Lie algebra $\mathfrak{k}$ such that $\mathfrak{g} \cong \mathfrak{k} \otimes \mathbb{C}$. If $\mathfrak{t}$ is a maximal commutative subalgebra of $\mathfrak{k}$, then $\mathfrak{t} \otimes \mathbb{C}$ is a Cartan subalgebra of $\mathfrak{g}$.*

Proof. The subalgebra $\mathfrak{h} := \mathfrak{t} \otimes \mathbb{C}$ of $\mathfrak{g}$ is clearly commutative. To see that it is maximal, suppose $X \in \mathfrak{g}$ satisfies $[T, X] = 0$ for $T \in \mathfrak{t}$. If $X = X_1 + iX_2$ with $X_1, X_2 \in \mathfrak{k}$, then we have

$$[T, X] = [T, X_1] + i[T, X_2] = 0.$$

Since these Lie brackets lie in $\mathfrak{k}$, we must have $[T, X_1] = [T, X_2] = 0$ for all $T \in \mathfrak{t}$. The maximality of $\mathfrak{t}$ then implies that X_1, X_2 must lie in $\mathfrak{t}$, so that X lies in $\mathfrak{h}$, as required.

To prove that $[H, -]$ is diagonalizable for $H \in \mathfrak{h}$, we consider an inner product as in Proposition 7.2.3. For T in $\mathfrak{t}$, the endomorphism $[T, -]$ of $\mathfrak{h}$ is skew-Hermitian with respect to this inner product, and therefore diagonalizable.[2] If $H = T_1 + iT_2$ with $T_1, T_2 \in$ then $[T_1, -]$ and $[T_2, -]$ are commuting diagonalizable endomorphisms, and therefore simultaneously diagonalizable. Then $[H, -] = [T_1, -] + i[T_2, -]$ is also diagonalizable, as required. □

Remark 7.3.5. One can also give a completely algebraic proof that a semisimple Lie algebra has a Cartan subalgebra, without appealing to the existence of a compact group K. This takes considerably more work to establish, however.

[2] This follows from the fact that Hermitian matrices are diagonalizable after multiplication by i.

Remark 7.3.6. We will only consider Cartan subalgebras of the form $\mathfrak{t} \otimes \mathbb{C}$ for $\mathfrak{t}$ a maximal commutative subalgebra in a compact real form. In fact, it can be shown that all Cartan subalgebras are of this form. Moreover, any two Cartan subalgebras are isomorphic under some automorphism of $\mathfrak{g}$, so that the following definition makes sense:

Definition 7.3.7. If $\mathfrak{g}$ is a complex semisimple Lie algebra, then the *rank* of $\mathfrak{g}$ is the dimension of a Cartan subalgebra.

Example 7.3.8. A Cartan subalgebra of $\mathfrak{sl}_n$ is given by the traceless diagonal matrices: If D is diagonal with entries $(\lambda_1, \dots, \lambda_n)$ then for any matrix X we have

$$[D, X]_{ij} = (\lambda_i - \lambda_j)X_{ij}.$$

In particular, using the basis from Proposition 7.1.11 consisting of E_{ij} $(i \neq j)$ and $E_{ii} - E_{nn}$ $(1 \leq i < n)$ we have

$$[D, E_{ij}] = (\lambda_i - \lambda_j)E_{ij}, \quad [D, E_{ii} - E_{nn}] = 0,$$

so D acts diagonally in this basis. The traceless diagonal matrices certainly form a commutative subalgebra of $\mathfrak{sl}_n$, so it remains to check that it is maximal. But if X satisfies $[D, X] = 0$ for any diagonal D, the computation above shows that we must have $X_{ij} = 0$ when $i \neq j$ (since we may choose D such that $\lambda_i \neq \lambda_j$) — in other words, X must be diagonal. If we think of $\mathfrak{sl}_n$ as $\mathfrak{su}_n \otimes \mathbb{C}$ then this Cartan subalgebra is $\mathfrak{t} \otimes \mathbb{C}$ where $\mathfrak{t}$ is the subalgebra of matrices of the form iD where D is a traceless *real* diagonal matrix, which it is easy to see is a maximal commutative subalgebra of $\mathfrak{su}_n$.

7.4 Roots and weights

Throughout this section we fix a semisimple complex Lie algebra $\mathfrak{g}$ with compact real form $\mathfrak{k}$, and a Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ of the form $\mathfrak{t} \otimes \mathbb{C}$ for $\mathfrak{t}$ a maximal commutative subalgebra of $\mathfrak{k}$. We can then extend the notion of *weights* to representations of $\mathfrak{g}$:

Definition 7.4.1. Suppose (V, ρ) is a finite-dimensional complex representation of $\mathfrak{g}$. An element $\lambda \in \mathfrak{h}^*$ is a *weight* of ρ if there exists $v \neq 0$ in V such that

$$\rho(H)v = \lambda(H) \cdot v$$

for all $H \in \mathfrak{h}$. In this case we call v a *weight vector* for λ. The *weight space* of λ is the subspace of V consisting of weight vectors for λ (including 0); the *multiplicity* of the weight λ is the dimension of its weight space.

Definition 7.4.2. A *root* of $\mathfrak{g}$ is a non-zero weight of the adjoint representation, i.e. an element $\alpha \in \mathfrak{h}^*$ such that there exists a non-zero $X \in \mathfrak{g}$ (a *root vector* for α) such that

$$[H, X] = \alpha(H) \cdot X$$

for all $H \in \mathfrak{h}$. The *root space* $\mathfrak{g}_\alpha$ of α is the vector subspace of $\mathfrak{g}$ consisting of all elements for which this equation holds (including 0).

Observation 7.4.3. Let us write $\mathfrak{g}_0$ also for the weight space of the 0 weight in the adjoint representation. This consists of those elements $X \in \mathfrak{g}$ such that $[H, X] = 0$ for $H \in \mathfrak{h}$. The maximality of $\mathfrak{h}$ then says that $\mathfrak{g}_0 = \mathfrak{h}$.

Proposition 7.4.4. *Suppose (V, ρ) is a finite-dimensional complex representation of $\mathfrak{g}$ and μ is a weight of ρ. If $\langle -, - \rangle$ is an inner product on $\mathfrak{g}$ as in Proposition 7.2.3, then $\mu = \langle M, - \rangle$ for a unique M, which lies in $i\mathfrak{t}$.*

Proof. The representation ρ of $\mathfrak{g}$ restricts to a representation $\rho|_{\mathfrak{k}}$ of $\mathfrak{k}$ (which determines ρ uniquely), and since $\mathfrak{k}$ is the Lie algebra of a compact group there exists an inner product on V such that this representation is unitary. In other words, we may assume that $\rho(X)$ is skew-Hermitian for every $X \in \mathfrak{k}$.

The inner product on $\mathfrak{g}$ identifies any element $\mu \in \mathfrak{h}^*$ with a unique linear functional of the form $\langle M, - \rangle$ where M is in $\mathfrak{h}$. If $X \in V$ is a weight vector for ρ, then $\langle M, T \rangle$ is an eigenvalue of $\rho(T)$ for $T \in \mathfrak{t}$, and since $\rho(T)$ is skew-Hermitian this eigenvalue must be pure imaginary.

We can write $M = M_1 + iM_2$ with $M_1, M_2 \in \mathfrak{t}$, so

$$\langle M_1, T \rangle + i \langle M_2, T \rangle$$

is pure imaginary for all $T \in \mathfrak{t}$. Since the inner product is real on $\mathfrak{t}$, this can only happen if $M = iM_2$, as required. □

Corollary 7.4.5. *Suppose α is a root of $\mathfrak{g}$. If X is a root vector for α, then X^* is a root vector for $-\alpha$. In particular, $-\alpha$ is also a root of $\mathfrak{g}$.*

Proof. Write $X = X_1 + iX_2$ with $X_1, X_2 \in \mathfrak{k}$. Then we compute for $T \in \mathfrak{t}$ that

$$[T, X^*] = [T, -X_1] + i[T, X_2] = [T, X]^*.$$

If X is a root vector for α, then we know from Proposition 7.4.4 that $\alpha(T)$ is pure imaginary, say $\alpha(T) = \lambda(T)i$ with $\lambda(T) \in \mathbb{R}$ for $T \in \mathfrak{t}$. Thus

$$[T, X] = \alpha(T)X = -\lambda(T)X_2 + \lambda(T)iX_1,$$

giving

$$[T, X^*] = [T, X]^* = \lambda(T)X_2 + \lambda(T)iX_1 = -\lambda(T)i \cdot (-X_1 + iX_2) = -\alpha(T)X^*.$$

By linearity, this identity extends to give

$$[H, X^*] = -\alpha(H)X^*$$

for all $H \in \mathfrak{h} = \mathfrak{t} \otimes \mathbb{C}$, as required. □

Proposition 7.4.6. *As a vector space, $\mathfrak{g}$ decomposes as a direct sum*

$$\mathfrak{g} \cong \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha,$$

where R is the set of roots.

Proof. By assumption, the endomorphisms $[H, -]$ are diagonalizable and commute. They are therefore simultaneously diagonalizable, and it is not hard to show that for any finite collection of simultaneously diagonalizable endomorphisms of a finite-dimensional vector space we get a decomposition as a direct sum of weight spaces. □

Corollary 7.4.7. *The roots span $\mathfrak{h}^*$.*

Proof. Suppose the roots do not span $\mathfrak{h}^*$. Then there would exist a non-zero $H \in \mathfrak{h}$ such that $\alpha(H) = 0$ for all roots α. But this H would then satisfy $[H, X] = \alpha(H)X = 0$ for $X \in \mathfrak{g}_\alpha$ and $[H, H'] = 0$ for all $H' \in \mathfrak{h}$. From the decomposition in Proposition 7.4.6 this implies that $[H, X] = 0$ for all $X \in \mathfrak{g}$, i.e. H is in the centre of $\mathfrak{g}$. But this is impossible since $\mathfrak{g}$ is semisimple. □

Proposition 7.4.8. *Let (V, ρ) be a representation of $\mathfrak{g}$. If $v \in V$ is a weight vector for a weight $\lambda \in \mathfrak{h}^*$ and $Z \in \mathfrak{g}$ is a root vector for a root α, then $\rho(Z)(v)$ is either 0 or a weight vector for $\lambda + \alpha$.*

Proof. Since $[\rho(H), \rho(Z)] = \rho([H, Z]) = \alpha(H)\rho(Z)$, we have

$$\rho(H)\rho(Z)(v) = \rho(Z)\rho(H)(v) + \alpha(H)\rho(Z)(v) = (\lambda(H) + \alpha(H))\rho(Z)(v),$$

as required. □

Corollary 7.4.9. *If $Z, Z' \in \mathfrak{g}$ are root vectors for $\alpha, \beta \in \mathfrak{h}^*$, respectively, then $[Z, Z']$ is either 0 or a root vector for $\alpha + \beta$. In other words, we have*

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$$

(where the right-hand side denotes $\mathfrak{h}$ if $\alpha + \beta = 0$ and 0 if $\alpha + \beta$ is not a root). □

Lemma 7.4.10. *Let α be a root.*

(i) Suppose $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$. Then $[X, Y]$ is in $\mathfrak{h}$ and satisfies

$$\langle [X, Y], H \rangle = \alpha(H) \langle Y, X^* \rangle.$$

(ii) Suppose $X \in \mathfrak{g}_\alpha$. Then

$$\alpha = \frac{\langle [X, X^*], - \rangle}{\|X^*\|^2}.$$

Proof. In part (i), we already know from Corollary 7.4.9 that $[X, Y]$ is in $\mathfrak{h} = \mathfrak{g}_0$. From Observation 7.2.4 and Corollary 7.4.5 we also know

$$\langle [X, Y], H\rangle = \langle Y, [X^*, H]\rangle = -\langle Y, [H, X^*]\rangle = -\langle Y, -\alpha(H)X^*\rangle = \alpha(H)\langle Y, X^*\rangle,$$

since X^* is a root vector for $-\alpha$. Putting $Y = X^*$ now gives part (ii). □

Theorem 7.4.11. *Let α be a root. Then there exists $X_\alpha \in \mathfrak{g}_\alpha$ such that if we set $Y_\alpha := X_\alpha^* \in \mathfrak{g}_{-\alpha}$ and $H_\alpha := [X_\alpha, X_\alpha^*]$ we have the relations*

$$[H_\alpha, X_\alpha] = 2X_\alpha, \quad [H_\alpha, Y_\alpha] = -2Y_\alpha, \quad [X_\alpha, Y_\alpha] = H_\alpha.$$

In other words, the subspace $\mathfrak{s}_\alpha$ spanned by these three elements is a subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2$. Moreover, $\langle H_\alpha, -\rangle$ is a multiple of α.

Proof. Choose a non-zero $X \in \mathfrak{g}_\alpha$. Then Lemma 7.4.10 shows that $\langle [X, X^*], -\rangle$ is a non-zero multiple of α. In particular, $[X, X^*] \neq 0$. Moreover, applying Lemma 7.4.10 with $H := [X, X^*]$ we get

$$\|H\|^2 = \alpha(H)\|X^*\|^2,$$

so $\alpha([X, X^*])$ must be real and > 0. Let us define

$$H_\alpha := \frac{2}{\alpha(H)}H, \quad X_\alpha := \sqrt{\frac{2}{\alpha(H)}}X, \quad Y_\alpha := X_\alpha^* = \sqrt{\frac{2}{\alpha(H)}}X^*.$$

Then we have

$$\begin{aligned}
[H_\alpha, X_\alpha] &= \alpha(H_\alpha)X_\alpha = 2X_\alpha, \\
[H_\alpha, Y_\alpha] &= -\alpha(H_\alpha)Y_\alpha = -2Y_\alpha, \\
[X_\alpha, Y_\alpha] &= \frac{2}{\alpha(H)}[X, X^*] = H_\alpha,
\end{aligned}$$

as required. Finally, we note that $X_\alpha, Y_\alpha, H_\alpha$ are linearly independent since they lie in weight spaces of $\mathfrak{g}$ for distinct weights. □

Observation 7.4.12. Let α be a root and write A_α for the unique element of $\mathfrak{h}$ such that $\alpha = \langle A_\alpha, -\rangle$. For $X_\alpha, Y_\alpha, H_\alpha$ as above, we have

$$[H_\alpha, X_\alpha] = \alpha(H_\alpha)X_\alpha = 2X_\alpha,$$

so $\langle A_\alpha, H_\alpha\rangle = \alpha(H_\alpha) = 2$. Meanwhile, H_α is a multiple of A_α, so that we must have

$$H_\alpha = \frac{2A_\alpha}{\langle A_\alpha, A_\alpha\rangle}.$$

Thus the element $H_\alpha \in \mathfrak{h}$ is uniquely determined; it is called the *coroot* of α.

Corollary 7.4.13. *Let (V, ρ) be a finite-dimensional complex representation of $\mathfrak{g}$, and suppose μ is a weight of ρ. Then for any root α, the value $\mu(H_\alpha)$ must be an integer. In particular, for any root β of $\mathfrak{g}$, we have that $\beta(H_\alpha)$ is an integer.*

Proof. Let $X \in V$ be a weight vector for μ. Then $\rho(H_\alpha)X = \mu(H_\alpha) \cdot X$, so that $\mu(H_\alpha)$ is an eigenvalue of H_α for the restriction of ρ to an action of $\mathfrak{s}_\alpha \cong \mathfrak{sl}_2$ on V. It must therefore be an integer by Proposition 6.1.3. □

Notation 7.4.14. At this point it starts to become very convenient to use the inner product to identify $\mathfrak{h}$ with $\mathfrak{h}^*$, so that we can talk about the inner product $\langle \alpha, \beta \rangle = \langle A_\alpha, A_\beta \rangle$ of two roots. (Alternatively, we can identify the root α with the element A_α in $\mathfrak{h}$.) Note also that since the roots lie in $i\mathfrak{t}$, their inner products are necessarily *real*

Observation 7.4.15. Using the formula for the coroot H_α in Observation 7.4.12, we get

$$\beta(H_\alpha) = \langle A_\beta, H_\alpha \rangle = 2\frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}.$$

We can interpret $\frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}\alpha$ as the orthogonal projection of β on α, so Corollary 7.4.13 says that for roots α and β, the orthogonal projection of β on α must be a half-integer multiple of α.

Observation 7.4.16. The fact that $\frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}$ is an integer severely constrains the possible values of $\langle \beta, \alpha \rangle$: Recall that if θ is the angle between α and β (in the *real* inner product space $i\mathfrak{t}$) then $\langle \beta, \alpha \rangle = \|\alpha\|\|\beta\| \cos\theta$, so that

$$4\cos^2\theta = \frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} \cdot \frac{2\langle \alpha, \beta \rangle}{\langle \beta, \beta \rangle}$$

is a non-negative integer. As $0 \leq \cos^2\theta \leq 1$, the only possible values of the left-hand side in this equation are $0, 1, 2, 3, 4$. If we write

$$m_1 := \frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}, \qquad m_2 := \frac{2\langle \alpha, \beta \rangle}{\langle \beta, \beta \rangle}$$

then m_1, m_2 are integers with the same sign, and if they are non-zero we have

$$\|\beta\|/\|\alpha\| = \sqrt{m_2/m_1}.$$

If we assume $|m_2| \geq |m_1|$, we get the following possible values (with $\theta \leq \pi$):

$4\cos^2\theta$	$\cos\theta$	θ	(m_1, m_2)	$\|\beta\|/\|\alpha\|$
0	0	$\pi/2$	$(0, 0)$	?
1	$\pm 1/2$	$\pi/3, 2\pi/3$	$(1, 1), (-1, -1)$	1
2	$\pm 1/\sqrt{2}$	$\pi/4, 3\pi/4$	$(1, 2), (-1, -2)$	$\sqrt{2}$
3	$\pm\sqrt{3}/2$	$\pi/6, 5\pi/6$	$(1, 3), (-1, -3)$	$\sqrt{3}$
4	± 1	$0, \pi$	$(\pm 1, \pm 4)$, $(\pm 2, \pm 2)$	2, 1

(Note however that we will see in a moment that in the final case, where β is a multiple of α, we must have $\alpha = \pm\beta$ — the case where $\|\beta\| = 2\|\alpha\|$ is impossible. All other possibilities in the table do occur, however.) We can summarize this calculation as:

Proposition 7.4.17. *Suppose α and β are roots such that α is not a multiple of β, and $\|\beta\| \geq \|\alpha\|$. Then one of the following must hold:*

(1) $\langle \alpha, \beta \rangle = 0$, i.e. α and β are orthogonal.

(2) $\|\beta\| = \|\alpha\|$ and the angle between α and β is $\pi/3$ or $2\pi/3$.

(3) $\|\beta\| = \sqrt{2}\|\alpha\|$ and the angle between α and β is $\pi/4$ or $3\pi/4$.

(4) $\|\beta\| = \sqrt{3}\|\alpha\|$ and the angle between α and β is $\pi/6$ or $5\pi/6$.

Proposition 7.4.18. *Let α be a root.*

(i) The only (real-number) multiples of α that are roots are α and $-\alpha$.

(ii) The root space $\mathfrak{g}_\alpha$ is 1-dimensional.

Proof. Suppose $c\alpha$ is a root. Since there are only finitely many roots, there is some minimal c such that this holds; by replacing α by this minimal multiple, we may assume $\|c\| \geq 1$. From the calculations in Observation 7.4.16 the only possibilities are then $c = \pm 2$ and $c = \pm 1$; the second case gives the known roots $\pm\alpha$, so we only need to exclude the case $c = \pm 2$.

Let V be the subspace $\mathbb{C} \cdot H_\alpha \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \mathfrak{g}_{2\alpha} \oplus \mathfrak{g}_{-2\alpha}$ of $\mathfrak{g}$. From Corollary 7.4.9 and the fact that there are no other multiples of α that are roots, we see that V is invariant under the adjoint action of $\mathfrak{s}_\alpha \cong \mathfrak{sl}_2$. Thus V is a finite-dimensional representation of $\mathfrak{sl}_2$, and $\mathfrak{s}_\alpha$ is an invariant subspace; as representations of $\mathfrak{sl}_2$ are completely reducible it therefore has an invariant complement U. Recall that $\alpha(H_\alpha) = 2$, so if $X \in \mathfrak{g}_{\pm 2\alpha}$ we have $[H_\alpha, X] = \pm 2\alpha(H_\alpha)X = \pm 4X$. The definition of V thus gives it as a direct sum of eigenspaces for the eigenvalues $0, \pm 2, \pm 4$. The endomorphism $[H_\alpha, -]$ of U must then have an eigenvector for one of these eigenvalues. Since they are all even, we see from Proposition 6.1.3 that 0 must also be an eigenvalue in U. But this is impossible, since multiples of H_α are the only eigenvectors for 0 in V and we have $U \cap \mathfrak{s}_\alpha = 0$. Thus we must have $U = 0$ and so $V = \mathfrak{s}_\alpha$. This means that $\mathfrak{g}_{\pm 2\alpha} = 0$ and that $\mathfrak{g}_\alpha$ is spanned by X_α and so is 1-dimensional. □

Combining Proposition 7.4.18 with Proposition 7.4.6, we get:

Corollary 7.4.19. *The dimension of $\mathfrak{g}$ is the sum of $|R|$ and the rank of $\mathfrak{g}$ (the dimension of $\mathfrak{h}$).* □

Lemma 7.4.20. *For the inner product $\langle X, Y \rangle = \operatorname{tr}(X^\dagger Y)$ on $\mathfrak{gl}_n(\mathbb{C})$ we have:*

(i) If X, Y are skew-Hermitian, then $\langle X, Y \rangle$ is real.

(ii) If X is skew-Hermitian, then

$$\langle [X,Y], Z\rangle = -\langle Y, [X,Z]\rangle$$

for all $Y, Z \in \mathfrak{gl}_n(\mathbb{C})$.

Proof. For (i), we have (using that $X^\dagger = -X$, $Y^\dagger = -Y$ and the cyclic symmetry of the trace, i.e. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$):

$$\overline{\operatorname{tr}(X^\dagger Y)} = \operatorname{tr}(X^\dagger Y)^\dagger = \operatorname{tr}(Y^\dagger X) = \operatorname{tr}(-YX) = \operatorname{tr}(-XY) = \operatorname{tr} X^\dagger Y.$$

For (ii), we get

$$\langle [X,Y], Z\rangle = \operatorname{tr}\,[X,Y]^\dagger Z = \operatorname{tr}\,(XY - YX)^\dagger Z = \operatorname{tr}(Y^\dagger X^\dagger - X^\dagger Y^\dagger)Z = \operatorname{tr}(YXZ) - \operatorname{tr}(XYZ),$$
$$-\langle Y, [X,Z]\rangle = -\operatorname{tr} Y^\dagger [X,Z] = -\operatorname{tr}(-Y)(XZ - ZX) = \operatorname{tr}(YXZ) - \operatorname{tr}(YZX).$$

These are equal, since the cyclic symmetry of the trace implies $\operatorname{tr}(XYZ) = \operatorname{tr}(YZX)$. □

Example 7.4.21 (Roots of $\mathfrak{sl}_n$)**.** Recall from Example 7.3.8 that the traceless diagonal matrices give a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{sl}_n$. Keeping the notation from Proposition 7.1.11, if D is a diagonal matrix with entries $(\lambda_1, \dots, \lambda_n)$, then we saw that

$$[D, E_{ij}] = (\lambda_i - \lambda_j)E_{ij}$$

for $i \neq j$. In other words, E_{ij} is a root vector for the root $L_{ij} \in \mathfrak{h}^*$ defined by $L_{ij}(D) = \lambda_i - \lambda_j$. Together with $\mathfrak{h}$, the matrices E_{ij} for $i \neq j$ span all of $\mathfrak{sl}_n$, so there is no room for other roots. (Note also that $L_{ij} = -L_{ji}$.)

If we think of $\mathfrak{sl}_n$ as $\mathfrak{su}_n \otimes \mathbb{C}$ and $\mathfrak{h}$ as $\mathfrak{t} \otimes \mathbb{C}$ where $\mathfrak{t}$ is the diagonal matrices in $\mathfrak{su}_n$, then $\mathfrak{t}$ consists of traceless diagonal matrices with pure imaginary entries, so that $i\mathfrak{t}$ consists of traceless diagonal matrices with *real* entries.

Consider the usual inner product $\langle X, Y\rangle = \operatorname{tr}(X^\dagger Y)$ on $\mathfrak{gl}_n(\mathbb{C})$. This real inner product extends to the standard complex inner product on $\mathfrak{sl}_n$. For this inner product, the matrices E_{ij} satisfy $\langle E_{ij}, X\rangle = X_{ij}$, so that on $\mathfrak{h}$ we have

$$L_{ij} = \langle E_{ii} - E_{jj}, -\rangle.$$

Using the notation

$$\delta_{ij} = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases}$$

we get

$$\langle L_{ij}, L_{kl}\rangle = \langle E_{ii} - E_{jj}, E_{kk} - E_{ll}\rangle = \delta_{ik} - \delta_{il} - \delta_{jk} + \delta_{jl}.$$

Since $i \neq j$ and $k \neq l$, this inner product takes the values

- 0, if i, j, k, l are all distinct, in which case L_{ij} and L_{kl} are orthogonal,

- ± 1 if $\{i,j\}$ and $\{k,l\}$ have a single element in common, in which case the angle between L_{ij} and L_{kl} is $\pi/3$ or $2\pi/3$,
- ± 2 if $\{i,j\} = \{k,l\}$ (in which case $L_{ij} = \pm L_{kl}$); in particular the length of each root L_{ij} is $\sqrt{2}$.

Let's look at the two smallest cases: For $n = 2$ we just have the two roots L_{21} and $L_{12} = -L_{21}$ and there's not much more to say. When $n = 3$ we have the six roots $L_{12} = -L_{21}$, $L_{13} = -L_{31}$, $L_{23} = -L_{32}$. Previously we discussed of roots for $\mathfrak{sl}_3$ as pairs of integers (m_1, m_2), corresponding to the linear functional taking the value m_i on H_i. Recalling that

$$H_1 = E_{11} - E_{22}, \qquad H_2 = E_{22} - E_{33},$$

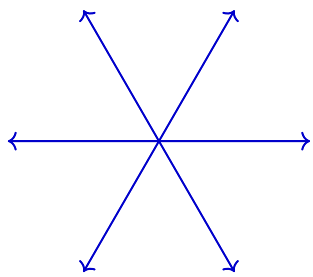
we see that

$$L_{ij}(H_1) = \delta_{i1} - \delta_{i2} - \delta_{j1} + \delta_{j2}, \quad L_{ij}(H_2) = \delta_{i2} - \delta_{i3} - \delta_{j2} + \delta_{j3}.$$

This gives the following correspondence between our new and old descriptions of the roots:

L_{12}	$(2,-1)$
L_{21}	$(-2,1)$
L_{13}	$(1,1)$
L_{31}	$(-1,-1)$
L_{23}	$(-1,2)$
L_{32}	$(1,-2)$.

In terms of our inner product, all the roots have length $\sqrt{2}$ — note that this is *not* the same as what we get by thinking of the pair of integers in the second column as a subset of $\mathbb{R}^2$! Computing the angles, we get the following geometric depiction of the roots:



In other words, the roots are the vertices of a regular hexagon. If the root along the positive x-axis represents L_{13} then the other two roots with positive x-values are L_{12} and L_{23}.

7.5 Weyl groups

We now introduce the *Weyl group*, which acts on the Cartan subalgebra $\mathfrak{h}$ and always permutes the weights of a representation. Its action on the roots will allow us to deduce further important information about their structure.

Definition 7.5.1. For a root α, we define $s_\alpha\colon \mathfrak{h} \to \mathfrak{h}$ by the formula

$$s_\alpha(H) = H - 2\frac{\langle \alpha, H\rangle}{\langle \alpha, \alpha\rangle}\alpha.$$

The *Weyl group* $W_\mathfrak{g}$ of $\mathfrak{g}$ is the subgroup of $\mathrm{GL}(\mathfrak{h})$ generated by all the maps s_α for $\alpha \in R$.

Observation 7.5.2. We know that any root α lies in $i\mathfrak{t}$, so s_α will take the real subspace $i\mathfrak{t}$ to itself. As an automorphism of $i\mathfrak{t}$, s_α is the reflection in the hyperplane orthogonal to α. In particular, $s_\alpha(T) = T$ if $\langle \alpha, T\rangle = 0$ and $s_\alpha(\alpha) = -\alpha$. Since any reflection is orthogonal, we may regard $W_\mathfrak{g}$ as a subgroup of $\mathrm{O}(i\mathfrak{t})$.

Proposition 7.5.3. *Let (V, ρ) be a finite-dimensional complex representation of $\mathfrak{g}$. If μ is a weight of ρ, then so is $s_\alpha(\mu)$ for any root α. Moreover, their multiplicities are the same.*

Proof. We define an endomorphism S_α of V by

$$S_\alpha := e^{\rho(X_\alpha)}e^{-\rho(Y_\alpha)}e^{-\rho(X_\alpha)}.$$

We want to show that $S_\alpha\rho(H)S_\alpha^{-1} = \rho(s_\alpha H)$ for all $H \in \mathfrak{h}$. Since both sides are linear in H, it suffices to check this separately in the cases where H is orthogonal to α and parallel to α.

If $H \in \mathfrak{h}$ satisfies $\langle \alpha, H\rangle = 0$, then $[H, X_\alpha] = \langle \alpha, H\rangle X_\alpha = 0$, which implies that $\rho(H)$ and $\rho(X_\alpha)$ commute. Similarly $\rho(H)$ commutes with $\rho(Y_\alpha)$, and hence S_α commutes with $\rho(H)$, i.e. $S_\alpha\rho(H)S_\alpha^{-1} = \rho(H) = \rho(s_\alpha H)$ since we also have $s_\alpha H = H$.

On the other hand, applying Proposition 6.1.3(iv) to the restriction of ρ to $\mathfrak{s}_\alpha$, it follows that we have

$$S_\alpha\rho(H_\alpha)S_\alpha^{-1} = -\rho(H_\alpha) = \rho(s_\alpha H_\alpha),$$

since s_α acts by multiplication by -1 on multiples of α.

Now suppose v is a weight vector for μ. Then we get

$$\rho(H)S_\alpha^{-1}v = S_\alpha^{-1}\rho(s_\alpha H)v = \langle \mu, s_\alpha H\rangle S_\alpha^{-1}v = \langle s_\alpha\mu, H\rangle S_\alpha^{-1}v,$$

so that $S_\alpha^{-1}v$ is a weight vector for $s_\alpha\mu$. (Here we used that s_α is orthogonal on $i\mathfrak{t}$ with $s_\alpha^2 = \mathrm{id}$, and the weight μ lies in $i\mathfrak{t}$.) This means that S_α^{-1} gives an isomorphism between the weight spaces for μ and $s_\alpha\mu$, so their dimensions must be equal. □

It follows that the Weyl group acts on the set of weights of any representation ρ. This action always restricts to the non-zero weights, from which we in particular get an action of $W_{\mathfrak{g}}$ on the set R of roots of $\mathfrak{g}$.

Corollary 7.5.4. *The Weyl group $W_{\mathfrak{g}}$ is finite.*

Proof. Since the roots span $\mathfrak{h}$ by Corollary 7.4.7, every $w \in W_{\mathfrak{g}}$ is determined by its values on R. But w takes R to itself, so we may identify $W_{\mathfrak{g}}$ with a subgroup of the permutation group of R, which is finite. □

Proposition 7.5.5. *Suppose α and β are roots, and let $0 \leq \theta \leq \pi$ be the angle between them.*

(i) *If the angle is strictly obtuse ($\pi/2 < \theta < \pi$ or $\langle\alpha, \beta\rangle < 0$) then $\alpha + \beta$ is a root.*

(ii) *If the angle is strictly acute ($0 < \theta < \pi/2$ or $\langle\alpha, \beta\rangle > 0$) then $\alpha - \beta$ is a root.*

Proof. Suppose $\|\beta\| \geq \|\alpha\|$. We already computed the possible angles and corresponding values of the integer $m := 2\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}$ in Observation 7.4.16; omitting those that are not strictly acute or obtuse, we have:

θ	m
$\pi/6, \pi/4, \pi/3$	1
$2\pi/3, 3\pi/4, 5\pi/6$	-1

We know that $s_\alpha\beta = \beta - m\alpha$ is a root, and this equals $\beta - \alpha$ for the three possible strictly acute angles, and $\beta + \alpha$ for the three strictly obtuse angles. □

Observation 7.5.6. Let α, β be roots such that $\alpha \neq \pm\beta$. We call the roots of the form $\beta + n\alpha$ for $n \in \mathbb{Z}$ the *α-string* of roots through β. Since there are finitely many roots, there exist maximal non-negative integers s, t such that $\beta - s\alpha$ and $\beta + t\alpha$ are roots. We claim that then $\beta + n\alpha$ is a root for any integer n with $-s \leq n \leq t$. Indeed, suppose $\beta + n\alpha$ is not a root, and $-s \leq n \leq t$. If we let $-s \leq j < n$ be maximal such that $\beta + j\alpha$ is a root and $n < k \leq t$ be minimal such that $\beta + k\alpha$ is a root, then we have $j < k$ and $\beta + (j + 1)\alpha$ and $\beta + (k - 1)\alpha$ are *not* roots. For this not to contradict Proposition 7.5.5 we must have

$$\langle\alpha, \beta + j\alpha\rangle = \langle\alpha, \beta\rangle + j\|\alpha\|^2 \geq 0, \quad \langle\alpha, \beta + k\alpha\rangle = \langle\alpha, \beta\rangle + k\|\alpha\|^2 \leq 0.$$

Since $j < k$ this is clearly impossible, which means that the α-string of roots through β consists of $\beta + n\alpha$ for all $-s \leq n \leq t$. Since s_α acts on roots by adding a multiple of α, the α-string is invariant under the action of s_α. It is easy to check (and intuitively clear) that the reflection s_α must reverse the string, so that

$$\beta - s\alpha = s_\alpha(\beta + t\alpha) = \beta - 2\frac{\langle\beta, \alpha\rangle}{\langle\alpha, \alpha\rangle}\alpha - t\alpha,$$

which means that

$$2\frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} = s - t.$$

Thus we see in particular that the length of a root string can be at most 4.

Example 7.5.7 (The Weyl group of $\mathfrak{sl}_n$)**.** Keeping the notation from Example 7.4.21, we see that for the root $L_{ij} = E_{ii} - E_{jj}$ and $D \in \mathfrak{h}$, we have

$$s_{L_{ij}} D = D - (D_{ii} - D_{jj})(E_{ii} - E_{jj}) = D + (D_{jj} - D_{ii})E_{ii} + (D_{ii} - D_{jj})E_{jj}.$$

Thus $s_{L_{ij}}$ acts on D by swapping the ith and jth diagonal entries. This means that we can identify $W_{\mathfrak{sl}_n}$ with the symmetric group S_n on n letters.

7.6 Bases

In order to define the notion of a "highest" weight for $\mathfrak{sl}_3$, we singled out two special roots. This choice is an instance of the general notion of a *base*, which we now consider for a general semisimple Lie algebra $\mathfrak{g}$.

We keep the notation from before, so $\mathfrak{g}$ is a fixed semisimple complex Lie algebra with Cartan subalgebra $\mathfrak{h}$, etc.

Definition 7.6.1. A subset Δ of the set of roots R is called a *base* if

(1) Δ is a basis of $i\mathfrak{t}$ as an $\mathbb{R}$-vector space (or of $\mathfrak{h}$ as a $\mathbb{C}$-vector space),

(2) when we write any root as a linear combination of the elements of Δ, then the coefficients are integers and are moreover either all non-positive or all non-negative.

The roots where the coefficients are all non-negative are called the *positive* roots, and the others the *negative* roots. We also call the elements of Δ the *(positive) simple* roots.

Exercise 7.3. Check that the weights $L_{i(i+1)} = E_{ii} - E_{(i+1)(i+1)}$ for $i = 1, \ldots, n-1$ form a base for $\mathfrak{sl}_n$. (In particular, for $n = 3$ the two roots L_{12} and L_{23} are a base, and as we saw in Example 7.4.21 these correspond to the two positive simple roots we used in our previous discussion of $\mathfrak{sl}_3$-representations.)

Lemma 7.6.2. *If α, β are distinct elements of a base Δ, then $\langle \alpha, \beta \rangle \leq 0$ (i.e. the angle between α and β is obtuse).*

Proof. If $\langle \alpha, \beta \rangle > 0$ then Proposition 7.5.5 implies that $\alpha - \beta$ is a root. But this contradicts the assumption that for a root the coefficients of the simple roots have the same sign. □

We want to show that there always exists a base — in fact, we will find all possible choices of one. The key idea is to divide the roots into sets of "positive" and "negative" ones by taking the positive roots to be those that lie on one side of a suitable hyperplane.

Lemma 7.6.3. *There exists a hyperplane V through the origin in $i\mathfrak{t}$ that does not contain any roots. (In other words, there exists some $X \in i\mathfrak{t}$ such that $\langle X, \alpha \rangle \neq 0$ for all $\alpha \in R$, so that no roots are contained in the hyperplane orthogonal to X.)*

Proof. For $\alpha \in R$, let V_α be the hyperplane orthogonal to α. Since there are finitely many roots, there exists some $X \in i\mathfrak{t}$ that is not contained in any of these hyperplanes. But then the hyperplane orthogonal to X does not contain any roots, as required. □

Definition 7.6.4. Suppose V is a hyperplane through the origin in $i\mathfrak{t}$ that does not contain any roots. Choose one "side" of V, and let R^+ denote the set of roots on this side of V. (Equivalently, if we choose some $X \neq 0$ that is orthogonal to V, we can define R^+ to be those roots α such that $\langle X, \alpha \rangle > 0$.) We then say that an element $\alpha \in R^+$ is *decomposable* if $\alpha = \beta + \gamma$ for some $\beta, \gamma \in R^+$. If no such β, γ exist we say that α is *indecomposable*.

Theorem 7.6.5. *Given a hyperplane V through the origin that does not contain any roots and a choice of a positive "side" of V, the set Δ of indecomposable roots in R^+ is a base for R.*

Proof. Choose $X \neq 0$ such that V is the hyperplane orthogonal to X, and R^+ consists of the roots α such that $\langle X, \alpha \rangle > 0$.

We first check that every element of R^+ is a linear combination of Δ with non-negative integer coefficients. Suppose this is false. Then we can find $\alpha \in R^+$ that cannot be expressed in this way and such that $\langle X, \alpha \rangle$ is minimal. Since α is certainly not contained in Δ, it must be decomposable, so that we can write $\alpha = \beta + \gamma$ with $\beta, \gamma \in R^+$. Then at least one of β, γ cannot be expressed as a linear combination of Δ with non-negative integer coefficents, since otherwise the same would be true of the sum α. We also have

$$\langle X, \alpha \rangle = \langle X, \beta \rangle + \langle X, \gamma \rangle,$$

where the terms on the right-hand side are both strictly positive. This contradicts the assumption that $\langle X, \alpha \rangle$ was minimal.

Next, we show that for $\alpha, \beta \in \Delta$ we must have $\langle \alpha, \beta \rangle \leq 0$. Indeed, if $\langle \alpha, \beta \rangle > 0$ then $\alpha - \beta$ and $\beta - \alpha$ are roots by Proposition 7.5.5. One of these roots must be positive, but if $\alpha - \beta$ is positive then $\alpha = (\alpha - \beta) + \beta$, which contradicts the assumption that α is indecomposable. Similarly if $\beta - \alpha$ is a root then β would be decomposable.

Using this, we can show that the elements of Δ are linearly independent: Suppose we have $\sum_{\alpha \in \Delta} c_\alpha \alpha = 0$. Let I be the set of roots in Δ such that c_α is non-negative, and J the set of α where $c_\alpha < 0$. Then we have

$$v := \sum_{\alpha \in I} c_\alpha \alpha = \sum_{\beta \in J} (-c_\beta) \beta.$$

The vector v satisfies

$$\langle v, v\rangle = \langle \sum_{\alpha\in I} c_\alpha \alpha, \sum_{\beta\in J}(-c_\beta)\beta\rangle = \sum_{\alpha\in I,\beta\in J} c_\alpha(-c_\beta)\langle\alpha,\beta\rangle.$$

In this sum the coeffiecients $c_\alpha(-c_\beta)$ are non-negative, but $\langle\alpha,\beta\rangle \leq 0$. Since $\langle v, v\rangle \geq 0$ this is only possible if $v = 0$. But then we have

$$0 = \langle X, v\rangle = \sum_{\alpha\in I} c_\alpha\langle X, \alpha\rangle$$

where $\langle X, \alpha\rangle > 0$, which means that each coefficient c_α has to be 0. Similarly, we cannot have any negative coefficients, so $J = \emptyset$.

It only remains to show that Δ spans $i\mathfrak{t}$. But we already showed that the positive roots are linear combinations of the elements of Δ, hence so are their negatives. Since the roots span by Corollary 7.4.7 this completes the proof. □

Combining this with Lemma 7.6.3, we get:

Corollary 7.6.6. *There exists a base for the roots of the complex semisimple Lie algebra* $\mathfrak{g}$. □

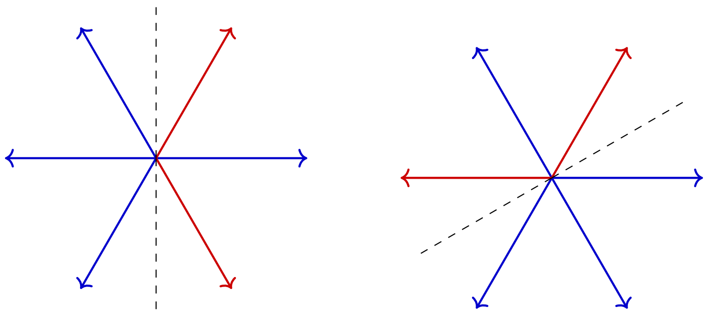
In fact, any choice of base arises from the construction in Theorem 7.6.5:

Proposition 7.6.7. *If Δ is a base of R, then there exists a hyperplane V through the origin of $i\mathfrak{t}$ and a side of V with respect to which Δ is precisely the indecomposable positive roots.*

Proof. Suppose $\Delta = \{\alpha_1, \ldots, \alpha_r\}$. Since the roots are a basis for $i\mathfrak{t}$, for any real numbers $c_1, \ldots, c_r$ there exists a unique γ such that $\langle\gamma, \alpha_i\rangle = c_i$ for all i. In particular, we can find γ such that $\langle\gamma, \alpha_i\rangle > 0$ for all i. Then we must have $\langle\gamma, \alpha\rangle > 0$ for all $\alpha \in R^+$, since the positive roots are linear combinations of the α_i with non-negative integer coefficients. This also means that $\langle\gamma, \alpha\rangle < 0$ if α is a negative root, so the positive roots are precisely those such that $\langle\gamma, \alpha\rangle > 0$.

It remains to show that the elements of Δ are precisely the indecomposable positive roots. First suppose $\alpha \in \Delta$ and we can write $\alpha = \beta + \gamma$ for positive roots β, γ. But β and γ are expressible uniquely as non-negative integer linear combinations of the elements of Δ, and since they cannot both be multiples of α this gives an expression of α as a linear combination of elements of Δ with a non-zero coefficient for an element other than α. This contradicts the linear independence of Δ, and so cannot happen. Thus Δ is contained in the indecomposable roots. On the other hand, both Δ and the indecomposable roots are bases for the vector space $i\mathfrak{t}$, and so they must contain the same number of elements. □

Example 7.6.8. The following diagram illustrates two choices of base for $\mathfrak{sl}_3$:



The positive roots are those on one side of the dashed line, and the two red roots are the positive simple roots for one choice of positive side.

Definition 7.6.9. The *(closed) fundamental Weyl chamber* of a base Δ is the set of all $X \in i\mathfrak{t}$ such that $\langle \alpha, X \rangle \geq 0$ for all $\alpha \in \Delta$, while the *open fundamental Weyl chamber* is the subset of X where $\langle \alpha, X \rangle > 0$ for all $\alpha \in \Delta$.

Definition 7.6.10. The *open Weyl chambers* are the connected components of $i\mathfrak{t} \setminus \bigcup_{\alpha \in R} V_\alpha$, where V_α is the hyperplane orthogonal to the root α.

There is a canonical bijection between bases and Weyl chambers:

Proposition 7.6.11.

(i) The open fundamental Weyl chamber of a base is an open Weyl chamber.

(ii) For each open Weyl chamber C there exists a unique base Δ_C such that C is the open fundamental Weyl chamber of Δ_C.

(iii) The positive roots for Δ_C are precisely those $\alpha \in R$ such that $\langle \alpha, X \rangle > 0$ for all $X \in C$.

Proof. Let Δ be a base, and let C be its open fundamental Weyl chamber. We have $C \subseteq A := i\mathfrak{t} \setminus \bigcup_{\alpha \in R} V_\alpha$: If $\gamma \in C$ then by assumption $\langle \gamma, \alpha \rangle > 0$ for all $\alpha \in \Delta$, and hence $\langle \gamma, \alpha \rangle > 0$ for any positive root α, since these are linear combinations of the elements of Δ with non-negative integer coefficients. Then we must have $\langle \gamma, \alpha \rangle < 0$ for all the negative roots, so in particular $\langle \gamma, \alpha \rangle \neq 0$ for all roots α. We must show that C is a (path-)component of A.

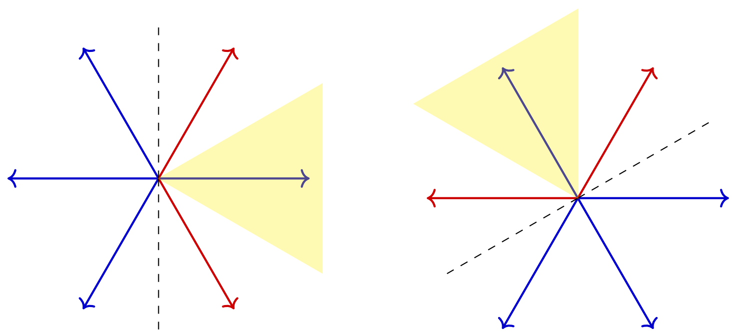
Since the elements of Δ are a basis for $i\mathfrak{t}$, we can (as we've seen before) find some γ such that $\langle \gamma, \alpha \rangle > 0$ for all $\alpha \in \Delta$, so that $C \neq \emptyset$. It is also clear that if $\gamma, \gamma' \in C$ then so is $t\gamma + (1-t)\gamma'$ for $0 \leq t \leq 1$, so that C is path-connected (and indeed convex). On the other hand, if C is not a path-component then there must exist $\gamma \in C$ and $\gamma' \in A \setminus C$ and a path p in A with $p(0) = \gamma, p(1) = \gamma'$. If $\gamma' \notin C$ then we must have $\langle \gamma', \alpha \rangle < 0$ for some $\alpha \in \Delta$. But then $\langle p(t), \alpha \rangle$ changes sign, so p must pass through V_α, contradicting the assumption that p lies in A. This proves (i).

To prove (ii), suppose C is an open Weyl chamber and pick $\gamma \in C$. Then the hyperplane orthogonal to γ does not contain any roots, so there is a base Δ

where the positive roots are those $\alpha \in R$ such that $\langle \alpha, \gamma \rangle > 0$. Since the sign of $\langle \alpha, - \rangle$ doesn't change in C, we then have $\langle \alpha, X \rangle > 0$ for all $X \in C$ and all positive roots α. Then the negative roots α must satisfy $\langle \alpha, X \rangle < 0$ for $X \in C$, so that the positive roots are precisely those roots α such that $\langle \alpha, X \rangle > 0$ for all $X \in C$.

Moreover, every element $X \in C$ satisfies $\langle \alpha, X \rangle > 0$ for $\alpha \in \Delta$, so that C is contained in the open fundamental Weyl chamber of Δ; since both C and the open fundamental Weyl chamber are path-components, they must be equal. This shows that there exists a base as in (ii), and that this satisfies (iii). It remains to show that this base is unique. To see this, suppose Δ' is a base whose open fundamental Weyl chamber is C. Then $\langle \alpha, \gamma \rangle > 0$ for $\alpha \in \Delta'$, so that the positive roots for Δ' are the same as those for Δ. Then the indecomposable positive roots must also be the same, i.e. $\Delta = \Delta'$. □

Example 7.6.12. The following diagram depicts the fundamental Weyl chambers for the two bases of $\mathfrak{sl}_3$ from Example 7.6.8:



Here the fundamental Weyl chamber is the yellow triangular region (extending off to infinity).

We state without proof some further results on the relation between the Weyl group and Weyl chambers or bases; see for instance [2, §8.5] for more details:

Proposition 7.6.13. *If Δ is a base, then the Weyl group $W_{\mathfrak{g}}$ is generated by the reflections s_α for $\alpha \in \Delta$.* □

Theorem 7.6.14. *The action of the Weyl group $W_{\mathfrak{g}}$ on $\mathfrak{t}$ induces a free and transitive action on the set of open Weyl chambers, i.e. for any pair of open Weyl chambers C, C' there exists a unique element $w \in W_{\mathfrak{g}}$ such that $C = w \cdot C'$.* □

The bijection between bases and Weyl chambers is compatible with the Weyl group action, so this implies:

Corollary 7.6.15. *The action of the Weyl group $W_{\mathfrak{g}}$ on $\mathfrak{t}$ induces a free and transitive action on the set of bases, i.e. for any pair of bases Δ, Δ' there exists a unique element $w \in W_{\mathfrak{g}}$ such that $\Delta = w \cdot \Delta'$.* □

Exercise 7.4. Verify that the Weyl group acts freely and transitively on the Weyl chambers and bases of $\mathfrak{sl}_3$.

7.7 Highest weights

Given a choice Δ of a base for the roots of our complex semisimple Lie algebra $\mathfrak{g}$, we can define notions of *dominant* and *integral* weights for this basis, as well as a partial order of weights that lets us define the *highest* weight of a representation.

Definition 7.7.1. An element $\mu \in i\mathfrak{t}$ is *integral* if

$$\langle \mu, H_\alpha \rangle = 2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle}$$

is an integer for all roots $\alpha \in R$. If Δ is a base for R, then we say that μ is *dominant* (relative to Δ) if $\langle \alpha, \mu \rangle \geq 0$ for all $\alpha \in \Delta$, and *strictly dominant* if $\langle \alpha, \mu \rangle > 0$ for all $\alpha \in \Delta$.

Observation 7.7.2. Let (V, ρ) be a finite-dimensional complex representation of $\mathfrak{g}$. Then it follows from Proposition 7.4.4 and Corollary 7.4.13 that the weights of ρ are integral elements of $i\mathfrak{t}$.

Remark 7.7.3. An element μ is thus strictly dominant if it lies in the open fundamental Weyl chamber of Δ, and dominant if it lies in the closed fundamental Weyl chamber. Note also that every root is integral, so any $\mathbb{Z}$-linear combination of roots is integral. However, it is typically not the case that *all* integral elements are $\mathbb{Z}$-linear combinations of roots.

Lemma 7.7.4. *An element $\mu \in i\mathfrak{t}$ is integral if and only if $\langle \mu, H_\alpha \rangle$ is an integer for all $\alpha \in \Delta$.*

Proof. It can be shown that for any root β, we can write H_β as a $\mathbb{Z}$-linear combination of H_α with $\alpha \in \Delta$. (This is because the coroots H_β for $\beta \in R$ form the roots of a *dual* root system, and the coroots for Δ form a dual base for this; see [2, §8.3 and Proposition 8.18].) □

Definition 7.7.5. Let $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ be a base. Then the *fundamental weights* ω_i $(i = 1, \ldots, r)$ are the elements of $i\mathfrak{t}$ characterized by

$$\langle \omega_j, H_{\alpha_i} \rangle = 2\frac{\langle \omega_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

for all $i = 1, \ldots, r$. (Since the elements of Δ form a basis, basic linear algebra implies that there exists a unique ω_j for which these equations hold for all i.) Then the integral elements of $i\mathfrak{t}$ are precisely the $\mathbb{Z}$-linear combinations of the fundamental weights, while the dominant integral elements are their linear combinations with non-negative integer coefficients.

Definition 7.7.6. The integral elements of $i\mathfrak{t}$ are precisely the $\mathbb{Z}$-linear combinations of the fundamental weights, and so form a lattice in $i\mathfrak{t}$. This is called the *weight lattice* (since it consists of the possible weights of representations), and is independent of the choice of a base.

Definition 7.7.7. Let $\Delta = \{\alpha_1, \dots, \alpha_r\}$ be a base. Then for $\mu, \lambda \in i\mathfrak{t}$ we say that μ is *higher* than λ (or that λ is *lower* than μ), relative to Δ, if when we write

$$\mu - \lambda = c_1\alpha_1 + \cdots + c_r\alpha_r,$$

we have $c_i \geq 0$ for all i. We write $\mu \geq \lambda$ to denote that μ is higher than λ (with the choice of Δ implicit); this gives a partial order on $i\mathfrak{t}$.

We can now state the theorem of the highest weight for a general semisimple Lie algebra, in two parts:

Theorem 7.7.8 (Highest weight 1)**.** *Let (V, ρ) be an irreducible finite-dimensional complexn representation of $\mathfrak{g}$. Then ρ has a unique highest weight μ, which is dominant integral, and its weight space is 1-dimensional. Moreover, any other irreducible finite-dimensional complex with highest weight μ is isomorphic to (V, ρ).*

Having spent a lot of effort to get to the point where we can state this part of the theorem, we will not actually say anything about the proof, since it is essentially the same as what we already did for $\mathfrak{sl}_3$:

Exercise 7.5. Check that, with appropriate minor changes, the definitions and proofs in Section 6.3 work for a general complex semisimple Lie algebra.

Theorem 7.7.9 (Highest weight 2)**.** *Suppose μ is a dominant integral weight for $\mathfrak{g}$. Then there exists an irreducible finite-dimensional complex representation of $\mathfrak{g}$ whose highest weight is μ.*

Remark 7.7.10. Note that this classification means there is a unique irreducible representation corresponding to each point in the weight lattice that lies in the (closed) fundamental Weyl chamber of our chosen base. Since the weights are symmetric under the action of the Weyl group, if we pick a different base this representation corresponds to the "same" point in the new Weyl chamber.

It is possible to give a general abstract construction of such a representation in terms of *Verma modules*, see [2, Chapter 9]. We can also give explicit descriptions of such irreducible representations for specific examples of (semi)simple Lie algebras. To show that they exist, the following is a useful starting point:

Observation 7.7.11. Suppose V and W are irreducible finite-dimensional complex representations of $\mathfrak{g}$ with highest weights λ and μ, respectively. Then, just as we saw for $\mathfrak{sl}_3$ in §6.4, the tensor product $V \otimes W$ contains an irreducible representation with highest weight $\lambda + \mu$. To show that there exists an irreducible representation for each possible highest weight, it therefore suffices to find an irreducible representation corresponding to each of the fundamental weights $\omega_1, \dots, \omega_r$.

Example 7.7.12. We compute the fundamental weights for our usual base for $\mathfrak{sl}_n$, which consists of the roots $L_{i(i+1)} = E_{ii} - E_{(i+1)(i+1)}$ where $i = 1, \dots, n-1$.

The fundamental weight ω_j is a traceless diagonal matrix; if its diagonal entries are $\lambda_1, \ldots, \lambda_n$, then we want to solve the equations

$$2\frac{\langle \omega_j, L_{i(i+1)}\rangle}{\langle L_{i(i+1)}, L_{i(i+1)}\rangle} = \delta_{ij} \quad \Longleftrightarrow \quad \langle \omega_j, E_{ii} - E_{(i+1)(i+1)}\rangle = \lambda_i - \lambda_{i+1} = \delta_{ij},$$

$$\lambda_1 + \cdots + \lambda_n = 0.$$

Thus $\lambda_i = \lambda_{i+1}$ if $i \neq j$, so that we have

$$\lambda_1 = \lambda_2 = \cdots = \lambda_j,$$

$$\lambda_{j+1} = \lambda_j - 1 = \lambda_1 - 1,$$

$$\lambda_{j+1} = \lambda_{j+2} = \cdots = \lambda_n = \lambda_1 - 1.$$

If we set $\lambda = \lambda_1$ then we get

$$\sum_{i=1}^{n} \lambda_i = n\lambda - (n-j) = 0,$$

so $\lambda = \frac{n-j}{n}$ and we have

$$\lambda_i = \begin{cases} \frac{n-j}{n}, & i \leq j, \\ -\frac{j}{n}, & i > j. \end{cases}$$

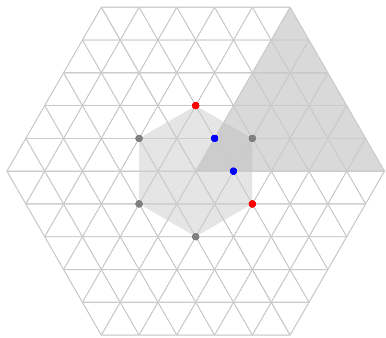
In terms of the basis $L_{i(i+1)}$, a traceless diagonal matrix with entries $(\lambda_1, \ldots, \lambda_n)$ is the linear combination

$$\lambda_1 L_{12} + (\lambda_2 + \lambda_1) L_{23} + (\lambda_3 + \lambda_2 + \lambda_1) L_{34} + \cdots + (\lambda_{n-1} + \cdots + \lambda_1) L_{(n-1)n} = \sum_{i=1}^{n-1} \left(\sum_{j=1}^{i} \lambda_j \right) L_{i(i+1)}.$$

For ω_k if we again set $\lambda = \frac{n-k}{n}$, the coefficient of $L_{i(i+1)}$ is given by

$$\begin{aligned} \sum_{j=1}^{i} \lambda_j &= \begin{cases} i\lambda, & i \leq k \\ i\lambda - (i-k), & i > k \end{cases} \\ &= \begin{cases} \frac{i(n-k)}{n}, & i \leq k \\ \frac{k(n-i)}{n}, & i > k \end{cases} \end{aligned}$$

The following diagram illustrates the weight lattice for $\mathfrak{sl}_2$, the roots (in grey) together with one choice of positive simple roots (in red), the corresponding fundamental Weyl chamber, and the fundamental weights (in blue):



Note that the fundamental weights lie on the boundary of the fundamental Weyl chamber (as is always the case). The highest weight of the adjoint representation is the unique root that is contained in the (closed) Weyl chamber.

Example 7.7.13 (Standard representation of $\mathfrak{sl}_n$)**.** Let's find the weights of the standard representation of $\mathfrak{sl}_n$ on $\mathbb{C}^n$. If $e_1, \ldots, e_n$ is the standard basis of $\mathbb{C}^n$, then $He_i = \lambda_i e_i$ if H is a diagonal matrix with entries $(\lambda_1, \ldots, \lambda_n)$. Thus e_i is a weight vector for the weight $\mu_i \in \mathfrak{h}^*$ given by $\mu_i(H) = H_{ii}$. Since these weight vectors span $\mathbb{C}^n$, these are all the weights. Next, we must describe these weights in terms of the fundamental weights. For this, we use our usual base and compute

$$\mu_j(L_{i(i+1)}) = \begin{cases} 1, & i = j \\ -1, & i = j-1, \\ 0, & i \neq j, j-1 \end{cases} \qquad \mu_j = \begin{cases} \omega_1, & j = 1, \\ \omega_j - \omega_{j-1}, & 1 < j < n, \\ -\omega_{n-1}, & j = n \end{cases}$$

We claim that ω_1 is the highest weight. To see this, we use the computation of ω_k in the basis $L_{i(i+1)}$ from Example 7.7.12. This gives in particular that

$$\omega_1 = \sum_{i=1}^{n-1} \frac{n-i}{n} L_{i(i+1)}, \qquad \omega_{n-1} = \sum_{i=1}^{n-1} \frac{i}{n} L_{i(i+1)},$$

$$\omega_j - \omega_{j-1} = \sum_{i=1}^{n} c_i L_{i(i+1)},$$

where

$$\begin{aligned} n \cdot c_i &= \begin{cases} i(n-j) - i(n-j+1), & i < j, \\ j(n-j) - (j-1)(n-j), & i = j, \\ j(n-i) - (j-1)(n-i), & i > j \end{cases} \\ &= \begin{cases} -i, & i < j, \\ (n-i) & i \geq j, \end{cases} \end{aligned}$$

We then have for all j that

$$\mu_j = \sum_{i<j} \frac{(-i)}{n} L_{i(i+1)} + \sum_{i=j}^{n-1} \frac{n-i}{n} L_{i(i+1)} = \sum_{i=j}^{n-1} L_{i(i+1)} - \sum_{i=1}^{n-1} \frac{i}{n} L_{i(i+1)} = L_{jn} - \sum_{i=1}^{n-1} \frac{i}{n} L_{i(i+1)}.$$

Thus if $j < k$ we have

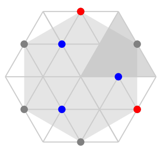
$$\mu_j - \mu_k = L_{jn} - L_{kn} = L_{jk} = \sum_{j \le i < k} L_{i(i+1)},$$

so that $\mu_j > \mu_k$ if $j < k$. In particular, this means that $\omega_1 = \mu_1 > \mu_j$ for all j, so that ω_1 is the highest weight. In fact, the standard representation is the *irreducible* representation with highest weight ω_1. This can be seen from the calculation of μ_j above, which shows that ω_1 is the *only* dominant weight of the standard representation. Since each piece in the decomposition into irreducible representations contributes at least one dominant integral weight (its highest weight), and the weight space of ω_1 is 1-dimensional, there is no room for more than one irreducible subrepresentation.

In the case $n = 3$, the standard representation has the 3 weights

$$\omega_1, \omega_2 - \omega_1, -\omega_2.$$

In the following picture we have drawn these weights in blue as well as the Weyl chamber and the positive simple roots (in red):



7.8 The root system of $\mathfrak{sp}_n$

Recall that the Lie algebra $\mathfrak{sp}_n$ consists of $2n \times 2n$-matrices X over $\mathbb{C}$ such that

$$\Omega X + X^T \Omega = 0,$$

where $\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$. According to Exercise 3.6, these are precisely the matrices with the block form

$$X = \begin{pmatrix} A & B \\ C & -A^T \end{pmatrix}$$

where B and C are symmetric $n \times n$ matrices and A is arbitrary. A basis for $\mathfrak{sp}_n$ is therefore given by the matrices

$$\begin{pmatrix} 0 & E_{jk} + E_{kj} \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ E_{jk} + E_{kj} & 0 \end{pmatrix} \quad (j \neq k), \quad \begin{pmatrix} 0 & E_{jj} \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ E_{jj} & 0 \end{pmatrix},$$

$$\begin{pmatrix} E_{jk} & 0 \\ 0 & -E_{kj} \end{pmatrix} \quad \text{(all } j,k\text{)}.$$

The diagonal matrices form a commutative subalgebra $\mathfrak{h}$ of $\mathfrak{sp}_n$; we claim that this is a Cartan subalgebra. To see this, let Λ be the diagonal matrix with diagonal entries $(\lambda_1, \dots, \lambda_n, -\lambda_1, \dots, -\lambda_n)$.

Exercise 7.6. Verify the following commutator calculations for the diagonal matrix Λ and the basis elements:

$$[\Lambda, \begin{pmatrix} 0 & E_{jk}+E_{kj} \\ 0 & 0 \end{pmatrix}] = (\lambda_j + \lambda_k) \begin{pmatrix} 0 & E_{jk}+E_{kj} \\ 0 & 0 \end{pmatrix},$$

$$[\Lambda, \begin{pmatrix} 0 & 0 \\ E_{jk}+E_{kj} & 0 \end{pmatrix}] = -(\lambda_j + \lambda_k) \begin{pmatrix} 0 & 0 \\ E_{jk}+E_{kj} & 0 \end{pmatrix}$$

$$[\Lambda, \begin{pmatrix} 0 & E_{jj} \\ 0 & 0 \end{pmatrix}] = 2\lambda_j \begin{pmatrix} 0 & E_{jj} \\ 0 & 0 \end{pmatrix},$$

$$[\Lambda, \begin{pmatrix} 0 & 0 \\ E_{jj} & 0 \end{pmatrix}] = -2\lambda_j \begin{pmatrix} 0 & 0 \\ E_{jj} & 0 \end{pmatrix},$$

$$[\Lambda, \begin{pmatrix} E_{jk} & 0 \\ 0 & -E_{kj} \end{pmatrix}] = (\lambda_k - \lambda_j) \begin{pmatrix} E_{jk} & 0 \\ 0 & -E_{kj} \end{pmatrix}.$$

We thus see that $[\Lambda, -]$ acts diagonally in our basis. Moreover, the basis vectors that do not lie in $\mathfrak{h}$ are weight vectors for non-zero weights, so since weight vectors for distinct weights are linearly independent there is no room to expand $\mathfrak{h}$ to a larger commutative subalgebra.

We use the compact real form $\mathfrak{usp}_n$, which consists of the matrices in $\mathfrak{sp}_n$ that are also skew-Hermitian. Then $\mathfrak{h} = \mathfrak{t} \otimes \mathbb{C}$ where $\mathfrak{t}$ consists of the pure-imaginary diagonal matrices in $\mathfrak{sp}_n$. Hence $i\mathfrak{t}$ consists of real diagonal matrices with diagonal $(\lambda_1, \dots, \lambda_n, -\lambda_1, \dots, -\lambda_n)$. As in Example 7.4.21 we can restrict the standard inner product on $\mathfrak{gl}_{2n}(\mathbb{C})$ to get an inner product on $\mathfrak{sp}_n$. By Lemma 7.4.20 this is real on $\mathfrak{usp}_n$, and $[X, -]$ for $X \in \mathfrak{usp}_n$ is skew-Hermitian, since $\mathfrak{usp}_n$ consists matrices that are in particular skew-Hermitian. If we write

$$L_j := \tfrac{1}{2} \begin{pmatrix} E_{jj} & 0 \\ 0 & -E_{jj} \end{pmatrix},$$

then for Λ as above we get

$$\langle L_j, \Lambda \rangle = \tfrac{1}{2}(\lambda_j + \lambda_j) = \lambda_j,$$

so that the roots are

$$\pm L_i \pm L_j \quad (i < j), \quad \pm 2L_j.$$

Since we also have $\langle L_i, L_j \rangle = \frac{1}{2}\delta_{ij}$, the inner products of the roots are given by

$$\langle (-1)^a L_i + (-1)^b L_j, (-1)^c L_k + (-1)^d L_l \rangle = \frac{1}{2}\left((-1)^{a+c}\delta_{ik} + (-1)^{a+d}\delta_{il} + (-1)^{b+c}\delta_{jk} + (-1)^{b+d}\delta_{jl} \right),$$

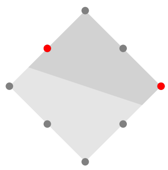
$$\langle (-1)^a L_i + (-1)^b L_j, (-1)^c 2L_k \rangle = (-1)^{a+c}\delta_{ik} + (-1)^{b+c}\delta_{jk},$$
$$\langle (-1)^a 2L_i, (-1)^b 2L_j \rangle = (-1)^{a+b} 2\delta_{ij}.$$

In particular, the roots $\pm 2L_{ij}$ have length $\sqrt{2}$ and those of the form $\pm L_i \pm L_j$ have length 1.

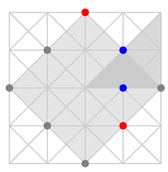
Exercise 7.7. Check that the roots $L_i - L_{i+1}$ $(i = 1, \ldots, n-1)$ and $2L_n$ form a base.

Let's look more closely at the smallest cases. For $n = 1$ there are only the two roots $\pm 2L_1$; these are the same as the roots of $\mathfrak{sl}_2$. Indeed, $\mathfrak{sp}_1$ consists of 2×2 matrices of the form $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$, so $\mathfrak{sp}_1$ is actually identical to the Lie algebra $\mathfrak{sl}_2$ of traceless 2×2 matrices.

In the case $n = 2$, there are 8 roots: $\pm L_1 \pm L_2$ (of length 1) and $\pm 2L_1$, $\pm 2L_2$ (of length $\sqrt{2}$). If we think of L_1 as lying on the x-axis and L_2 on the y-axis, we get the following picture of the roots:



Here the dark grey region indicates the positive side of one possible hyperplane dividing the roots, and the positive simple roots are labelled in red (they are $2L_1$ and $L_2 - L_1$). The next picture illustrates a Weyl chamber and the fundamental weights (though unfortunately for a different choice of positive simple roots, as indicated in red):



Here we have also drawn part of the weight lattice, but note that this is slightly misleading: the points where two diagonal lines cross in the centre of a square are *not* contained in the lattice.

7.9 The root system of $\mathfrak{so}_n$

The roots of $\mathfrak{so}_n$ turn out to behave somewhat differently according to whether n is odd or even. In order to consider the two cases in parallel, it is convenient to define m so that either $n = 2m$ or $n = 2m + 1$.

We know that $\mathfrak{so}_n$ consists of skew-symmetric $n \times n$ matrices over $\mathbb{C}$; as the compact real form we take $\mathfrak{so}_n(\mathbb{R})$, consisting of skew-symmetric real $n \times n$ matrices. As a basis we can take the matrices

$$\tilde{E}_{ij} := E_{ij} - E_{ji}$$

for $i < j$.

Lemma 7.9.1. *Let $\mathfrak{h}$ be the $\mathbb{C}$-subspace of $\mathfrak{so}_n$ spanned by the matrices*

$$L_i := \tilde{E}_{(2i-1)(2i)}$$

for $i = 1, \ldots, m$. This is a commutative subalgebra of $\mathfrak{so}_n$.

Proof. We always have $E_{ij}E_{kl} = \delta_{jk}E_{il}$, so that $[E_{ij}, E_{kl}] = \delta_{jk}E_{il} - \delta_{li}E_{kj}$. From this we compute

$$[\tilde{E}_{ij}, \tilde{E}_{kl}] = \delta_{jk}\tilde{E}_{il} + \delta_{il}\tilde{E}_{jk} + \delta_{ki}\tilde{E}_{li} + \delta_{lj}\tilde{E}_{ki}.$$

In particular we get

$$\begin{aligned}
[L_i, \tilde{E}_{kl}] &= [\tilde{E}_{(2i-1)(2i)}, \tilde{E}_{kl}] \\
&= \delta_{(2i)k}\tilde{E}_{(2i-1)l} + \delta_{(2i-1)l}\tilde{E}_{(2i)k} + \delta_{k(2i-1)}\tilde{E}_{l(2i)} + \delta_{l(2i)}\tilde{E}_{k(2i-1)} \\
&= \begin{cases} -\tilde{E}_{(k+1)l}, & k = 2i-1 \\ -\tilde{E}_{k(l+1)}, & l = 2i-1 \\ \tilde{E}_{(k-1)l}, & k = 2i \\ \tilde{E}_{k(l-1)}, & l = 2i, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

None of the four non-zero cases can occur for $[L_i, L_j] = [\tilde{E}_{(2i-1)(2i)}, \tilde{E}_{(2j-1)(2j)}]$, so this is indeed 0. □

If we take $\mathfrak{t} \subseteq \mathfrak{so}_n(\mathbb{R})$ to be the real vector space spanned by the L_i, then the same calculation shows this is a commutative subalgebra in $\mathfrak{so}_n(\mathbb{R})$, and $\mathfrak{h} \cong \mathfrak{t} \otimes \mathbb{C}$. In fact, $\mathfrak{t}$ is a maximal commutative subalgebra, so that $\mathfrak{h}$ is a Cartan subalgebra.

Exercise 7.8. Use the calculation above of $[L_i, -]$ on the basis $\tilde{E}_{kl}$ of $\mathfrak{so}_n(\mathbb{R})$ to show that if $[L_i, X] = 0$ for all i then $X \in \mathfrak{t}$.

Next, we want to find the roots of $\mathfrak{so}_n$. For this we unfortunately need to find a new basis, since our current one clearly doesn't consist of root vectors.

Exercise 7.9. Show that the four matrices

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

are linearly independent, and so form a basis for $\mathrm{M}_2(\mathbb{C})$. [Hint: To show that a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ can be written uniquely as $\alpha A + \beta B + \gamma C + \delta D$, think of this as a linear equation

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = M \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix}$$

for a 4×4 matrix M, and check that this M is invertible.]

If $n = 2m$, we can extend $L_1, \dots, L_m$ to a basis for $\mathfrak{so}_{2m}$ by thinking of the elements as $m \times m$ matrices of 2×2 blocks, and putting each of these 4 matrices as the single non-zero block above the diagonal. More precisely, let us define (also in the case $n = 2m + 1$) for $1 \leq i < j \leq m$

$$\begin{aligned} A_{ij} &:= \tilde{E}_{(2i-1)(2j-1)} + i\tilde{E}_{(2i-1)(2j)} + i\tilde{E}_{(2i)(2j-1)} - \tilde{E}_{(2i)(2j)} \\ B_{ij} &:= \tilde{E}_{(2i-1)(2j-1)} + i\tilde{E}_{(2i-1)(2j)} - i\tilde{E}_{(2i)(2j-1)} + \tilde{E}_{(2i)(2j)} \\ C_{ij} &:= \tilde{E}_{(2i-1)(2j-1)} + i\tilde{E}_{(2i-1)(2j)} - i\tilde{E}_{(2i)(2j-1)} - \tilde{E}_{(2i)(2j)} \\ D_{ij} &:= \tilde{E}_{(2i-1)(2j-1)} - i\tilde{E}_{(2i-1)(2j)} - i\tilde{E}_{(2i)(2j-1)} + \tilde{E}_{(2i)(2j)}. \end{aligned}$$

Proposition 7.9.2. *Write $H = \sum_{i=1}^m a_i L_i$ for an element of $\mathfrak{h}$. For all $1 \leq k < l \leq m$ we then have:*

(1) The matrix A_{kl} is a root vector for the root $\alpha_{kl}^A(H) = i(a_k + a_l)$.

(2) The matrix B_{kl} is a root vector for the root $\alpha_{kl}^B(H) = -i(a_k + a_l)$.

(3) The matrix C_{kl} is a root vector for the root $\alpha_{kl}^C(H) = i(a_k - a_l)$.

(4) The matrix D_{kl} is a root vector for the root $\alpha_{kl}^D(H) = -i(a_k - a_l)$.

Proof. We'll prove (1), and leave the sign changes needed for the 3 other cases to any exceptionally diligent readers. We compute (omitting the cases that are 0):

$$[L_i, \tilde{E}_{(2k-1)(2l-1)}] = \begin{cases} -\tilde{E}_{(2k)(2l-1)}, & i = k \\ -\tilde{E}_{(2k-1)(2l)}, & i = l \end{cases}$$

$$[L_i, \tilde{E}_{(2k-1)(2l)}] = \begin{cases} -\tilde{E}_{(2k)(2l)}, & i = k \\ \tilde{E}_{(2k-1)(2l-1)}, & i = l \end{cases}$$

$$[L_i, \tilde{E}_{(2k)(2l-1)}] = \begin{cases} \tilde{E}_{(2k-1)(2l-1)}, & i = k \\ -\tilde{E}_{(2k)(2l)}, & i = l \end{cases}$$

$$[L_i, \tilde{E}_{(2k)(2l)}] = \begin{cases} \tilde{E}_{(2k-1)(2l)}, & i = k \\ \tilde{E}_{(2k)(2l-1)}, & i = l \end{cases}$$

$$\begin{aligned} [L_i, A_{kl}] &= [L_i, \tilde{E}_{(2k-1)(2l-1)} + i\tilde{E}_{(2k-1)(2l)} + i\tilde{E}_{(2k)(2l-1)} - \tilde{E}_{(2k)(2l)}] \\ &= \begin{cases} -\tilde{E}_{(2k)(2l-1)} - i\tilde{E}_{(2k)(2l)} + i\tilde{E}_{(2k-1)(2l-1)} - \tilde{E}_{(2k-1)(2l)}, & i = k \\ -\tilde{E}_{(2k-1)(2l)} + i\tilde{E}_{(2k-1)(2l-1)} - i\tilde{E}_{(2k)(2l)} - \tilde{E}_{(2k)(2l-1)}, & i = l \end{cases} \\ &= \begin{cases} iA_{kl}, & i = k, l \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Thus $[L_i, A_{kl}] = \alpha^A_{kl}(L_i)A_{kl}$ for all i, and since the L_i are a basis for $\mathfrak{h}$ this means A_{kl} is a root vector for α^A_{kl}, as required. □

For $n = 2m$ we have now found a basis of root vectors, while we still need a few more vectors to span $\mathfrak{so}_{2m+1}$: We define

$$F_j := \tilde{E}_{(2j-1)(2m+1)} + i\tilde{E}_{(2j)(2m+1)}, \quad G_j := \tilde{E}_{(2j-1)(2m+1)} - i\tilde{E}_{(2j)(2m+1)}$$

Together these span the non-diagonal entries in the last $((2m+1)\text{th})$ column, so together with our previous roots vectors and the L_i they give a basis for $\mathfrak{so}_{2m+1}$.

Proposition 7.9.3. *Assume $n = 2m + 1$, and write $H = \sum_{i=1}^m a_i L_i$ for an element of $\mathfrak{h}$. For all $1 \leq j \leq m$ we then have:*

(1) The matrix F_j is a root vector for the root $\alpha^F_j(H) = ia_j$.

(2) The matrix G_j is a root vector for the root $\alpha^G_j(H) = -ia_j$.

Proof. We prove case (1) by computing:

$$[L_k, \tilde{E}_{(2j-1)(2m+1)}] = \begin{cases} -\tilde{E}_{(2j)(2m+1)}, & k = j \\ 0, & k \neq j \end{cases}$$

$$[L_k, \tilde{E}_{(2j)(2m+1)}] = \begin{cases} \tilde{E}_{(2j-1)(2m+1)}, & k = j \\ 0, & k \neq j \end{cases}$$

$$[L_k, F_j] = \begin{cases} -\tilde{E}_{(2j)(2m+1)} + i\tilde{E}_{(2j-1)(2m+1)}, & k = j \\ 0, & k \neq j \end{cases}$$

$$= \begin{cases} iF_j, & k = j \\ 0, & k \neq j \end{cases}$$

Thus $[L_k, F_j] = \alpha_j^F(L_k)F_j$ for all k. Since the L_k form a basis for $\mathfrak{h}$, this means that F_j is a root vector for α_j^F, as required. □

To get a good inner product on $\mathfrak{so}_n$ we can again use Lemma 7.4.20 to conclude that the restriction of the standard inner product on $\mathfrak{gl}_n(\mathbb{C})$ works, since our compact real form $\mathfrak{so}_n(\mathbb{R})$ consists of matrices that are in particular skew-Hermitian.

Since E_{ij} for all i, j is an orthonormal basis for $\mathfrak{gl}_n(\mathbb{C})$, we then get

$$\langle L_i, L_j \rangle = 2\delta_{ij},$$

so that in $i\mathfrak{t}$ we can identify the roots as

$$\alpha_{kl}^A = -\frac{i}{2}(L_k + L_l), \quad \alpha_{kl}^B = \frac{i}{2}(L_k + L_l), \quad \alpha_{kl}^C = -\frac{i}{2}(L_k - L_l), \quad \alpha_{kl}^D = \frac{i}{2}(L_k - L_l)$$

for $1 \leq k < l \leq m$, as well as

$$\alpha_j^F = -\frac{i}{2}L_j, \quad \alpha_j^G = \frac{i}{2}L_j$$

when n is odd. Thus the roots of $\mathfrak{so}_{2m}$ in $i\mathfrak{t}$ are

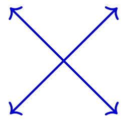
$$\frac{i}{2}(\pm L_k \pm L_l), \quad k < l.$$

In particular, these roots all have the same length. For $\mathfrak{so}_{2m+1}$ we also have the additional roots $\pm\frac{i}{2}L_k$ for $1 \leq k \leq m$; note that these are shorter than the first type of root by a factor of $\sqrt{2}$.

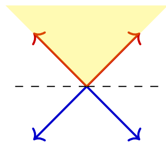
Let's see what the roots look like for small values of n:

Example 7.9.4. For $\mathfrak{so}_2$, $m = 1$ and there are no roots. (Indeed, $\mathfrak{so}_2$ is abelian.) For $\mathfrak{so}_3$, $m = 1$ and we have the two roots $\pm L_1$. (Of course, $\mathfrak{so}_3 \cong \mathfrak{sl}_2$ and their roots are the same.)

Example 7.9.5. For $\mathfrak{so}_4$, $m = 2$ and we have 4 roots: $\frac{i}{2}(\pm L_1 \pm L_2)$. The roots $\frac{i}{2}(L_1 + L_2)$ and $\frac{i}{2}(L_1 - L_2)$ are orthogonal, so we can draw the roots as follows:

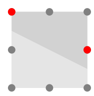


Any of the four possible choices of two orthogonal roots is a base, and the Weyl chambers are the four quadrants. Let's draw one option:

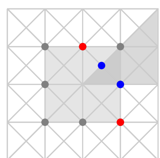


Recall that we had an isomorphism $\mathfrak{so}_4 \cong \mathfrak{sl}_2 \oplus \mathfrak{sl}_2$ — it is not a coincidence that the roots of $\mathfrak{so}_4$ look like two orthogonal copies of those of $\mathfrak{sl}_2$!

Example 7.9.6. For $\mathfrak{so}_5$, $m = 2$ and we have 8 roots: $\frac{i}{2}(\pm L_1 \pm L_2)$ as before, but now also $\pm\frac{i}{2}L_1, \pm\frac{i}{2}L_2$. We get the following picture, where the dark grey region indicates a choice of hyperplane dividing the roots, with the corresponding positive simple roots marked in red:



The next picture shows a Weyl chamber (again for a different choice of positive simple roots, in red) and the corresponding fundamental weights (in blue), and part of the weight lattice:



Note that in this picture the points where two diagonal lines cross *are* weights!

Remark 7.9.7. The roots of $\mathfrak{so}_5$ look suspiciously like those of $\mathfrak{sp}_2$, except for being rotated by 45 degrees. Indeed, there is an isomorphism of Lie algebras $\mathfrak{so}_5 \cong \mathfrak{sp}_2$. While we're at it, there's one more notable isomorphism worth mentioning: $\mathfrak{so}_6 \cong \mathfrak{sl}_4$.

7.10 (Anti)symmetric powers and $\mathfrak{sl}_n$-representations

Our goal in this section is to define the fundamental representations of $\mathfrak{sl}_n$ (i.e. those whose highest weights are the fundamental weights $\omega_1, \dots, \omega_{n-1}$). It turns out that we can obtain these by applying a general construction, the exterior or antisymmetric powers, to the standard representation of $\mathfrak{sl}_n$. We start by introducing this construction in vector spaces, together with its slightly simpler cousin, the symmetric powers.

Notation 7.10.1. We write $V^{\otimes n}$ for the n-fold iterated tensor product $V \otimes \cdots \otimes V$ of V with itself.

Definition 7.10.2. The symmetric group S_n acts on $V^{\otimes n}$ for any vector space V by permuting the n copies of V. More precisely, we define the action of a permutation σ by setting

$$\sigma \cdot v_1 \otimes \cdots \otimes v_n := v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$$

and extending this linearly to general elements of $V^{\otimes n}$. The *nth symmetric power* $\mathrm{Sym}^n V$ is the subspace of $V^{\otimes}$ consisting of elements that are fixed by this action, i.e. $X \in V^{\otimes n}$ lies in $\mathrm{Sym}^n V$ if and only if $\sigma \cdot X = X$ for all $\sigma \in S_n$.

Exercise 7.10. Show that we can define a linear map $s \colon V^{\otimes n} \to \mathrm{Sym}^n V$ by $s(X) = \sum_{\sigma \in S_n} \sigma \cdot X$. Check that if $X \in \mathrm{Sym}^n V$ then $s(X) = (n!)X$, so that (provided our base field is of characteristic zero) the composite

$$\mathrm{Sym}^n V \hookrightarrow V^{\otimes n} \xrightarrow{s} \mathrm{Sym}^n V$$

is an automorphism of $\mathrm{Sym}^n V$. In particular, s is surjective; show that its kernel is the subspace spanned by vectors of the form

$$v_1 \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$$

for $\sigma \in S_n$, $v_i \in V$, and conclude that $\mathrm{Sym}^n V$ is also the quotient of $V^{\otimes n}$ where we impose these relations.

Notation 7.10.3. We often denote the element $s(v_1 \otimes \cdots \otimes v_n)$ of $\mathrm{Sym}^n V$ as $v_1 \cdots v_n$. Then we have $v_1 \cdots v_n = v_{\sigma(1)} \cdots v_{\sigma(n)}$ for all $\sigma \in S_n$.

Exercise 7.11. Suppose $e_1, \ldots, e_m$ is a basis of V. Prove that the vectors $s(e_{i_1} \otimes \cdots \otimes e_{i_n})$ with $i_1 \leq i_2 \leq \cdots \leq i_n$ form a basis for $\mathrm{Sym}^n V$.

Now we turn to the exterior (or antisymmetric) powers:

Definition 7.10.4. Recall that the *sign* $\operatorname{sgn} \sigma \in \{\pm 1\}$ of a permutation $\sigma \in S_n$ is defined by taking the sign of a transposition to be -1 and requiring that $\operatorname{sgn}(\sigma\tau) = (\operatorname{sgn} \sigma)(\operatorname{sgn} \tau)$; alternatively, $\operatorname{sgn} \sigma$ is the determinant of σ when regarded as an automorphism of $\mathbb{R}^n$ that permutates the coordinates. We can twist the action of S_n on $V^{\otimes}$ by the sign to get a new action by defining

$$\sigma \cdot v_1 \otimes \cdots \otimes v_n := (\operatorname{sgn} \sigma) \cdot v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)},$$

and extending this linearly. The *nth exterior* (or *antisymmetric*) *power* $\Lambda^n V$ is the subspace of $V^{\otimes}$ consisting of elements that are fixed by this action, i.e. $X \in V^{\otimes n}$ lies in $\mathrm{Sym}^n V$ if and only if $\sigma \cdot X = X$ for all $\sigma \in S_n$.

Exercise 7.12. Show that we can define a linear map $a\colon V^{\otimes n} \to \Lambda^n V$ by $a(X) = \sum_{\sigma \in S_n} \sigma \cdot X$. Check that if $X \in \Lambda^n V$ then $a(X) = (n!)X$, so that (provided our base field is of characteristic zero) the composite

$$\Lambda^n V \hookrightarrow V^{\otimes n} \xrightarrow{a} \Lambda^n V$$

is an automorphism of $\Lambda^n V$. In particular, a is surjective; show that its kernel is the subspace spanned by vectors of the form

$$v_1 \otimes \cdots \otimes v_n - \operatorname{sgn}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$$

for $\sigma \in S_n$, $v_i \in V$, and conclude that $\Lambda^n V$ is also the quotient of $V^{\otimes n}$ where we impose these relations.

Notation 7.10.5. We often denote the element $a(v_1 \otimes \cdots \otimes v_n)$ of $\Lambda^n V$ as $v_1 \wedge \cdots \wedge v_n$. Note that $v_1 \wedge \cdots \wedge v_n = (\operatorname{sgn}\sigma) v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)}$ for all $\sigma \in S_n$. In particular, if two of the v_i's are equal this must be 0.

Exercise 7.13. Suppose $e_1, \ldots, e_m$ is a basis of V. Prove that the vectors $a(e_{i_1} \otimes \cdots \otimes e_{i_n})$ with $i_1 < i_2 < \cdots < i_n$ form a basis for $\Lambda^n V$.

Now we come apply these constructions to representations:

Proposition 7.10.6. *Let (V, ρ) be a representation of a matrix group G or a Lie algebra $\mathfrak{g}$. Then* $\operatorname{Sym} V$ *and* $\Lambda^n V$ *are invariant subspaces of the tensor product representation* $(V^{\otimes n}, \rho^{\otimes n})$.

Proof. We consider the symmetric case; the antisymmetric case is almost identical. We first assume ρ is a group representation. For $X \in V^{\otimes}$, $g \in G$, $\sigma \in S_n$, we claim that $\rho^{\otimes n}(g)(\sigma \cdot X) = \sigma \cdot \rho^{\otimes}(g)(X)$. Since both sides are linear, it suffices to check this when $X = v_1 \otimes \cdots \otimes v_n$, in which case we have

$$\rho^{\otimes n}(g)(\sigma \cdot X) = \rho^{\otimes n}(g)(v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}) = \rho(g) v_{\sigma(1)} \otimes \cdots \rho(g) v_{\sigma(n)} = \sigma \cdot \rho(g)(X).$$

If $X \in \operatorname{Sym}^n V$ we thus have $\sigma \cdot (\rho(g)X) = \rho(g)(\sigma \cdot X) = \rho(g)(X)$, so that $\rho(g)X$ is also in $\operatorname{Sym}^n V$, as required. The Lie algebra version is almost the same, except that we then compute for $Q \in \mathfrak{g}$ that

$$\begin{aligned}
\rho^{\otimes n}(Q)(\sigma \cdot X) &= \rho^{\otimes n}(Q)(v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}) \\
&= \sum_{i=1}^{n} v_{\sigma(1)} \otimes \cdots \otimes \rho(Q) v_{\sigma(i)} \otimes \cdots \otimes v_{\sigma(n)} \\
&= \sum_{j=1}^{n} \sigma \cdot (v_1 \otimes \cdots \otimes \rho(Q) v_j \otimes \cdots \otimes v_n) \\
&= \sigma \cdot \rho^{\otimes n}(Q)(X),
\end{aligned}$$

where we use that σ is a permutation to reorder the sum with $j = \sigma(i)$. □

Exercise 7.14. Show that for any vector space V we have a direct sum decomposition $V \otimes V = \mathrm{Sym}^2 V \oplus \Lambda^2 V$. [Hint: write $u \otimes v$ as $\frac{1}{2}(u \otimes v + v \otimes u) + \frac{1}{2}(u \otimes v - v \otimes u)$.] Conclude that $V \otimes V$ is never an irreducible representation.

Exercise 7.15. Let (V, ρ) be a finite-dimensional complex representation of a complex semisimple Lie algebra $\mathfrak{g}$. Suppose $v_1, \dots, v_n$ are weight vectors in V for weights $\mu_1, \dots, \mu_n$, respectively. Show that:

(i) $v_1 \otimes \cdots \otimes v_n$ is a weight vector for $\mu_1 + \cdots + \mu_n$ in $V^{\otimes n}$.

(ii) $v_1 \cdots v_n$ is a weight vector for $\mu_1 + \cdots + \mu_n$ in $\mathrm{Sym}^n V$.

(iii) If the v_i are linearly independent (for example if the weights μ_i are all distinct), then $v_1 \wedge \cdots \wedge v_n$ is a weight vector for $\mu_1 + \cdots + \mu_n$ in $\Lambda^n V$.

Using this exercise we immediately conclude the following:

Proposition 7.10.7. *Suppose (V, ρ) is a finite-dimensional complex representation of a complex semisimple Lie algebra $\mathfrak{g}$ such that V is the direct sum of its weight spaces, i.e. there exists a basis of weight vectors $v_1, \dots, v_k$ with weights $\mu_1, \dots, \mu_k$ (possibly repeated). Then*

(i) The vectors $v_{i_1} \otimes \cdots \otimes v_{i_n}$ form a basis of weight vectors for $V^{\otimes n}$ with weights $\mu_{i_1} + \cdots + \mu_{i_n}$, respectively.

(ii) The vectors $v_{i_1} \cdots v_{i_n}$ with $i_1 \leq \cdots \leq i_n$ form a basis of weight vectors for $\mathrm{Sym}^n V$ with weights $\mu_{i_1} + \cdots + \mu_{i_n}$, respectively.

(iii) The vectors $v_{i_1} \wedge \cdots \wedge v_{i_n}$ with $i_1 < \cdots < i_n$ form a basis of weight vectors for $\Lambda^n V$ with weights $\mu_{i_1} + \cdots + \mu_{i_n}$, respectively.

Corollary 7.10.8. *Let $V = \mathbb{C}^n$ be the standard representation of $\mathfrak{sl}_n$. Then $\Lambda^k V$ is an irreducible representation with highest weight ω_k for $k = 1, \dots, n-1$.*

Proof. Recall from Example 7.7.13 that the standard basis $e_1, \dots, e_n$ for V consists of weight vectors for the (distinct) weights

$$\mu_j = \begin{cases} \omega_1, & j = 1, \\ \omega_j - \omega_{j-1}, & 1 < j < n, \\ -\omega_{n-1}, & j = n \end{cases}$$

From Proposition 7.10.7 it follows that $e_{i_1} \wedge \cdots \wedge e_{i_k}$ for $i_1 < \cdots < i_k$ is a basis for $\Lambda^k V$ consisting of weight vectors for the weights $\mu_{i_1} + \cdots + \mu_{i_k}$. Note in particular that $e_1 \wedge e_2 \wedge \cdots \wedge e_k$ is a weight vector for

$$\mu_1 + \mu_2 + \cdots + \mu_k = \omega_1 + (\omega_2 - \omega_1) + \cdots + (\omega_k - \omega_{k-1}) = \omega_k,$$

so ω_k is a weight of $\Lambda^k V$. To see that it is the highest weight, recall from Example 7.7.13 that if $i < j$ we have $\mu_i - \mu_j = L_{ij}$. For any sequence $1 \leq i_1 < \cdots < i_k$ we must have $i_j \geq j$ (with equality possible only if it holds for all j). Therefore

$$\omega_k - (\mu_{i_1} + \cdots + \mu_{i_k}) = (\mu_1 - \mu_{i_1}) + \cdots + (\mu_k - \mu_{i_k}) = L_{1 i_1} + \cdots + L_{k i_k},$$

which is clearly a sum of the positive simple roots with non-negative coefficients, so that $\omega_k \geq (\mu_{i_1} + \cdots + \mu_{i_k})$. We leave the proof of irreducibility as an exercise. □

Exercise 7.16. Recall that in $\mathfrak{sl}_n$, the matrix E_{ij} is a root vector for L_{ij}. In the standard representation $V = \mathbb{C}^n$ we have

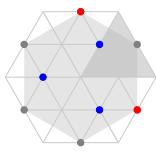
$$E_{ij}e_k = \begin{cases} e_i, & j = k, \\ 0, & \text{otherwise.} \end{cases}$$

Show from this that we can obtain all basis vectors $e_{i_1} \wedge \cdots \wedge e_{i_k}$ of $\Lambda^k V$ from $e_1 \wedge e_2 \wedge \cdots \wedge e_k$ by acting with the matrices $E_{i_j j}$. Conclude that $\Lambda^k V$ is highest weight cyclic, and so irreducible.

Example 7.10.9. For $n = 3$, the representation $\Lambda^2 V$ of $\mathfrak{sl}_3$ has the weights $\mu_i + \mu_k$ for $1 \leq i < k \leq 3$, i.e.

$$\begin{aligned} \mu_1 + \mu_2 &= \omega_1 + (\omega_2 - \omega_1) = \omega_2, \\ \mu_2 + \mu_3 &= (\omega_2 - \omega_1) - \omega_2 = -\omega_1, \\ \mu_1 + \mu_3 &= \omega_1 - \omega_2. \end{aligned}$$

We can draw these weights (in blue) in our usual diagram as follows:



Exercise 7.17. Show that if V is the standard representation of $\mathfrak{sl}_3$, then the representations $\Lambda^2 V$ and V^* are isomorphic.

Remark 7.10.10. In fact, this duality isomorphism is true in general: for any unitary representation (V, ρ) (of a matrix group or Lie algebra) the dual representation V^* is isomorphic to the representation $\Lambda^{n-1} V$. More generally, $(\Lambda^k V)^*$ is isomorphic to $\Lambda^{n-k} V$.

Remark 7.10.11. For $\mathfrak{sp}_n$ we can also build the fundamental representations out of the standard representation, by a similar (but not quite as straightforward) procedure. In the case of $\mathfrak{so}_n$, however, only *half* of the representations can be built from the standard representation. This is because the standard representation lifts to the (complexified) standard representation of $\mathrm{SO}_n(\mathbb{R})$, but not all representations lift (since $\mathrm{SO}_n(\mathbb{R})$ is not simply connected). We saw the case $n = 3$ of this phenomenon back in Proposition 6.1.5. To find all representations of $\mathfrak{so}_n$ we therefore need to work with the universal covers of $\mathrm{SO}_n(\mathbb{R})$, which are the so-called *spin* groups. These are also matrix groups, and we can use *their* standard representations to find all irreducible $\mathfrak{so}_n$-representations.

Chapter 8

(⋆) Classification of complex simple Lie algebras

8.1 Abstract root systems

The following definition abstracts some of the key properties we've seen for the roots of a complex semisimple Lie algebra:

Definition 8.1.1. A *root system* consists of a real inner product space $(E, \langle -, - \rangle)$ together with a finite set R of elements of E (called *roots*) satisfying the following properties:

(1) The roots span E.

(2) If $\alpha \in R$ then $-\alpha \in R$, and no other multiples of α lie in R.

(3) If α and β lie in R then
$$s_\alpha \beta := \beta - 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha$$
also lies in R.

(4) For all $\alpha, \beta \in R$, the real number
$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}$$
is an integer.

The dimension of E is called the *rank* of the root system.

Note that many of the results about roots we proved earlier, such as the restrictions on their possible angles and relative lengths and the existence of a base, only used the properties contained in this definition.

Definition 8.1.2. An *isomorphism* between root systems (E, R) and (E', R') is a linear isomorphism $\phi\colon E \xrightarrow{\sim} E'$ such that $\phi(R) = R'$ and such that for $\alpha \in R$ and $x \in E$ we have

$$\phi(s_\alpha x) = s_{\phi(\alpha)}\phi(x).$$

(Note that ϕ only needs to preserve the reflections in the roots, *not* the inner product. In particular, an isomorphism of root systems is allowed to rescale lengths.[1])

Observation 8.1.3. If ϕ is an isomorphism of root systems from (E, R) to (E', R'), then

$$\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = \frac{\langle \phi\alpha, \phi\beta \rangle}{\langle \phi\alpha, \phi\alpha \rangle}$$

for all $\alpha, \beta \in R$. In particular, ϕ preserves the angles between roots.

Examples 8.1.4. Of course, the roots of a complex semisimple Lie algebra form a root system. For the simple Lie algebras these have special names:

- The root system of $\mathfrak{sl}_{n+1}$ $(n \geq 1)$ is called A_n.
- The root system of $\mathfrak{so}_{2n+1}$ $(n \geq 2)$ is called B_n.
- The root system of $\mathfrak{sp}_n$ $(n \geq 2)$ is called C_n.
- The root system of $\mathfrak{so}_{2n}$ $(n \geq 3)$ is called D_n.

Note that the subscript here denotes the *rank* of the root system.

Example 8.1.5. If (E, R) is a root system, we can define the *coroot* H_α of $\alpha \in R$ to be $\frac{2\alpha}{\langle \alpha, \alpha \rangle}$. Let $R^\vee$ be the set of coroots. Then it can be shown (see [2, §8.3]) that $(E, R^\vee)$ is also a root system, known as the *dual root system* of (E, R). Moreover, $(R^\vee)^\vee = R$. Of the root systems in the previous example, both A_n and D_n are self-dual ($A_n^\vee \cong A_n$, $D_n^\vee \cong D_n$) while B_n is dual to C_n.

Definition 8.1.6. Suppose (E, R) and (E', R') are root systems. Then their *direct sum* $(E, R) \oplus (E', R')$ is given by the vector space $E \oplus E'$, equipped with the inner product that restricts to those on E, E' and makes E orthogonal to E', together with the union of the roots $R \cup R'$. A root system (E'', R'') is called *reducible* if it is isomorphic to a direct sum $(E, R) \oplus (E', R')$. If no such decomposition exists, we say that (E'', R'') is *irreducible*.

Theorem 8.1.7. *The root system of a complex semisimple Lie algebra* $\mathfrak{g}$ *is irreducible if and only if* $\mathfrak{g}$ *is simple.*

More precisely, a decomposition of the root system as a direct sum corresponds precisely to a decomposition of $\mathfrak{g}$ as a direct sum of Lie algebras; see [2, §7.6] for a proof.

[1] Potentially we can even scale by different factors in orthogonal directions, but this can only happen if the root system is decomposable.

Example 8.1.8. From our description of the roots of $\mathfrak{so}_4$ in Example 7.9.5 we see that these are a direct sum of two copies of the roots of $\mathfrak{sl}_2$, corresponding to the Lie algebra isomorphism $\mathfrak{so}_4 \cong \mathfrak{sl}_2 \oplus \mathfrak{sl}_2$.

We've already seen that the root system encodes much of the structure of a semisimple Lie algebra, but in fact we can recover the entire Lie algebra from it:

Theorem 8.1.9 (Serre)**.** *Let (E, R) be a root system, and suppose $\Delta = \{\alpha_1, \ldots, \alpha_r\}$ is a base for it. Let $\mathfrak{g}$ be the Lie algebra generated by elements H_i, X_i, Y_i for $i = 1, \ldots, r$, subject to the following relations:*

(1) $[H_i, H_j] = 0.$

(2) $[X_i, Y_j] = \begin{cases} H_i, & i = j, \\ 0, & i \neq j. \end{cases}$

(3) $[H_i, X_j] = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle} X_j.$

(4) $[H_i, Y_j] = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle} Y_j.$

(5) $\mathrm{ad}_{\mathfrak{g}}(X_i)^m(X_j) = 0$ *for* $m = 1 - 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle}.$

(6) $\mathrm{ad}_{\mathfrak{g}}(Y_i)^m(Y_j) = 0$ *for* $m = 1 - 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle}.$

Then $\mathfrak{g}$ is a (finite-dimensional) complex semisimple Lie algebra whose root system is (E, R), and up to isomorphism this is the unique *such Lie algebra.*

8.2 Examples in rank 2

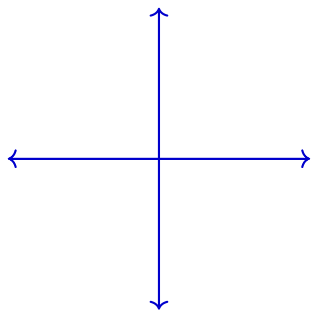
We have already seen 3 distinct (non-isomorphic) root systems in rank 2: $A_1 \oplus A_1$ (the roots of $\mathfrak{so}_4$), A_2 (the roots of $\mathfrak{sl}_3$), and B_2 (the roots of $\mathfrak{so}_5 \cong \mathfrak{sp}_2$). We will now classify the possible roots systems with rank 2 — we will see that there is one further "exotic" example.

To that end, suppose (E, R) is a root system of rank 2; we may as well assume that E is $\mathbb{R}^2$ with the standard inner product. Let θ be the smallest angle between two elements of R. We claim that $\theta \leq \pi/2$. Indeed, if α, β are two linearly independent elements of R and the angle between them is $> \pi/2$, then the angle between α and $-\beta$ is necessarily $< \pi/2$. From Observation 7.4.16 we see that the only possible options for θ are $\pi/2, \pi/3, \pi/4$, and $\pi/6$.

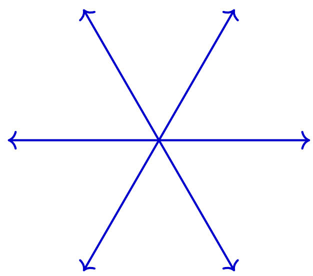
Now suppose α and β are two roots such that the angle between them is the minimal angle θ. Then $\gamma := -s_\beta(\alpha)$ is also a root. Here $-s_\beta$ is an orientation-reversing orthogonal automorphism of $\mathbb{R}^2$, so the angle between γ and β is also θ, but γ lies on the opposite side of β than α. Thus the angle between α and γ is 2θ. Similarly $-s_\gamma(\beta)$ is at an angle 3θ from α, and continuing in the same way we see that there are roots with angles $n\theta$ from α for all n. Since $2\pi/\theta$ is an integer, we must eventually come back to α (since no multiple other than $-\alpha$ can be a root). Moreover, this process must produce all the roots, as any additional root would form an angle $< \theta$ with one of these roots.

We conclude that R consists of n evenly spaced vectors where $\theta = 2\pi/n$, with n either $4, 6, 8$ or 12.

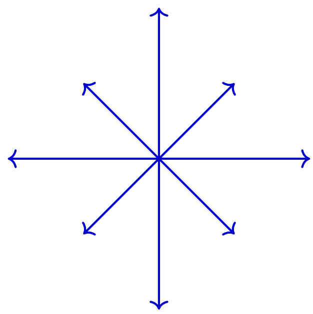
If $n = 4$, we get two orthogonal roots and their negatives; the orthogonal roots may have any relative lengths, but we can also rescale them independently in an isomorphism of root systems, so up to isomorphism there is only one example. This gives $A_1 \oplus A_1$:



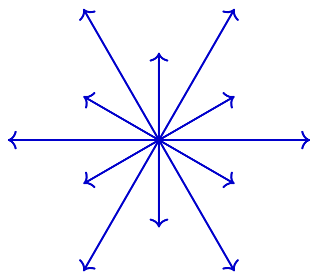
If $n = 6$, the angle between adjacent roots is $\pi/3$, and Observation 7.4.16 implies that two roots with this angle between them must have the same length. Thus all roots have the same length, and and isomorphism must be a constant rescaling of an orthogonal isomorphism. This gives A_2:



If $n = 8$, the angle between adjacent roots is $\pi/4$, and Observation 7.4.16 implies that for two roots with this angle between them, one root must be longer than the other by a factor of $\sqrt{2}$. must have the same length. Again this forces an isomorphism to by a constant rescaling of an orthogonal isomorphism. This gives B_2:



The remaining case is where $n = 12$. Here the angle between adjacent roots is $\pi/6$, and Observation 7.4.16 implies that for two roots with this angle between them, one root must be longer than the other by a factor of $\sqrt{3}$. Once more, this forces an isomorphism to by a constant rescaling of an orthogonal isomorphism. This gives a new example of a root system, called G_2:



Exercise 8.1. Check that G_2 is a root system.

Remark 8.2.1. In rank 3 there are three irreducible root systems: A_3, B_3, C_3. See [2, §8.9] for illustrations of these.

8.3 Dynkin diagrams and classification

According to Theorem 8.1.9, there exists a unique simple Lie algebra, unimaginatively called $\mathfrak{g}_2$, whose root system is G_2; since G_2 has 12 roots, this has dimension $14 = 12+2$. This is the smallest example of an *exceptional simple Lie algebra*, meaning one that is not one of the *classical* simple Lie algebras $\mathfrak{sl}_n$, $\mathfrak{sp}_n$, $\mathfrak{so}_n$. The reason other examples are called "exceptional" is that it turns out that there exist precisely 5 such exceptions.[2]

Since a simple Lie algebra is determined (up to isomorphism) by its root system, this classification of simple Lie algebras reduces to a combinatorial classification of the irreducible root systems. Our goal here is simply to give a precise statement of the result, for which it is convenient to first introduce *Dynkin diagrams*, which give a simple way to encode the combinatorial data of a root system; see for instance [3, §11] for details of the proof.

[2] G_2 is also exceptional in the additional sense that it is the only root system to feature roots at an angle of $\pi/6$.

Definition 8.3.1. Let (E, R) be a root system with a base $\Delta = \{\alpha_1, \ldots, \alpha_r\}$. The corresponding *Dynkin diagram* is an (unoriented) graph with r vertices $v_1, \ldots, v_r$. The vertices v_i and v_j are connected by n edges where n is determined by the angle θ between α_i and α_j as follows (recall that $\theta \geq \pi/2$ by Lemma 7.6.2):

θ	n
$\pi/2$	0
$2\pi/3$	1
$3\pi/4$	2
$5\pi/6$	3

In addition, if α_i and α_j are not orthogonal and have different lengths, the edges between them are decorated with an arrow pointing from the longer root to the shorter root.[3] An *isomorphism* between two Dynkin diagrams is a bijection between the vertices that preserves the numbers of edges and the directions of any arrows.

Observation 8.3.2. Any pair of bases for the same root system are related by an element of the Weyl group, which preserves angles and lengths. This must therefore give an isomorphism of the corresponding Dynkin diagrams, so that (up to isomorphism) the Dynkin diagram only depends on the root system.

Examples 8.3.3. The Dynkin diagrams of the three irreducible rank 2 root systems are as follows:

- A_2:
- B_2:
- G_2:

The Dynkin diagram for $A_1 \oplus A_1$ is simply two vertices with no edge between them.

Proposition 8.3.4.

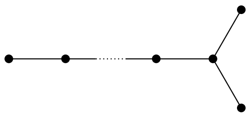
(i) A root system is irreducible if and only if its Dynkin diagram is connected.

(ii) If the Dynkin diagrams of two roots systems are isomorphic, then the root systems are themselves isomorphic.

More precisely, the connected components of the Dynkin diagram correspond to a direct sum decomposition of the root system; see [2, Proposition 8.32] for the proof.

Examples 8.3.5. The Dynkin diagrams of the classical simple Lie algebras are the following:

3Think of the arrowhead as a “>” symbol.

- A_n ($\mathfrak{sl}_{n+1}$):
- B_n ($\mathfrak{so}_{2n+1}$):
- C_n ($\mathfrak{sp}_n$):
- D_n ($\mathfrak{so}_{2n}$): 

Here we can see several evident isomorphisms of Dynkin diagrams of small ranks, reflecting isomorphims of the corresponding Lie algebras:

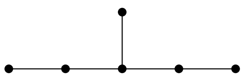
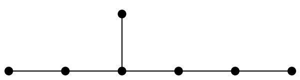
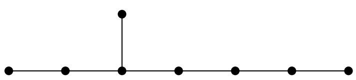
- $A_1 \cong B_1 \cong C_1$, corresponding to $\mathfrak{sl}_2 \cong \mathfrak{so}_3 \cong \mathfrak{sp}_2$. ("$D_1$" is undefined, reflecting the fact that $\mathfrak{so}_2$ is abelian.)
- $B_2 \cong C_2$, corresponding to $\mathfrak{so}_5 \cong \mathfrak{sp}_2$.
- $D_2 \cong A_1 \amalg A_1$, corresponding to $\mathfrak{so}_4 \cong \mathfrak{sl}_2 \oplus \mathfrak{sl}_2$.
- $D_3 \cong A_3$, corresponding to $\mathfrak{so}_6 \cong \mathfrak{sl}_4$.

(Note also that according to Proposition 8.3.4, the fact that these Dynkin diagrams are connected immediately implies that the Lie algebras $\mathfrak{sl}_n$, $\mathfrak{sp}_n$ and $\mathfrak{so}_n$ ($n > 4$) are not just semisimple, but actually simple.)

We are now ready to state the full classification of simple complex Lie algebras, or equivalently of irreducible root systems:

Theorem 8.3.6. *An irreducible root system is isomorphic to exactly one of the following: A_n ($n \geq 1$), B_n ($n \geq 2$), C_n ($n \geq 3$), D_n ($n \geq 4$), E_6, E_7, E_8, F_4, G_2.*

Here the five exceptional cases have the following Dynkin diagrams:

- E_6: 
- E_7: 
- E_8: 
- F_4:
- G_2:

Appendix A

Quick review of some background

A.1 Topological spaces

Definition A.1.1. A *topological space* is a set X equipped with a collection $\mathcal{T}_X$ of subsets of X, such that

- $\emptyset, X \in \mathcal{T}_X$,
- if $U_i \in \mathcal{T}_X$ for all $i \in I$ (where I can be any set) then $\bigcup_{i \in I} U_i \in \mathcal{T}_X$,
- if $U, U' \in \mathcal{T}_X$ then $U \cap U' \in \mathcal{T}_X$.

The collection $\mathcal{T}_X$ is called a *topology* on X and the elements of $\mathcal{T}_X$ are the *open* subsets of X. We usually just say that X is a topological space without mentioning $\mathcal{T}_X$ explicitly.

Terminology A.1.2.

- A subset $U \subseteq X$ is called *closed* if $X \setminus U$ is open.
- If x is a point of X, an *open neighbourhood* of x is an open subset U of X such that $x \in U$. (A *neighbourhood* of x is a subset $S \subseteq X$ that contains an open neighbourhood of x.)

Examples A.1.3.

(i) The standard topology on $\mathbb{R}^n$ is defined by saying that a subset U of $\mathbb{R}^n$ is *open* if for every $x \in U$ there exists $\epsilon > 0$ such that when $|x - y| < \epsilon$ we have $y \in U$ (i.e. U contains the open ball of radius ϵ around x).

(ii) Similarly, if (X, d) is a metric space, a subset $U \subseteq X$ is *open* if for every $x \in U$ there exists $\epsilon > 0$ such that U contains the open ball of radius ϵ around x. This defines a topology on any metric space.

(iii) We can equip any set X with the *discrete topology* where all sets are open ($\mathcal{T}_X$ is the power set of X).

(iv) We can equip any set X with the *coarse* (or *indiscrete*) topology, where $\mathcal{T}_X := \{\emptyset, X\}$.

Definition A.1.4. Let X be a topological space and $Y \subseteq X$ any subset. Then the *subspace topology* on Y is given by

$$\mathcal{T}_Y := \{V \subseteq Y : V = U \cap Y \text{ for some } U \in \mathcal{T}_X\}.$$

Example A.1.5. The *n-sphere* S^n can be defined as the subset

$$S^n := \left\{ x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1 \right\}$$

equipped with the subspace topology from $\mathbb{R}^{n+1}$.

Exercise A.1. Let X be a topological space and $U \subseteq X$ a subset. Show that the subspace topology on U is characterized by the following property: if T is a topological space, then a continuous map from T to U is a map of sets $T \to U$ such that the composite $T \to U \hookrightarrow X$ is continuous.

Definition A.1.6. Let X and Y be topological spaces. A *continuous map* from X to Y is a function $f\colon X \to Y$ such that if $U \subseteq Y$ is open, then $f^{-1}U \subseteq X$ is also open.

Exercise A.2. Let (X, d_X) and (Y, d_Y) be metric spaces. Show that a function $f\colon X \to Y$ is continuous if and only if for every $x \in X$ and every $\epsilon > 0$ there exists $\delta > 0$ such that if $d_X(x, x') < \delta$ then $d_Y(f(x), f(x')) < \epsilon$. (Note that δ may depend on x.)

A.2 Groups

Definition A.2.1. A *group* is a set G equipped with a binary multiplication operation, that is a function[1] $- \cdot -\colon G \times G \to G$, and a unit element $1 \in G$ such that

- $\cdot$ is associative: for $a, b, c \in G$ we have $a(bc) = (ab)c$.
- 1 is a unit for $\cdot$: for $a \in G$ we have $a \cdot 1 = a = 1 \cdot a$.
- $\cdot$ has inverses: for $a \in G$ there exists $a^{-1} \in G$ such that $aa^{-1} = 1 = a^{-1}a$.

Definition A.2.2. If G is a group, then a *subgroup* of G is a subset $H \subseteq G$ that is closed under the group operations, i.e. $1 \in H$, if $a, b \in H$ then $ab \in H$, and if $a \in H$ then $a^{-1} \in H$. A subgroup $H \subseteq G$ is called *normal* if for $h \in H$ we have $ghg^{-1} \in H$ for all $g \in G$.

Definition A.2.3. If G and H are groups, then a *homomorphism* $\phi\colon G \to H$ is a function such that $\phi(ab) = \phi(a)\phi(b)$ for all $a, b \in G$ and $\phi(1) = 1$.

[1] We usually write the product $a \cdot b$ as just ab.

A.3 Tensor products

Definition A.3.1. Let $\mathbb{K}$ be a field (such as $\mathbb{R}$ or $\mathbb{C}$). If U, V, W are $\mathbb{K}$-vector spaces, a *($\mathbb{K}$)-bilinear map*

$$\phi \colon U \times V \to W$$

is a map of sets that satisfies

$$\phi(\lambda u + \mu u', v) = \lambda\phi(u, v) + \mu\phi(u', v), \quad \phi(u, \lambda v + \mu v') = \phi(u, \lambda v) + \phi(u, \mu v')$$

for all $u, u' \in U, v, v \in V, \lambda, \mu \in \mathbb{K}$. A *tensor product* of U and V is a $\mathbb{K}$-vector space $U \otimes_{\mathbb{K}} V$ (or just $U \otimes V$ if $\mathbb{K}$ is obvious from the context) together with a *universal* bilinear map $u \colon U \times V \to U \otimes V$ in the following sense: for every bilinear map $\phi \colon U \times V \to W$ there exists a *unique* linear map $f \colon U \otimes V \to W$ such that $\phi = f \circ u$, i.e.

$$\begin{array}{ccc} U \times V & \xrightarrow{u} & U \otimes V \\ & \searrow^{\phi} & \downarrow{\scriptstyle \exists!} \\ & & W. \end{array}$$

Exercise A.3. Show that if $u \colon U \times V \to W$ and $u' \colon U \times V \to W'$ are both tensor products of U and V, then there exists a unique isomorphism $w \colon W \xrightarrow{\sim} W'$ such that $w \circ u = u'$.

Proposition A.3.2. *For any $\mathbb{K}$-vector spaces U, V, their tensor product exists.*

Proof. Let $F = \mathbb{K}(U \times V)$ be the free vector space on the set $U \times V$, i.e. the set of formal $\mathbb{K}$-linear combinations of elements of $U \times V$ as a *set*. Then any map of sets $\phi \colon U \times V \to W$ where W is a $\mathbb{K}$-vector spaces extends uniquely to a $\mathbb{K}$-linear map $\phi' \colon F \to W$. Let R be the subspace of F spanned by $(\lambda u + \mu u', v) - (\lambda u, v) - (\mu u', v)$ and $(u, \lambda v + \mu v') - (u, \lambda v) - (u, \mu v')$ for all $u, u' \in U, v, v' \in V, \lambda, \mu \in \mathbb{K}$; then the map ϕ is bilinear if and only if R is contained in the kernel of ϕ'. Setting $T := F/R$, then by the universal property of the quotient we get a correspondence between bilinear maps $U \times V \to W$ and linear maps $T \to W$, given by composing with the composite map $u \colon U \times V \to F \to T$ where the first map is the inclusion of the generators and the second is the quotient map. This shows that (T, u) is a tensor product of U and V. □

Remark A.3.3. For $x \in U, y \in V$, we write $x \otimes y \in U \otimes V$ for $u(x, y)$. From the explicit construction we see that an element of $U \otimes V$ can be written as a finite sum $\sum u_i \otimes v_i$ of elements of this form. Since u is bilinear, we have

$$(u + u') \otimes v = u \otimes v + u' \otimes v,$$

$$u \otimes (v + v') = u \otimes v + u \otimes v',$$

$$(\lambda u) \otimes v = \lambda(u \otimes v) = u \otimes \lambda v, \ \lambda \in \mathbb{K}.$$

Warning A.3.4. A general element of $U \otimes V$ can *not* be written as $u \otimes v$ for $u \in U, v \in V$!

Lemma A.3.5. *Suppose $e_1, \dots, e_n$ is a basis for U and $f_1, \dots, f_m$ is a basis for V. Then $e_i \otimes f_j$ for $i = 1, \dots, n$, $j = 1, \dots, m$ is a basis for $U \otimes V$.*

Proof. From Remark A.3.3 we see that these vectors span $U \otimes V$. To see that they are linearly independent, suppose we have $\sum a_{ij} e_i \otimes f_j = 0$. For $k = 1, \dots, m$, define a bilinear map $\phi_k \colon U \times V \to U$ by $\phi_k(u, \sum b_j f_j) = b_k u$. If ϕ'_k is the induced linear map $U \otimes V \to U$, then

$$\phi'_k\left(\sum a_{ij} e_i \otimes f_j\right) = \sum a_{ik} e_i.$$

Since the e_i's are linearly independent, we must have $a_{ik} = 0$ for all i, and this goes for any k. □

Exercise A.4. Prove the following formal properties of tensor products using only the universal property:

(i) For $\mathbb{K}$-vector spaces U, V, there is a canonical isomorphism

$$U \otimes V \xrightarrow{\sim} V \otimes U$$

taking $u \otimes v$ to $v \otimes u$.

(ii) For $\mathbb{K}$-vector spaces U, V, W there is a canonical isomorphism

$$U \otimes (V \otimes W) \xrightarrow{\sim} (U \otimes V) \otimes W.$$

(iii) For $\mathbb{K}$-vector spaces U_i, $i \in I$ and V, there is a canonical isomorphism

$$\bigoplus_{i \in I} (U_i \otimes V) \xrightarrow{\sim} \left(\bigoplus_{i \in I} U_i\right) \otimes V.$$

(iv) For a $\mathbb{K}$-vector space V, there are canonical isomorphisms

$$\mathbb{K} \otimes V \xrightarrow{\sim} V,$$
$$0 \otimes V \xrightarrow{\sim} 0.$$

Definition A.3.6. Suppose $f \colon U \to U'$ and $g \colon V \to V'$ are $\mathbb{K}$-linear maps between $\mathbb{K}$-vector spaces. Then $f \otimes g \colon U \otimes V \to U' \otimes V'$ is the unique linear map arising from the bilinear map

$$U \times V \xrightarrow{f \times g} U' \times V' \to U' \otimes V',$$

so that $(f \otimes g)(u \otimes v) = f(u) \otimes g(v)$.

Exercise A.5. Prove the following formal properties of tensor products of linear maps, using only the universal property of tensor products:

(i) For $\mathbb{K}$-linear maps $f \colon U \to U'$, $f' \colon U' \to U''$, $g \colon V \to V'$, $g' \colon V' \to V'$", we have

$$(f' \otimes g') \circ (f \otimes g) = (f' \circ f) \otimes (g \circ g).$$

(ii) For $\mathbb{K}$-vector spaces U, U', V, V' and a linear map $f \colon U \to U'$ the tensor product $f \otimes 0$ with the zero map $0 \colon V \to V'$ is the zero map $0 \colon U \otimes V \to U' \otimes V'$.

(iii) For homomorphisms $f, g \colon U \to U'$, $h \colon V \to V'$, and $\lambda, \mu \in \mathbb{K}$, we have $(\lambda f + \mu g) \otimes h = \lambda(f \otimes h) + \mu(g \otimes h)$ as homomorphisms $U \otimes V \to U' \otimes V'$.

Bibliography

[1] William Fulton and Joe Harris, *Representation theory: A first course*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991. MR1153249

[2] Brian Hall, *Lie groups, Lie algebras, and representations*, 2nd ed., Graduate Texts in Mathematics, vol. 222, Springer, Cham, 2015. An elementary introduction. MR3331229

[3] James E. Humphreys, *Introduction to Lie algebras and representation theory*, Graduate Texts in Mathematics, Vol. 9, Springer-Verlag, New York-Berlin, 1972. MR0323842

[4] John Stillwell, *Naive Lie theory*, Undergraduate Texts in Mathematics, Springer, New York, 2008. MR2431402

[5] Qiaochu Yuan, *Annoying Precision*. Blog at https://qchu.wordpress.com.